DEPARTMENT OF EDUCATION

WORKING AT THE NEXUS OF COGNITIVE ACTIVATION AND DIFFERENTIATION: EXPLORING THE EFFECTIVENESS OF A PROFESSIONAL LEARNING AND DEVELOPMENT PROGRAM

DOCTOR OF PHILOSOPHY DISSERTATION

EVRIDIKI KASAPI

2024

University of Cyprus

DEPARTMENT OF EDUCATION

WORKING AT THE NEXUS OF COGNITIVE ACTIVATION AND DIFFERENTIATION: EXPLORING THE EFFECTIVENESS OF A PROFESSIONAL LEARNING AND DEVELOPMENT PROGRAM

EVRIDIKI KASAPI

A Dissertation Submitted to the University of Cyprus in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

May 2024

VALIDATION PAGE

Doctoral Candidate: Evridiki Kasapi

Doctoral Thesis Title: Working At the Nexus of Cognitive Activation and Differentiation: Exploring the Effectiveness of a Professional Learning and Development Program

The present Doctoral Dissertation was submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy at the ***Department of Education*** *and was approved on the 10th of May 2024 by the members of the* ***Examination Committee****.*

Examination Committee:

Research Supervisor: Charalambos Y. Charalambous, Associate Professor, Department of Education, University of Cyprus

____

Committee Member: Stavroula Philippou, Associate Professor, Department of Education, University of Cyprus (Chair of the Examination Committee)

____

Committee Member: Demetra Pitta-Pantazi, Professor, Department of Education, University of Cyprus

____

Committee Member: Despina Potari, Professor, Department of Mathematics, National and Kapodistrian University of Athens

____

Committee Member: Seán Delaney, Registrar and Vice President, Marino Institute of Education, Ireland

____

DECLARATION OF DOCTORAL CANDIDATE

The present doctoral dissertation was submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy of the University of Cyprus. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

Evridiki Kasapi

…………………………………..

ABSTRACT [in Greek]

Η παρούσα μελέτη επιδιώκει να συνεισφέρει σε δύο σημαντικά κενά στον τομέα της εκπαιδευτικής έρευνας. Το πρώτο αφορά τη σύζευξη της γνωστικής ενεργοποίησης—δηλαδή, την εμπλοκή των μαθητών/τριών σε σκέψη υψηλού επιπέδου μέσω γνωστικά απαιτητικών έργων—και της διαφοροποίησης—δηλαδή, της διαμόρφωσης της διδασκαλίας ώστε να συμβαδίζει με τις διάφορες ανάγκες των μαθητών/τριών. Αν και οι δύο αυτοί άξονες έχουν εξεταστεί χωριστά στο παρελθόν, η δυναμική συνεργασία τους δεν έχει ακόμη πλήρως εξερευνηθεί. Η συνένωση των δύο ερευνητικών αυτών πεδίων προτείνει ένα μοντέλο φιλόδοξης διδασκαλίας στα Μαθηματικά, με στόχο την εμπλοκή όλων των μαθητών/τριών σε γνωστικά απαιτητικά έργα. Ενώ υπάρχει ευρεία συναίνεση στον ερευνητικό χώρο ότι η φιλόδοξη διδασκαλία στα Μαθηματικά είναι επωφελής για τη μάθηση των μαθητών/τριών, παραμένει ανοιχτό το ζήτημα της υποστήριξης των εκπαιδευτικών στην αποτελεσματική εφαρμογή της.

Το δεύτερο ερευνητικό κενό που διερευνά η μελέτη σχετίζεται με την εμπειρική αξιολόγηση της επίδρασης των προγραμμάτων Επαγγελματικής Μάθησης και Ανάπτυξης (ΕΜΑ) των εκπαιδευτικών στην ενίσχυση των φιλόδοξων διδακτικών πρακτικών τους στα Μαθηματικά. Προηγούμενες μελέτες έχουν καταγράψει ανομοιογενή αποτελεσματικότητα των προγραμμάτων ΕΜΑ στην επίτευξη αλλαγής από τους/τις εκπαιδευτικούς, μελετώντας κυρίως αλλαγές στους/στις εκπαιδευτικούς ως ομάδα, κάτι που τονίζει την ανάγκη για πιο ολοκληρωμένες προσεγγίσεις μελέτης της αποτελεσματικότητας. Αυτές θα πρέπει να αξιολογούν τα αποτελέσματα των προγραμμάτων, αλλά και να αναλύουν τις διαδικασίες και τις προκλήσεις που επηρεάζουν την αλλαγή στη διδασκαλία.

Επικεντρωμένη σε αυτά τα κενά, η μελέτη εξετάζει την αποτελεσματικότητα ενός προγράμματος ΕΜΑ που εστιάζει στη γνωστική ενεργοποίηση, τη διαφοροποίηση και την αλληλεπίδρασή τους, χρησιμοποιώντας το μοντέλο αξιολόγησης του Kirkpatrick (2007). Αυτή η προσέγγιση διερευνά την αποτελεσματικότητα του προγράμματος σε διάφορα επίπεδα: τις αντιδράσεις των εκπαιδευτικών για το πρόγραμμα, τη διδακτική τους συμπεριφορά κατά τη διάρκεια του προγράμματος, τη μάθησή τους γύρω από τους άξονες, καθώς και τη διδακτική απόδοσή τους στο τελευταίο τους μάθημα.

Μέσω μιας διαδικασίας δειγματοληψίας δύο σταδίων, η μελέτη εστίασε αρχικά σε μια ομάδα οκτώ Κύπριων εκπαιδευτικών δημοτικής εκπαίδευσης που συμμετείχαν

σε μία Λέσχη Ανάλυσης Οπτικογραφημένων Διδασκαλιών (ΛΑΟΔ) ως σύνολο και στη συνέχεια, σε τρεις ετερογενείς μελέτες περίπτωσης για πιο εμβριθή ανάλυση των διαφορετικών διαδρομών μάθησης και βελτίωσης της διδασκαλίας τους. Μέσω της εφαρμογής ενός μοντέλου μικτών μεθόδων, η μελέτη συνδύασε ποσοτικές μεθόδους—με τη χρήση περιγραφικής και επαγωγικής στατιστικής—για τη μελέτη της διδακτικής απόδοσης και του πειραματισμού των εκπαιδευτικών, με ποιοτικές μεθόδους—με τη χρήση θεματικής ανάλυσης για τη μελέτη της κατανόησης των εκπαιδευτικών σε θέματα γνωστικής ενεργοποίησης, διαφοροποίησης και της σύζευξής τους, των αλλαγών στη διδακτική συμπεριφορά τους, των προκλήσεων που αντιμετώπιζαν και της στάσης τους απέναντι στο πρόγραμμα. Τα δεδομένα περιλάμβαναν οπτικογραφημένα μαθήματα, σχέδια μαθημάτων, συνεντεύξεις, οπτικογραφημένες ΛΑΟΔ και κάρτες αναστοχασμού των εκπαιδευτικών.

Αυτή η διττή ανάλυση αποκάλυψε τόσο τις συλλογικές τάσεις όσο και τις ατομικές διαδρομές επαγγελματικής ανάπτυξης μεταξύ των συμμετεχουσών εκπαιδευτικών. Στο επίπεδο των συλλογικών τάσεων, οι εκπαιδευτικοί εμφάνισαν γενικά θετική στάση προς το πρόγραμμα, εντοπίζοντας ταυτόχρονα περιοχές που χρήζουν βελτίωσης. Η αξιολόγηση της τελικής διδακτικής απόδοσής τους στα μαθήματα έδωσε μικτά αποτελέσματα, δυσχεραίνοντας τον καθορισμό μιας σαφούς πορείας αλλαγής. Ανάλογα μικτά αποτελέσματα παρατηρήθηκαν και κατά τη διάρκεια του προγράμματος ΕΜΑ στο επίπεδο της διδακτικής συμπεριφοράς. Ωστόσο, στατιστικά σημαντικές διαφορές εντοπίστηκαν μέσω του Wilcoxon signed-rank τεστ στην επίδοση των εκπαιδευτικών σε διάφορα χρονικά σημεία του προγράμματος, ιδίως όσον αφορά στην εξέταση της καλύτερής τους απόδοσης σε συγκεκριμένες πρακτικές διδασκαλίας.

Οι τρεις ενδελεχείς μελέτες περίπτωσης επισήμαναν διαφορετικά ατομικά μοτίβα αλλαγής στη μάθηση και στη συμπεριφορά. Η περίπτωση της Πίνας χαρακτηρίζεται από συνεχή βελτίωση, με έμφαση στη σύζευξη της γνωστικής ενεργοποίησης και της διαφοροποίησης. Η διαδρομή της Κέιτ υπογραμμίζει τις δυσκολίες στη διατήρηση συνέπειας και ποιότητας κατά την ενασχόληση με πολλαπλές διαφορετικές πρακτικές. Η αφήγηση της Μισέλ αναδεικνύει τις περιπλοκότητες της έννοιας της «φροντίδας» για τους μαθητές, επισημαίνοντας πώς η καλοπροαίρετη υποστήριξη του/της εκπαιδευτικού μπορεί να παρεμποδίσει ακούσια τη σκέψη των μαθητών/τριών.

Τα κύρια ευρήματα της μελέτης φώτισαν τρία σημαντικά θέματα σχετικά με την ΕΜΑ στη φιλόδοξη διδασκαλία στα Μαθηματικά. Αρχικά, αποκάλυψαν την πολυπλοκότητα της αποτελεσματικότητας, αναδεικνύοντας ότι η αποτελεσματικότητα δεν είναι μονοδιάστατη, αλλά αποτελεί ένα συνονθύλευμα επιπέδων, περιλαμβανομένων των αντιδράσεων των εκπαιδευτικών, των αλλαγών στην κατανόησή τους γύρω από τους άξονες του προγράμματος, των αλλαγών στη διδακτική συμπεριφορά τους, και των διδακτικών αποδόσεών τους στο τέλος του προγράμματος. Δεύτερο, φανέρωσε τις δυνατότητες και τις αδυναμίες του εξεταζόμενου προγράμματος ΕΜΑ, εμπλουτίζοντας τη βιβλιογραφία σχετικά με τα χαρακτηριστικά της αποτελεσματικής ΕΜΑ και ενημερώνοντας μελλοντικούς σχεδιασμούς και εφαρμογές ανάλογων προγραμμάτων. Τέλος, η μελέτη ανέδειξε τις διαφορετικές διαδρομές μάθησης και αλλαγής στη διδασκαλία που ακολούθησαν οι εκπαιδευτικοί κατά τη συμμετοχή τους στο πρόγραμμα.

Η μελέτη προσφέρει σημαντικές προεκτάσεις για τη φιλόδοξη διδασκαλία στα Μαθηματικά μέσω προγραμμάτων ΕΜΑ. Σε θεωρητικό επίπεδο, υπογραμμίζει την πολυπλοκότητα της αλλαγής των εκπαιδευτικών και τονίζει την ανάγκη για τη σύζευξη της γνωστικής ενεργοποίησης με τη διαφοροποίησης. Από μεθοδολογικής πλευράς, τονίζει τη σημασία της ολιστικής αξιολόγησης των προγραμμάτων ΕΜΑ, τις δυνατότητες των ΛΑΟΔ στην υποστήριξη της φιλόδοξης διδασκαλίας σε εκπαιδευτικούς δημοτικής εκπαίδευσης, και της χρήσης διάφορων μετρήσεων για την αξιολόγηση της διδακτικής απόδοσης, όπως μέσων όρων και μέγιστων βαθμολογιών. Πρακτικά, δείχνει τη δυνατότητα ενσωμάτωσης της γνωστικής ενεργοποίησης με τη διαφοροποίηση στη διδασκαλία, επισημαίνοντας τη σημασία των χαρακτηριστικών των προγραμμάτων ΕΜΑ που υποστηρίζουν αυτήν την ενσωμάτωση και την ανάγκη για προγράμματα ΕΜΑ που ανταποκρίνονται στις μοναδικές ανάγκες κάθε εκπαιδευτικού. Τα ευρήματα αυτά εμπλουτίζουν τη συζήτηση για την αποτελεσματικότητα των προγραμμάτων ΕΜΑ στη φιλόδοξη διδασκαλία στα Μαθηματικά και διανοίγουν νέους δρόμους για μελλοντικές έρευνες.

ABSTRACT [in English]

This study addresses two significant gaps in educational research. The first gap pertains to the interplay of *cognitive activation*—enhancing higher-level thinking through challenging mathematical tasks—and *differentiation*—customizing teaching to accommodate diverse student needs. Although these approaches have historically been explored separately, their synergistic potential is yet to be fully harnessed. The nexus of these research streams suggests a model of ambitious mathematics teaching, which aims to engage *all* students in challenging mathematical endeavors. While there is a broad consensus among researchers that ambitious mathematics teaching effectively promotes student learning, the question of how practicing teachers can be supported in adopting this approach remains unresolved.

Linked to this open issue, the second research gap this study explores relates to the empirical examination of the impact of Professional Learning and Development (PLD) programs in fostering ambitious teaching practices among teachers. Previous studies have shown inconsistent results regarding the effectiveness of PLD initiatives in driving teacher change, either by concentrating on collective transformations or analyzing selected teacher cases. This inconsistency highlights an urgent call for more comprehensive studies that not only evaluate the end results of such programs but also dissect the underlying processes and challenges influencing teacher change.

Building upon these gaps, the study scrutinizes the effectiveness of a PLD program focusing on cognitive activation, differentiation, and their interplay, utilizing Kirkpatrick's (2007) evaluation model. This model approach provided a detailed exploration of the program's effectiveness across multiple dimensions: teachers' reactions, their learning processes, behavioral improvements in teaching practices, and the final performance at the final timepoint of the PLD.

Employing a strategic two-stage sampling process, this study first focused on a cohort of eight Cypriot elementary school teachers engaged in the video-club program, subsequently narrowing down to three heterogeneous individual case studies for an in-depth examination of unique learning trajectories and teaching evolutions. Through the convergence model of mixed-methods triangulation design, the study combined quantitative assessments (using descriptive and inferential statistics) of teachers' teaching performance and experimentation, with qualitative evaluations (thematic analysis) of their conceptual development, behavioral changes, challenges, and

reactions to the program, drawing on a rich dataset of videotaped lessons, lesson plans, interviews, video-club sessions, and reflection cards.

This dual-level analysis illuminated both collective trends and individual pathways of professional growth among participating teachers. In terms of collective trends, teachers displayed a generally positive attitude towards the program while identifying areas for enhancement at the Reactions level. The evaluation of their performances in final lessons yielded mixed outcomes, rendering it challenging to delineate a clear pathway of change at the Results level. A mixed picture also emerged from the examination of their teaching performance throughout the PLD program at the Behavior level. Nevertheless, Wilcoxon signed-rank tests uncovered statistically significant differences in teacher performance across various program timepoints, particularly in their maximum performance scores.

The three case studies revealed varied individual patterns of change in learning and behavior at the Learning and Behavior level. Pina's case exemplifies continuous improvement, focusing on the synergistic interplay of cognitive activation and differentiation. Kate's journey underscores the difficulties in sustaining consistency and quality while juggling multiple different practices. Michelle's narrative, on the other hand, illuminates the intricacies of caring for students, illustrating how well-intentioned teacher support can inadvertently impede student thinking.

The main findings of the study illuminated three pivotal topics in PLD for ambitious mathematics teaching. Firstly, they unpacked the concept of effectiveness within teacher PLD, revealing that effectiveness is not a monolithic outcome but a complex interplay of factors, including teachers' reactions, their evolving understanding, behavioral changes, and the end results of the PLD program on teaching practices. Secondly, they highlighted the strengths and limitations of the PLD program under consideration, enriching the literature on the features of effective PLD, crucial for informing future designs and implementations. Lastly, the study underscored the diverse learning and behavioral change paths traversed by teachers participating in the PLD program.

This study offers insightful implications for ambitious mathematics teaching through PLD. Theoretically, it highlights the complexity of teacher change, advocating for an integrated approach to cognitive activation and differentiation. Methodologically, it values holistic evaluations of PLD programs, emphasizing the effectiveness of video clubs and advocating for mixed metrics (both mean and maxima) to assess teacher performance. Practically, it demonstrates the viability of merging cognitive activation with differentiation in teaching, underlining the importance of PLD features that support this integration and the need for PLD programs to cater to individual teacher needs. These findings contribute to the discussion on the effectiveness of PLD aimed at ambitious mathematics teaching while opening the door to future research.

ACKNOWLEDGMENTS

Completing my PhD dissertation has been a long and challenging journey, one that I could not have navigated without the guidance, support, and encouragement of a constellation of individuals. This dissertation, the culmination of years of effort, would not have come to fruition without their valuable and collective contributions.

Foremost, I am profoundly grateful to my advisor, Associate Professor Charalambos Y. Charalambous, for his tireless support, immense knowledge, and insightful guidance, which have been paramount in navigating the intricate pathways of research. His tremendous patience in meticulously reading hundreds of pages of my writing, engaging in extensive discussions, providing constructive feedback, offering challenging yet enlightening comments, and affording me the opportunity to participate in academic conferences, has been invaluable. Moreover, entrusting me with the role of Assistant Coordinator for the project, from which this study's data were derived, signifies a level of trust and respect that I deeply appreciate. His unwavering dedication and mentorship have not only been instrumental in shaping this dissertation but have also played a significant role in my personal and professional development. I could not have envisioned a better advisor and mentor for my PhD journey!

Besides my advisor, I would like to express my heartfelt appreciation to the rest of the internal members of my Examination Committee: Associate Professor Stavroula Philippou (the Chair) and Professor Demetra Pitta-Pantazi. Their expertise and constructive suggestions have immensely enriched the quality of this dissertation. My sincere thanks also go to the two external members of my Examination Committee, Professor Despina Potari, and Registrar and Vice President Seán Delaney, for their expert evaluation and critical perspectives that significantly contributed to the depth and breadth of my study.

Special acknowledgment is due to the European Commission for funding the program “Enhancing Differentiated Instruction and Cognitive Activation in Mathematics Lessons by Supporting Teacher Learning (EDUCATE)” from which the study drew its data. My gratitude extends to the collaborative efforts of the consortium of the European ERASMUS+ Project EDUCATE, comprising seven organizations across four partner European countries. Many thanks to the individuals within these organizations who have contributed their time, expertise, and resources to collect quality data and meet all the goals of this project.

Special thanks to the Ministry of Education, Sport, and Youth of Cyprus and the Centre of Educational Research and Evaluation for granting permission to conduct this study. Their support was crucial in facilitating the research process.

This journey could not have been undertaken without the participating teachers who volunteered to participate in the professional learning and development program of the project, and who generously opened their classrooms for videotaping, offering a direct glimpse into their teaching. Equally, the students deserve thanks for their participation and time while videotaping their teacher's teaching mathematics. I also thank the school principals for their consent, enabling this research to proceed.

My journey was also enriched by my fellow students—Dr Thekla Iakovou, Sergios Sergiou, Kassandra Georgiou, Yiola Ioannou, Angeliki Makri, Nicolas Antoniades, and Maria Perikli—whose camaraderie and shared insights were a source of inspiration and encouragement, while our inside jokes were sources of joy and laughter. I am confident that this is not where our shared journey ends!

Heartfelt thanks to my parents, Michalis and Eleni, and sisters, Diamanto and Lydia, whose unwavering support and understanding have been my backbone. Their belief in me and my endeavors provided me with the strength to fulfil my dreams and persevere through challenges. Many thanks for the meals, the humorous videos shared on social media, the planning of outings, and all the love you have shown.

Special thanks to my friends, who have mastered the art of simultaneously being the reason for my procrastination and the ones who dragged me across the finish line. Your ability to turn every stressful moment into a laugh has been the secret ingredient to surviving this PhD journey!

Finally, I reserve special gratitude for my future husband, Michalis Machallekides, for his tireless support and courage, especially during moments when I doubted my own. His love and patience were the lights guiding me through this journey. As I always jokingly remind him: "It's not Miss, Ms, or Mrs... it's Dr, actually!"

To all mentioned and unmentioned, who contributed to this journey, I am eternally grateful.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

AMT	Ambitious Mathematics Teaching
CA	Cognitive Activation
DIF	Differentiation
EPI	End-of-program Interview
INT	Interplay of Cognitive Activation and Differentiation
L	Lesson
MKT	Mathematical Knowledge for Teaching
MTF	Mathematical Task Framework
PCK	Pedagogical Content Knowledge
PLD	Professional Learning and Development
POI	Post-Lesson Interview
PRI	Pre-Lesson Interview
SAW	Student Autonomous Work
T	Timepoint
TAG	Task Analysis Guide
TL	Task Launching
TRC	Teacher Reflection Cards
VCS	Video-Club Session
WCI	Whole-Class Interactions

CHAPTER 1. THE PROBLEM

Calls for mathematics reform prioritize enhancing teachers' capacity to respond to the needs of a diverse student population, alongside engaging all students in cognitively activating work—thus, making their mathematics lessons both challenging and engaging. Concurrently attending to both cognitive activation and differentiation, that is teaching ambitiously, is a way to promote this dual goal. Yet, these two teaching goals have largely been studied separately. Carefully designed professional learning development (PLD) programs can support teachers in learning to enact cognitively activating and differentiated instruction. However, the effectiveness of PLD programs targeting ambitious teaching on teacher learning and practice remains inconsistent, with outcomes ranging from positive shifts to negligible impacts. Utilizing Kirkpatrick's model, this research zeroes in on teachers' end results and their developmental journey throughout a PLD on ambitious teaching. The study aims to explore the practices with which teachers experiment and the challenges they confront during such a PLD program. It also investigates potential changes in their conceptualizations and teaching practice concerning cognitive activation, differentiation, and their interplay. The findings of this study can inform the design of future PLD initiatives and advance our understanding of ambitious teaching, crucial for enhancing mathematics teaching and student outcomes in diverse classrooms.

1.1 Introduction

Contemporary classrooms and educational systems are characterized by a multitude of complexities. On the one hand, most educational systems worldwide emphasize the importance of developing students' subject-matter proficiency, involving the mastery of challenging concepts and skills across all grade levels. Teachers are facing increasing pressure to teach ambitiously across all academic subjects, particularly the subjects of language arts or mathematics, which are usually at the forefront of many educational policies (Anagnostopoulos et al., 2020).

At the same time, teachers are called upon to attend to the student diversity within their classrooms. International assessments consistently highlight disparities in student performance across diverse racial, cultural, and socioeconomic backgrounds, as evident in long-term trends seen in Grade 4 and Grade 8 TIMSS results in mathematics (Mullis et al., 2019). A recent OECD report (Organisation for Economic Co-operation and Development, [OECD], 2019) indicates that, on average, only 2.4% of students achieve a minimum level of proficiency in mathematics by the age of 15, demonstrating their ability to conceptualize, generalize, investigate, and model complex problem situations in relatively unfamiliar contexts. Conversely, in one-third of the participating countries, over 50% of students score at a low proficiency level, capable only of identifying information and executing routine procedures in familiar situations. Furthermore, significant disparities persist in the performance of students from disadvantaged backgrounds, both within and across countries. Notably, socio-economically advantaged students score 180 points higher in mathematics, equivalent to nearly four additional years of schooling, compared to their less-advantaged peers (Schleicher, 2019).

In our ever-evolving societies, it has become imperative that *all students*—despite their backgrounds—acquire the knowledge and skills necessary for full participation in contemporary economies and societies (Lampert et al., 2010; OECD, 2023). The key to these complexities that teachers are dealing with may lie at the crossroads of two research strands that have largely progressed independently, much like two parallel roads: *cognitive activation* (e.g., Doyle, 1983; Tekkumru-Kisa et al., 2020; Stein et al., 2009) and *differentiation* (e.g., Guild & Garger, 1998; Taylor, 2015; Tomlinson, 2001; Van Geel et al, 2019). Although a considerable amount of literature has been published on either cognitive activation or differentiation, studies focusing on both axes have been relatively scant, and more recent (e.g., Charalambous et al., 2023a; 2023b; Delaney & Gurhy, 2019; Mellroth et al., 2021; Psycharis et al., 2019; Sullivan et al., 2016a).

The current study attempts to bring these two strands together. Although doing so might be challenging[1], it can offer significant benefits both for research and practice. First, merging these strands can lead to a more comprehensive understanding of effective teaching practices. Approaching these two axes more synergistically can enable scholars "to see the complementarities of their work and realize how these complementarities can help better capture teaching quality and its effects on student learning" (Charalambous & Praetorius, 2020, p.3). Differentiation addresses the diverse needs of students, while cognitive activation, through the use of challenging tasks, can support higher-order thinking and problem-solving skills. Together, these two strands provide a holistic perspective on how to engage and support all students effectively. Second, teachers can benefit from a more integrated approach to teaching, equipping them with a wider range of practices and tools to meet their students'

[1] Integrating the research strands of cognitive activation and differentiation might be challenging due to various reasons (e.g., each is rooted in distinct theoretical frameworks; the different terminologies used by each strand can lead to conceptual differences and/or create an illusion of discussing entirely different ideas, even if they overlap).

diverse needs, all while reducing potential conflicts between addressing those needs and challenging students mathematically. Lastly, research has already demonstrated that cognitively activating teaching, when combined with *appropriate adaptations* of mathematically challenging tasks to scaffold different student groups, is a key condition for addressing the learning needs of all students, resulting in higher levels of understanding and equitable outcomes for diverse students (Boaler & Staples, 2008; Choppin, 2011; Tomlinson, 2015).

In this respect, concurrently attending to cognitive activation and differentiation is a pioneering way to work towards *ambitious teaching,* which aims *at* concurrently attending to both issues of teaching quality and equity (Waddell, 2014). Ambitious teaching is defined as teaching responsive to what students do and need while immersing them in intellectually, authentic, and challenging educational experiences (e.g., Cohen, 2011; Lampert et al., 2010; McDonald et al., 2013; Newmann & Wehlage, 1996). Both earlier and recent empirical research points to the critical nature of this kind of teaching in supporting student learning and understanding, transcending factors such as teaching and learning standards, curriculum, school type, student age, or other contextual variables (e.g., Boaler & Staples, 2008; Silver & Kenney, 2016; Silver & Stein, 1996; Stein et al., 2007). Engaging students in challenging tasks that foster meaning-making, reasoning, and problem-solving, along with differentiated instruction to ensure equitable access to a high-quality curriculum, constitute essential components of effective mathematics teaching and learning (Spangler & Wanko, 2017). The complexity of teaching ambitiously in a classroom full of diversity and the scarcity of research bringing together the two independently developed research strands (i.e., cognitive activation and differentiation) underscores the need for further research in this area.

1.2 Background

A substantial body of research has focused on understanding what knowledge and skills teachers need to develop to effectively implement ambitious teaching (e.g., Boston, 2013; Franke et al., 2007; Kazemi et al., 2009; Lampert et al., 2013; Smith & Stein, 2011; Sullivan et al., 2015). However, the development of ambitious teaching practices is complex and demanding for teachers who grapple with a shortage of understanding and proficiency in navigating this complex terrain (Wager et al., 2017). Jackson and Cobb (2010) argued that teachers require high-quality PLD “informed by an ambitious and equitable vision of mathematics instruction if they are to cultivate the recommended instructional practices" (p. 29).

Indeed, over the past decade, efforts to *improve ambitious teaching in mathematics*—the subject matter of this study—have often relied on teacher PLD to achieve this objective. However, the number of PLD programs focusing on ambitious mathematics teaching is scarce, with limited attention to practicing teachers and their teaching practice (cf. Santagata et al., 2021). Even more critically, evaluation studies examining the effectiveness of such PLD programs have certain limitations, as discussed below. As the landscape of ambitious mathematics teaching continues to evolve, understanding the impact of these programs on both teachers and students becomes increasingly crucial for informing future educational endeavors.

Quantitative PLD program evaluation studies have primarily determined program success by assessing its *final outcomes*, such as improvements in teacher and/or student performance (e.g., Dash et al., 2012; Hill et al., 2018; Jacob et al., 2017; Kraft et al., 2020; Lindvall et al., 2022; 2023; Sun et al., 2014). These quantitative investigations reported varying levels of program effectiveness, with some showing only moderate effects (e.g., Jacob et al., 2017), others indicating null effects (e.g., Hill et al., 2018; Lindvall et al., 2023), and many demonstrating mixed effects (e.g., Dash et al., 2012; Kraft et al., 2020; Lindvall et al., 2022; Sun et al., 2014). A recent quantitative meta-analysis of studies focusing on practicing teachers teaching from kindergarten to Grade 12 (Garrett et al., 2019) underscored the substantial *heterogeneity* in program effects, despite some improvements identified in various dimensions of teaching practice through practice-oriented interventions.

On the contrary, qualitative PLD program evaluation studies have often centered on more nuanced aspects of program effectiveness. Specifically, they have provided insights into *collective* changes in teacher noticing regarding ambitious mathematics teaching practices (e.g., Jakopovic, 2021) and in teacher learning of these practices, through detailed discourse analysis (e.g., Wæge & Fauskanger, 2021; 2023; Gibbons & Okun, 2023). Other studies investigated *collective* teacher learning, focusing on aspects of lesson planning, rehearsal, enactment, and discussion after enactment (e.g., Fauskanger & Bjuland, 2019; Gibbons et al., 2017). These studies consistently reported positive results in these domains, showcasing the potential of PLD programs to bring about meaningful changes in teaching and learning.

Despite their considerable merit, these studies primarily spotlighted collective changes and outcomes, leaving a gap in the comprehensive exploration of individual teacher learning and practice, which is crucial for understanding the adoption and implementation of new teaching practices. The disparities stemming from individual teachers' backgrounds, experiences, and conceptions can notably influence the effectiveness of PLD initiatives. Furthermore, these studies did not quantify or compare any changes in teacher learning, an impediment that could be mitigated by incorporating quantitative methodologies. Lastly, there was a lack of focus on the challenges faced by teachers in implementing ambitious teaching practices learned through PLD programs. Identifying and understanding the types of potential challenges, such as time constraints, lack of resources, or resistance to change, is essential. This understanding allows for adapting PLD programs to different contexts and developing tailored interventions and strategies to address the unique needs and constraints of different teacher participants.

Among the qualitative studies, two case studies examined changes in individual teachers' practices: the first observed the use of tools for ambitious mathematics teaching across four cases but lacked clarity in depicting specific changes (Charalambous et al., 2023a). The second study presented a teacher's experience in a PLD program, indicating minimal change in her teaching approach (Anthony et al., 2018). Both studies highlighted challenges in enacting ambitious teaching, emphasizing the need for clearer articulation of change and deeper exploration of teaching complexities.

Mixed-methods studies considering final outcomes typically reported positive gains in student learning from pre-test to post-test and an overall positive experience with the teaching treatment (e.g., Shumway et al., 2020). They also reported greater gains in ambitious mathematics teaching for coach-teacher pairs discussing practice implementation timing and rationale, offering teacher input, and considering conditions for specific teaching moves (Witherspoon et al., 2021). These studies, however, did not systematically monitor teachers' teaching practice through lesson observations.

It is evident from these findings that *there remains considerable variability in the effectiveness of PLD programs aiming to equip teachers with the knowledge and skills necessary for ambitious mathematics teaching*. The mixed results from quantitative studies suggest that while PLD programs may lead to improvements in some areas, they may not consistently yield significant changes across the board. Based on the results of these studies, one could argue that the programs are "unsuccessful." However, Kennedy (2016) urges us to explore such discouraging results and reconsider the concept of "failure" for a PLD program, questioning whether null or mixed effects are indeed indicative of failure. On the other side, qualitative and mixed-methods studies offer valuable insights into the complexities of teacher learning and the challenges they face during the implementation of ambitious mathematics teaching practices, but, like the quantitative studies, are also limited in other respects. Overall, then, it appears that the effectiveness of PLD programs demands a more comprehensive investigation that attends to the specific context, goals, and outcomes being measured.

This diversity in outcomes underscores the complexity of designing, implementing, and evaluating effective PLD initiatives, emphasizing the need for further research to develop a more nuanced understanding of what teachers ultimately gain from such programs. This complexity is aggravated by taking into consideration that the programs reviewed above have been developed based on key quality characteristics proposed to render PLD programs successful in improving teacher learning and practice, and ultimately, student outcomes (cf. Desimone, 2011). In this context, two lingering questions arise: What do teachers gain from such PLD programs? How similar/different are these gains for participating teachers and why?

A widely used model for evaluating the effectiveness of PLD programs is Kirkpatrick's model/framework (Kirkpatrick & Kirkpatrick, 2007). Kirkpatrick's framework seems to integrate both quantitative and qualitative research approaches, encompassing four levels[2] of evaluation. These levels, detailed further in Section 2.7, can facilitate a structured understanding and analysis of a PLD program's effectiveness. Briefly, they explore teacher participants' *reactions* to the PLD program; the extent and depth of their *learning;* changes in their *behavior*; and the *final results* achieved. The first three levels align with the focus of qualitative studies, while the last level was mostly addressed by quantitative investigations. The four levels of Kirkpatrick's model provided both a framework for exploring the effectiveness of a PLD program on ambitious mathematics teaching as well as to organize the research questions of this program evaluation study.

---

[2] The four levels have been referred to as stages, criteria, types, steps, and categories of measures (Reio et al., 2017).

1.3 Statement of the Problem

While there is a growing emphasis on ambitious teaching practices to cater to the diverse needs of students in mixed-ability classrooms, the effectiveness of PLD programs designed to equip teachers with the necessary skills remains inconsistent. Quantitative studies present mixed results, and while qualitative research offers deeper insights into teacher experiences, there is a notable gap in understanding individual teacher learning and its translation into practice. Furthermore, the integration of cognitive activation and differentiation, two pivotal research strands, has been scarcely explored. The challenge lies not only in understanding the effectiveness of these programs collectively but also in discerning the different shades in the experiences, challenges, and learning trajectories of individual teachers. This study sought to bridge these gaps by employing Kirkpatrick's model to holistically evaluate the impact of a PLD program on ambitious mathematics teaching, focusing on both collective and individual teacher outcomes.

1.4 Purpose of the Study

The purpose of this study was to examine the effectiveness of a PLD program on teachers' conceptualization and implementation of cognitive activation, differentiation, and the interplay between the two. Specifically, the study aimed to assess the teachers' performance in their culminating videotaped lessons and identify the practices with which they experimented (i.e., the Results). Furthermore, the study aimed to trace the evolution of teachers' practices and conceptualizations throughout their participation in the PLD program, identifying any challenges they encounter and comparing their evolving understandings with their actual teaching practices (i.e., the Learning and Behavior Processes). Also, the study focused on exploring teachers' perceptions of the PLD program (i.e., their Reactions). Through this investigation, it was explored if and how an extensive PLD program, sharing many characteristics of effective PLD (see Section 2.5.2) and capitalizing on the video-clubs approach (see Section 2.6), can potentially create a platform to help teachers reconceptualize the two axes and their interplay and introduce changes in their teaching.

1.5 Research Questions

In this context, this study delved into the following—previously formulated—query: *What do teachers gain from participating in a PLD program focusing on aspects of cognitive activation, differentiation, and the interplay between the two?* This overarching query was split into four research questions, corresponding to the four levels of the Kirkpatrick model.

Regarding the *Results*, this study asked:

1. How does teachers' performance look at the final timepoint?
   1.1. What teaching practices do teachers experiment with in their concluding lessons?
   1.2. What is the teachers' performance in their concluding lessons?

In terms of *Teachers' Behavior*, the study examined:

2. How does their teaching behavior evolve over time?
   2.1. What teaching practices do teachers experiment with most frequently?
   2.2. What changes do teachers introduce in their practice, during their participation in the PLD program?
   2.3. What challenges do teachers encounter during this process?

With respect to *Teachers' Learning*, the study investigated:

3. How does teachers' learning evolve over time?
   3.1. How do teachers (re)conceptualize cognitive activation, differentiation, and their interplay throughout their participation in a relevant PLD program?
   3.2. How do teachers' (re)conceptualizations and practices compare?

Concerning *Teachers' Reactions*, the study wondered:

4. What reactions do teachers have regarding the PLD program?

1.6 Significance of the Study

This study has theoretical, methodological, and practical implications. In terms of *theory*, by merging the strands of cognitive activation and differentiation, this study can make an important contribution to the field of research on ambitious mathematics teaching. The lack of studies that concurrently attend to cognitive activation and differentiation, let alone discuss teachers' experimentation, and challenges around both issues and how they conceptualize and enact both axes, stresses the need for conducting further research toward this end. By analyzing and describing teachers' (re)conceptualizations and the changes they introduce in their practice, we can identify several areas and dimensions that warrant attention and development.

Additionally, by focusing on teacher experiences and learning trajectories, this study can offer a deeper insight into the complexities of teacher learning and the impact of PLD on teachers' practice, enriching our theoretical understanding of teacher change. Focusing on both the end results and the learning and behavioral processes of teachers, we acknowledge that teachers are learners themselves and that teacher change may not be a linear process but one that is influenced by various personal, contextual, and systemic factors (Clark & Hollingsworth, 2002). A sole focus on end results would overlook the significance of "illuminating the black box of teachers' learning […] to develop more general understandings about how certain catalysts for change affect the pathways of teachers' learning" (Goldsmith et al., 2014, p.24). This quote underscores the importance of process-oriented research in education, emphasizing that merely examining the end results or outcomes of teacher learning is insufficient. In essence, it advocates for a more comprehensive approach to educational research, one that values the journey as much as the destination.

From a *methodological* standpoint, this study offers two significant contributions to the evaluation of PLD programs. The first pertains to the potential added value of collecting data not only at the beginning and the end of the PLD program. By leveraging the Kirkpatrick model, the study goes beyond traditional outcome-based evaluations, by collecting data (both quantitative and qualitative) throughout the implementation of the program. The second notable methodological contribution of this study includes the exploration of the potential role of a particular form of PLD, namely video-clubs (see Section 2.6), in supporting elementary practicing teachers in this area. Remarkably, no previous video-club study has focused on the conceptual and teaching changes of elementary practicing teachers concerning cognitive activation, differentiation, and their interplay. Given the inherent complexities of teaching for both cognitive activation and differentiation, the video-club approach holds promise, offering a platform for teachers to examine their teaching gradually and critically, and experiment with different ways to improve it.

Besides, the study's emphasis on different teacher trajectories and the challenges they face can pave the way for more personalized and adaptive research methodologies. Instead of treating teachers as a homogenous group, future research can adopt methodologies that recognize and account for individual differences, yielding richer and more detailed findings.

On the *practical* front, investigating any potential changes in teachers' teaching practice within a PLD program can help teacher educators and researchers develop several insights regarding the materialization of cognitive activation and differentiation. First, teacher educators can understand what is feasible and realistic to expect from elementary practicing teachers in terms of changes in teachers' teaching quality as a result of their participation in the program. Also, the findings of this study can inform the design and implementation of future PLD programs. If certain challenges are consistently faced by teachers, PLD initiatives can be tailored to address these challenges directly, ensuring that teachers are better prepared to adopt ambitious teaching practices in their classrooms. This is critical to advancing our understanding of the demands of such work and developing more effective PLD opportunities for teachers (Boaler, 2002).

In addition, PLD programs should be designed "based on a more nuanced understanding of what teachers do, what motivates them, and how they learn and grow." (Kennedy 2016, p. 974). As Kennedy (2019) suggests, this is "an area in which we need to learn more about how to learn about teacher learning, how to design our studies, and how to map exposure to PLD with changes in practice" (p.153). This study offers a perspective to researchers on how PLD can support teachers to concurrently attend to both issues. Positive shifts in practice could be partly explained by the self-reported opportunities that teachers had while addressing both goals and certain PLD program characteristics. Also, any progress can possibly suggest that both goals are learnable in the context of high-quality teacher PLD. Conversely, if negative or no change is observed, it might imply that effective teaching for both cognitive activation and differentiation cannot simply emanate from a PLD program (Pfister et al., 2015) or that the challenges faced, or other contextual factors may prevent teachers from implementing such teaching in their classes. To better understand the variability of the effectiveness of PLD programs on ambitious teaching researchers need to develop their understanding of teacher learning and how to support them in integrating new ideas into their so-called systems of practice (Kennedy, 2016).

1.7 Outline of the Study

The rest of the thesis is composed of eight chapters. Chapter 2, the literature review, begins by establishing the theoretical foundations of *Cognitive Activation*, discussing its critical role in student engagement, and exploring task design, implementation challenges, and the support offered by PLD programs. The chapter then transitions to defining *Differentiation*, addressing related misconceptions, and highlighting its indispensability in catering to diverse student needs alongside the facilitators and barriers teachers encounter. The discussion extends to strategies for maintaining task complexity in differentiated teaching with a focus on the concept of *ambitious mathematics teaching*, underscoring the importance of resources, teacher actions, and existing research gaps. The chapter culminates with an in-depth examination of *PLD programs*, detailing their effective characteristics, and outlining *Kirkpatrick's model* as a framework for evaluating their effectiveness. Within this exploration, the chapter also examines *Video Clubs* as a form of PLD, analyzing their contribution to teacher learning and research gaps.

Chapter 3 outlines the methodology of this study on evaluating the effectiveness of the EDUCATE PLD video-club program, focusing on cognitive activation, differentiation, and their interplay, for elementary school teachers in Cyprus. It details the mixed-methods triangulation design pursued to integrate quantitative evaluations of teaching performance with qualitative analyses of teachers' teaching behaviors, conceptual development, and reactions. The chapter describes the EDUCATE project background, the participant sampling, the intervention process, and the data collection methods. It then explains the data analysis procedures, leveraging Kirkpatrick's model for assessing outcomes at various levels. Finally, it discusses the methodological safeguards to enhance study trustworthiness while acknowledging its limitations.

Chapter 4 delves into the *group-level* effectiveness of the PLD program through the lens of Kirkpatrick's model, focusing on teachers' Reactions, Results, and Behavior. It includes a thematic analysis of teachers' perceptions and feedback on the program; quantitative evaluations of their teaching performance across various teaching practices in their culminating lesson; and an examination of how their teaching behaviors evolved across teachers' lessons over time. The chapter navigates through the complexities of implementing cognitive activation, differentiation, and their interplay, offering insights into the areas of strength and opportunities for further PLD.

Chapters 5 to 7 each offer an in-depth exploration of one of the three selected *case studies*—Pina, Kate, and Michelle—delineating the PLD effectiveness through their individual trajectories within the program. Despite focusing on different individuals, each chapter is meticulously structured in the same manner for comparative coherence. It begins with the teacher's background, followed by their evaluation of the program (Reactions), and details their teaching performance in the concluding lesson (Results). This progression continues with an analysis of the evolution of their teaching performance, dissected through both quantitative and qualitative lenses (Behavior), and delves into the Learning level by examining the evolution of each teacher's conceptualizations around cognitive activation, differentiation, and their interplay, from initial, through evolved, to final conceptualizations. Each chapter wraps up by comparing the teacher's conceptual evolution and teaching practice, offering a comprehensive view of their development.

Chapter 8 concludes the study and discusses its main findings. The chapter begins with an introduction that sets the stage for the analysis of the main findings. Rather than isolating the findings related to each research question, the chapter synthesizes the outcomes around three pivotal topics, offering a collective examination of the four research questions posed. Through this synthesis, the chapter articulates implications for theory, methodology, and educational practice. The chapter culminates with the limitations of the study and outlines directions for future research, thereby paving the way for continued advancement in the field of teacher PLD.

CHAPTER 2. REVIEW OF THE LITERATURE

Chapter 2 delves into the complexities of teacher PLD, focusing on cognitive activation and differentiation within the context of mathematics education. It outlines the challenges that teachers face in implementing ambitious teaching practices while catering to the diverse needs of students. Additionally, it discusses the role of PLD in supporting these efforts. The chapter significantly contributes to identifying current research gaps by exploring the effectiveness of various PLD programs in focusing on ambitious teaching, revealing inconsistent contributions to teacher learning and practice. The chapter underscores the necessity for more detailed research on both individual and group learning in PLD and emphasizes the importance of well-designed PLD programs for enhancing teaching quality. Current research often overlooks the depth and complexity of teacher learning and practice, indicating a need for mixed-method research approaches. Literature advocates for studies that examine the co-evolution of teachers' conceptualizations and teaching practices through PLD. Recognizing the differential impact that PLD could have on teachers, this study explores both teacher collective learning and how different teachers progress through cycles of learning and experimentation.

2.1 Introduction

Chapter 2 aims to unravel the intricate fabric of teacher PLD, with a focus on cognitive activation and differentiation, as well as their interplay. This exploration builds upon the foundational insights established in the preceding chapter, which portrayed the complex landscape of contemporary educational systems. It emphasized the dual challenge teachers face when delivering ambitious mathematics teaching while catering to the diverse needs of students. The first chapter highlighted a crucial, yet underexplored, synergy between two pivotal research domains—cognitive activation and differentiation. Individually, these strands have significantly enriched educational research; however, their intersection presents an uncharted territory ripe for investigation. The chapter concluded by emphasizing the need for PLD programs that effectively support teachers in implementing ambitious mathematics teaching practices while acknowledging the variability in their effectiveness and the gap in a comprehensive understanding of their impact.

This initial exploration sets the stage for a thorough examination in Chapter 2 of the existing literature related to the six key sections of this study, each vital to understanding the overarching subject. Specifically, the first section, 'Cognitive Activation,' begins by defining the concept and then delves into its significance in engaging students with challenging work. This is followed by an examination of the critical aspects of task design, selection, and implementation, fundamental in driving cognitive activation. The section also explores the facilitators and barriers faced by teachers in maintaining cognitive activation, concluding with the role of PLD in supporting these efforts.

The second major section, 'Differentiation,' starts by defining the concept and arguing about its essential role in teaching. It addresses common misconceptions, underlining the necessity of differentiation in teaching. This section then identifies factors that support or inhibit teachers in this endeavor and discusses the importance of PLD.

The third section, 'Differentiating Teaching without Reducing Task Complexity,' shifts focus to the concept of ambitious mathematics teaching. It discusses resources and teacher actions that facilitate such teaching, outlines the challenges faced by the teachers, and reviews the supporting research, highlighting unresolved issues.

The review culminates with an in-depth exploration of 'Teacher Professional Learning and Development.' This section defines the concept and outlines key features of effective PLD programs related to the individual teacher-learner, the learning group, and the program itself. It also examines 'Video Clubs' as a model of PLD, discussing their structure, goals, effectiveness, and research evidence. The literature review concludes with an evaluation of these programs using Kirkpatrick's model (2007), evaluating the PLD program from different perspectives.

This chapter seeks to map out the current state of knowledge in these areas, identify gaps or emerging trends in literature, and understand the challenges and successes documented in implementing ambitious mathematics teaching practices and PLD programs. By doing so, this chapter aims to provide a theoretical background against which the findings of this study can be contextualized and understood, and the methodological choices made in subsequent chapters can be justified (see Chapter 3).

2.2 Cognitive Activation

2.2.1 Definition of Cognitive Activation

Cognitive activation is a complex, multidimensional concept that is associated with students' opportunities to learn and develop a deep conceptual understanding of an idea or a subject (Praetorius et al., 2014; Schlesinger & Jentsch, 2016). It encompasses the required type and level of thinking and the *cognitive processes* utilized by the students as they interact with the teaching content (Hamre et al., 2007; Stein et al., 2009). Educational psychologists categorize cognitive processes into *basic* *mental processes*, such as memory recall, routine problem-solving, attention, and perception, and *higher mental processes*, such as reasoning, complex problem-solving, and abstract thinking (Smith & Kelly, 2015). Basic cognitive processes occur in an automated manner with lower level of consciousness, such as receiving, interpreting, and responding to different stimuli, or encoding, storing, and retrieving information. Higher cognitive processes involve extensive mental effort, including thinking more deeply about the teaching content; solving non-routine problems, using complex and creative thinking strategies; and reasoning and reflecting on learning (Boston & Smith, 2009; Henningsen & Stein, 1997; Stein et al., 2009). Both high and low cognitive processes are essential for learning and functioning effectively.

Although cognitive activation has been extensively investigated in various academic disciplines, including literature (e.g., Applebee et al., 2003; Winkler, 2020), science (e.g., Förtsch et al., 2016; Minner et al., 2010; Tekkumru-Kisa & Stein, 2015), and predominantly mathematics—the subject matter of the current study (e.g., Baumert et al., 2010; Henningsen & Stein, 1997; Hiebert & Grows, 2007; Stein et al., 2009)—meta-analyses of the past decades have underscored the challenges in defining and interpreting cognitive activation due to its varied terminology and conceptualizations, leading to empirical complexities (Hattie, 2009; Seidel & Shavelson, 2007; Schlesinger & Jentsch, 2016). In particular, several authors use different terms to refer to the notion of cognitive activation, such as challenging tasks[3] (e.g., Sullivan et al., 2015), intellectually demanding tasks (e.g., Doyle, 1988), cognitively demanding tasks/instruction (e.g., Boston & Smith, 2011), exploratory tasks (e.g., Chapman, 2013), cognitive demand (e.g., Hsu & Yao, 2023), cognitive or mathematical challenge (in mathematics education, e.g., Applebaum & Leikin, 2014), and inquiry-based teaching (mainly used in science education, e.g., Tekkumru-Kisa et al., 2019). Acknowledging that there are some differences between the terms, in this study, the term 'cognitive activation' specifically denotes the process by which teachers engage students in challenging mathematical tasks and discourse, thereby stimulating critical thinking, problem-solving, and deeper understanding of the mathematical concepts (Kunter et al., 2013). This term primarily focuses on the teacher's actions, which are not inherently present in the rest of the terms.[4] The terms 'mathematical

[3] The rationale for selecting and using challenging tasks in teaching is discussed in a subsequent section.
[4] When the focus shifts to the students' engagement rather than the teacher's actions, the term 'cognitive activity' is used in the literature (Charalambous & Praetorius, 2020)—but it is not utilized in the current study.

challenge', 'cognitive demand', 'cognitively demanding tasks', and 'challenging tasks' are closely related and typically refer to the inherent challenge of the tasks (Hsu & Yao, 2023). To avoid confusion the term 'mathematically challenging tasks' is used when referring to tasks in this study.

The roots of cognitive activation in mathematics can be traced back to several key educational theories and research studies that have emphasized the importance of engaging students in deep, thoughtful learning processes. Originating from the field of psychology, particularly constructivist theories, the work of Jean Piaget (1936) stands out. Piaget proposed that learners are not passive recipients of knowledge; instead, they actively construct their own understanding and knowledge of the world, through experiencing things and reflecting on those experiences. Applied to mathematics education, this theory transforms the traditional learning paradigm, encouraging students to actively engage with mathematical concepts and problems. Lev Vygotsky (1978), another important figure in psychology, introduced a different perspective by emphasizing that social interactions with others play a crucial role in cognitive development. His theory implications can be seen in a classroom where students can think out loud with their peers and the teacher.

Building upon these foundational theories from psychology and their implications for mathematics education, the focus shifts to their practical application in the classroom. Acknowledging the pivotal role of teachers and their teaching in stimulating cognitive activation, cognitive activation is considered as a teaching practice[5] (Lipowsky et al., 2009). Since students' cognitive activity cannot be observed directly, it is usually described by the kind of teaching offered to students (Förtsch et al., 2016). A higher level of cognitive activation is potentially achievable when teachers encourage students to decode information; explore concepts and relationships; solve non-routine problems, share, explain, and compare solutions; and connect prior to new knowledge, by engaging them with challenging tasks, differing ideas, and cognitive conflicts and by participating in content-related discourse (Boston & Smith, 2011; Hsu & Yao, 2023; Kunter et al., 2013; Lipowsky et al., 2009). Conversely, a focus on non-challenging tasks may enhance a low level of cognitive activation. Such teaching is related to fluency with facts and procedures, which is definitely important in reducing

[5] A teaching practice is defined as an action that a teacher does "constantly and habitually" to support learning (Lampert, 2010, p.25).

students' cognitive load of working memory and in focusing their attention on selecting and devising a solution plan (Smith & Stein, 1998).

Tasks are classroom-based elements serving as a bridge between teaching and learning, and a vehicle to understand how the former influence the latter (Tekkumru-Kisa et al., 2020). Research has suggested that cognitive activation is greatly determined by the cognitive demands of the selected tasks as such; the ways the teacher and the students negotiate their work around the task; and the teacher-students' (inter)actions (Herbst, 2006; Herbst et al., 2017; Kunter & Baumert, 2006; Stein et al., 2007). Specifically, the nature of a task, the intellectual processes and products expected by the students when they work with it, and the resources and scaffolding that are available to assist students in doing so, all direct their attention on particular mathematical ideas (Stein et al., 1996), and determine whether they will be involved in high-level thinking and develop a well-structured and elaborated knowledge base (Lipowsky et al., 2009).

In summary, linking back to the foundational psychological theories of cognitive activation it appears that "the demand of tasks and the support from the external environment as well as the past experiences that individuals have can determine cognitive workload" (Hsu & Yao, 2023, p. 222). Smith and Stein (2011) highlighted the crucial role of teachers' decisions in lesson planning and enactment to either maintain or reduce the cognitive demands placed on students and the degree to which they will "develop their ideas for themselves" (Marshall & Horton, 2011, p. 99). Thinking for themselves rather than having someone else (e.g., the teacher, a classmate, or the task itself) doing the thinking for them is the fundamental idea of treating students as sense-makers who can engage in serious intellectual activity and nudging them toward higher-order thinking skills and goals (Cohen, 2011; Lampert et al., 2013; Sullivan et al., 2015). The next section elaborates more on the benefits of student engagement in challenging work.

2.2.2 The Importance of Engaging Students in Challenging Work

Over the last decades, the significance of cognitively activating students has been highlighted by several conceptual frameworks (e.g., Bolhuis, 2003; Hiebert & Grouws, 2007; Seidel & Shavelson, 2007), as well as a wealth of empirical research findings (e.g., Boaler & Staples, 2008; Boston & Smith, 2011; Hiebert & Wearne, 1993; Stein & Lane, 1996; Stein et al., 2007). These studies demonstrate that classrooms where cognitive activation is high and the demands of mathematically challenging tasks are consistently sustained during lesson enactment, yield significantly better learning outcomes compared to classrooms in which this is not the case. In particular, learning environments that prompt students to use multiple strategies, engage in complex and creative thinking, make connections between mathematical concepts, and provide explanations tend to also exhibit higher performance in mathematical thinking, reasoning, and problem-solving abilities compared to environments lacking these characteristics (Stein & Lane, 1996; Tarr et al., 2008). These findings hold true across both elementary (e.g., Hiebert & Wearne, 1993; Stein & Kaufman, 2010) and secondary education (e.g., Hollingsworth et al., 2003; Stigler & Hiebert, 2004; Jackson et al., 2013; Stein & Lane, 1996), and is evident even in disadvantaged school settings or in schools employing various curriculum types (e.g., Boaler & Staples, 2008; Croninger et al., 2006). Moreover, working with mathematically challenging tasks within a supportive classroom culture has been shown to academically benefit students from diverse linguistic, cultural, and economic backgrounds, significantly narrowing the achievement gap between groups of students of different nationalities (Boaler & Staples, 2008). This aspect ties into the concept of differentiation, a topic that will be revisited later on.

A lack of deep conceptual understanding can lead to mechanically solving tasks and dealing with mathematical concepts, operations, relations, and ideas, without truly grasping their underlying meaning (Resnick & Zurawsky, 2006; Watson & Sullivan, 2008). The rote learning of rules and procedures promotes low cognitive activity even if the content under consideration is of an advanced level (e.g., integral calculus in advanced mathematics classes in high school); in contrast, the comprehension of mathematical concepts can stimulate high cognitive activity, even when dealing with basic mathematical content (e.g., the addition of one-digit numbers without regrouping in first grade, Resnick & Zurawsky, 2006). Thus, prioritizing high cognitive activation in teaching becomes imperative, as it equips students with robust conceptual understanding, advanced mathematical thinking, and reasoning skills, empowering them to tackle complex mathematical problems rooted in real-world contexts.

Sparked by teachers' concerns that such a level of demands may induce anxiety, research has delved into students' responses to mathematically challenging tasks (Cheeseman et al., 2013; Henningsen & Stein, 1997). The literature presents mixed findings. A U.S. longitudinal, large-scale multiple case study highlighted that students in classes with a sustained mathematical challenge (i.e., in which teachers assigned challenging tasks and asked thought-provoking questions, emphasizing that success in mathematics comes from effort rather than innate ability) reported enjoying mathematics more and achieved higher learning outcomes (Boaler & Staples, 2008). However, a Swiss-German study found that when confronted with mathematically challenging tasks during autonomous work, students developed negative feelings about their involvement and believed they had not grasped the content of the lesson well (Hugener et al., 2009). Additionally, other research linked engagement with less mathematically challenging tasks with student motivation aspects, such as interest and effectiveness in mathematics, less tendency to avoid task solving, and fewer experiences of negative feelings (Gilbert, 2016; Stipek et al., 1998). The tension among the findings suggests that engaging students in solving mathematically challenging tasks does not suffice; while these tasks inherently carry risk and ambiguity, a supportive learning climate in the classroom is warranted in order to encourage student persistence and achievement (Lipowsky et al., 2009; Sullivan et al., 2013; Sullivan & Mornane, 2013; Tarr et al., 2008).

As students explore a mathematical topic, they experience cognitive conflict which does not necessarily mean that they have not understood the new content sufficiently. Rather, it is a normal part of the knowledge-building process. Students begin to question their existing knowledge, necessitating teacher facilitation and support to create links between their prior and new knowledge (Pogrow, 1988). Although challenging and potentially frustrating, especially for first-timers, this process of dismantling confidence in prior conceptions through contradictory experiences, and reshaping them into more accurate understandings, yields significant benefits. With consistent encouragement from teachers to tackle challenging tasks and persevere, students at all educational levels show marked improvements in conceptual understanding (e.g., Adnyani, 2020; Lipowsky et al., 2009; Wu & Lin, 2016).

Consequently, the variance in cognitive activation opportunities appears to be linked to disparities in curriculum materials, the tasks used, and the teaching methodologies employed in each classroom. Recognizing the paramount importance of tasks, along with their selection and enactment, the subsequent section delves into examining the critical impact of designing, choosing, and implementing mathematical tasks on student cognitive activity.

2.2.3 Task Design, Selection, and Implementation: Key Drivers of Cognitive Activation

Almost four decades ago, Doyle (1988) introduced the concept of the 'task', highlighting the key role of task selection in teaching and in facilitating student learning. He emphasized how task selection is pivotal in directing student attention towards specific content aspects, demonstrating ways to process new and unfamiliar content and problems, and determining the teacher-student relationship. Defined as a single or complex problem, question, activity, or a series thereof, a task is assigned to students with the goal of directing their attention to particular subject-matter concepts, ideas, and skills[6] (Stein et al., 1996; 2009). For example, a mathematical task could be an exploratory activity found in the student textbook or a classroom activity organized by the teacher. Stein and colleagues (2009) emphasized that a task qualifies as *mathematical*, if it focuses on mathematical concepts, contributing to a better exploration and understanding of the content. Conversely, activities, such as solely cutting out geometric shapes without a mathematical purpose, are not deemed mathematical tasks since they disorient students from mathematics. Focusing on tasks is important, as teaching predominantly revolves around them, and engaging students in demanding tasks has been shown to significantly enhance their learning outcomes (Stein et al., 2009).

Analyzing a task's cognitive level can be particularly complex, as it may comprise a blend of both challenging and less demanding sub-tasks (Silver et al., 2009; Smith & Stein, 1998). The 'cognitive level of a task' refers to the mental processes required from the initiation to the completion of the task, culminating in the final outcome (Doyle, 1988; Stein et al., 2009). As previously discussed, the cognitive demands of a task can either stimulate high or low levels of cognitive activation. Thus, the nature of the tasks creates opportunities for student learning (Sullivan, 2011).

Acknowledging the critical role of tasks in student learning and drawing from an extensive research body on the influence of mathematical tasks on student learning, Stein and Lane (1996) introduced the *Task Analysis Guide (TAG*). This tool aids teachers and researchers in defining and elucidating the nature of students' cognitive activity around a task, by classifying a mathematical task as presented in the curriculum materials and textbooks, based on its characteristics and the level of thinking required

[6] A task "is not classified as a different or new task unless the underlying mathematical idea toward which the activity is oriented changes" (Stein et al., 1996, p. 460).

by students (Hsu & Yao, 2023). Additionally, TAG provides teachers and teacher educators with a common language to discuss the selection and enactment of tasks. Within this framework, tasks are classified into two main categories, high and low mathematical challenging, each further divided into two subcategories, as depicted in Table 1.

Table 1

The Task Analysis Guide (adapted from Stein et al., 2009, p.16)

Low Mathematical Challenge		**High Mathematical Challenge**	
Memorization Tasks	***Procedures without Connections Tasks***	***Procedures with Connections Tasks***	***Doing Mathematics Tasks***
▪ Involve the exact reproduction of previously learned facts, rules, formulas, or definitions. ▪ Are not amenable to solution through procedures. ▪ Present the information to be recalled clearly and straightforwardly. ▪ Lack a link to the underlying concepts or meanings associated with the facts, rules, formulas, or definitions being learned or reproduced.	▪ Follow an algorithm, with the task either explicitly requiring its use or making it apparent based on previous lessons, or the task's context. ▪ Prioritize arriving at the correct solution over fostering deeper understanding of mathematical concepts. ▪ Lack ties to the meanings that underpin the employed procedure. ▪ Do not require explanations or only require explanations that detail the steps of the utilized procedure.	▪ Direct students' focus on procedural applications that are closely linked to underlying conceptual ideas. ▪ Require students to interact with and comprehend the concepts of the procedures for successful task completion and enhanced understanding. ▪ Typically offer multiple representations such as visual diagrams, manipulatives, or symbols, fostering meaning through connections among these varied forms.	▪ Do not hint at a predictable, well-rehearsed method or solution pathway explicitly suggested by the task, or a worked-out example. ▪ Require complex and non-algorithmic thinking. ▪ Require students to: ○ explore and understand mathematical concepts, processes, or relationships. ○ draw upon their existing knowledge and experiences, applying them judiciously to navigate the task. ○ analyze the task and its constraints that could limit possible solutions.

The high mathematical challenge category includes tasks that either lack a predictable solution path or are procedural but closely linked to mathematical concepts and ideas, namely 'doing mathematics' and 'procedures with connections' tasks. In 'doing mathematics'[7] tasks (which are considered to induce the highest level of mathematical challenge compared to other task types), students engage in complex, non-algorithmic thinking, without a predefined solution path, leading to potentially multiple methods of solution (see Figure 1). This cognitive process mirrors the problem-solving process of mathematicians, fostering open-ended exploration, diverse solutions, and creative responses. While some tasks may be broadly open-ended, others may target specific mathematical aspects or curricular topics (Sullivan, 2011). These tasks might also encourage students to develop strategies and grapple with mathematical ideas within real-life scenarios. However, a problem's context is a surface-level attribute; while it can aid in understanding and solving the problem, it does not inherently elevate the task's cognitive demands.

Figure 1

An Example of a 'Doing Mathematics' Task on the Introduction to Equivalent Fractions

At a school event, there were three pizzas of the same size. Each pizza was divided into equal parts.

- Sophia ate 2 slices of Pizza A.
- Andrew ate 3 slices of Pizza B.
- Michael ate 4 slices of Pizza C.
- All children ate the same amount of pizza.

How is this possible? Explain your thinking.

Pizza A Pizza B Pizza C

(Mathematics Curriculum, 2017, Grade 4, Unit 8, p.36, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

Figure 1 represents an example of a 'doing mathematics' task, where students delve into an inquiry learning process to explore the concept of equivalent fractions within a daily life context. They are expected to leverage both their fraction models and observational skills to articulate their explanations, engage in mathematical reasoning, and draw conclusions about the relationship between the numerator (number of pizza slices eaten) and the denominator (total number of pizza slices). The task's complexity

[7] Previous research has referred to these tasks using terms such as "worthwhile tasks" (NCTM, 2000), "powerful tasks" (Krainer, 1993), "cognitively challenging" tasks, "high-level" tasks (Stein et al., 1996), and "making connections" tasks (Hiebert et al., 2005).

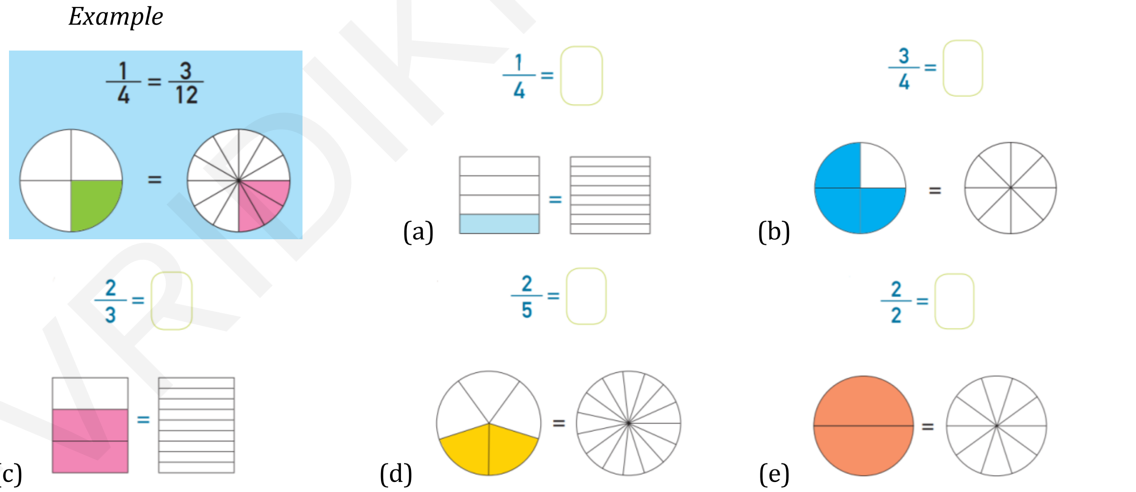
is heightened as students come to realize the multitude of valid solution paths and answers available. For instance, they might observe that Sophia, Andrew, and Michael consumed two out of four slices of Pizza A, three out of six slices of Pizza B, and four out of eight slices of Pizza C, respectively. Alternatively, they could express these quantities as 2/8 of Pizza A, 3/12 of Pizza B, and 4/16 of Pizza C. This engaging task not only deepens students' understanding of fractions but also cultivates their ability to think critically and embrace diverse mathematical approaches.

In the second category, 'procedures with connections tasks,' students adhere to a suggested solution pathway, implementing a broad process or algorithm that has a close connection to the underlying mathematical concepts, as depicted in Figure 2. These tasks might incorporate tangible or other representations to demystify abstract mathematical ideas. This approach enables students to cultivate relevant vocabulary, explore various representations of a concept, and provides teachers with insight into their students' understanding through their solution methods (Sullivan, 2011).

Figure 2

An Example of a 'Procedures with Connections' Task on Equivalent Fractions

Shade appropriately to form equivalent fractions, as in the example.



(Mathematics Curriculum, 2017, Grade 4, Unit 8, p.38, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

Figure 2 illustrates an example of a 'procedures with connections' task, in which students are expected to follow a clearly outlined procedure, demonstrated through a worked-out example, to identify pairs of equivalent fractions. This method is directly tied to the underlying meaning of equivalent fractions, encouraging students to establish links between symbolic and visual representations, and enhancing their understanding of the topic.

The category of low mathematical challenge encompasses two distinct types of tasks: 'procedures without connections' and 'memorization.' In 'procedures without connections' tasks, students engage in practicing a well-known process or algorithm through structurally similar examples, yet these tasks offer minimal, if any, connection to underlying mathematical concepts and provoke limited cognitive engagement (see Figure 3). It is noteworthy that tasks aimed at fostering procedural fluency represent one of the most prevalent types of tasks found in mathematics classrooms (Sullivan, 2011).

Figure 3

An Example of a 'Procedures Without Connections' Task on Equivalent Fractions

Circle all fractions that are equivalent to the fraction on the left card.

$\frac{2}{5}$	$\frac{6}{10}$	$\frac{8}{20}$	$\frac{4}{10}$	$\frac{6}{15}$
$\frac{12}{20}$	$\frac{6}{10}$	$\frac{4}{8}$	$\frac{3}{5}$	$\frac{24}{40}$
$\frac{6}{8}$	$\frac{18}{32}$	$\frac{30}{40}$	$\frac{12}{16}$	$\frac{3}{4}$

(Mathematics Curriculum, 2017, Grade 4, Unit 8, p.41, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

For example, to solve the task in Figure 3, students must identify equivalent fractions by multiplying or dividing the numerator and denominator of the green card's fraction by the same whole number. Despite multiple equivalent options, the procedure is straightforward and learned from prior teaching. Students aim to find correct answers without explaining or showing an understanding of the procedure.

'Memorization' tasks, the last task type, offer a slightly lower mathematical challenge than 'procedures without connections' tasks. Students recall and reproduce previously learned rules, definitions, or processes, without engaging with the underlying mathematical concepts or meanings, resulting in minimal cognitive activation. For instance, in Figure 4's task, students fill in the blank with a familiar term and then reproduce a memorized definition of equivalent fractions. The task requires simple recall rather than procedural follow-through.

Figure 4

An Example of a 'Memorization' Task on Equivalent Fractions

Complete the sentences.

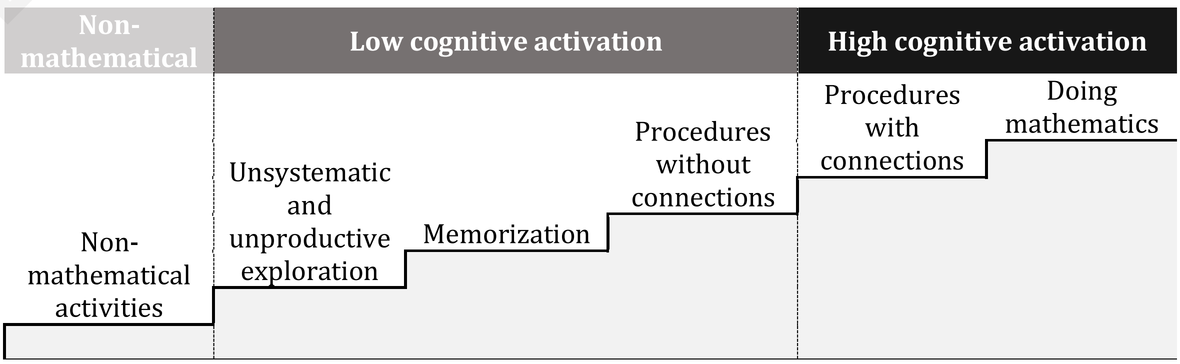
- $\frac{1}{2} = \frac{2}{4}$ The fraction one-half is ……………………………….. to two-fourths.

- Equivalent fractions are those fractions that ……………………………………………………………….
……

Stein and Lane (1996) identified two additional task categories in mathematics classrooms beyond the previously mentioned four. The first, 'unsystematic and unproductive exploration,' falls under the low mathematical challenge umbrella (refer to Figure 5). Teachers select these tasks aiming for high cognitive activation; however, during their presentation and implementation, efforts to immerse students in mathematical processes fall short. The engagement with mathematical concepts becomes erratic, hindering students' ability to develop mathematical strategies and understanding. Literature suggests that tasks typically devolve into non-systematic explorations due to factors such as the task's unsuitability for the specific students (e.g., low motivation levels or insufficient prior knowledge), unrealistic teacher expectations, improper time allocation for task completion (too little or too much), or reduced student accountability (Henningsen & Stein, 1997). The final category, 'non-mathematical activities,' involves tasks in which students do not engage in processes centered around mathematical content and goals, such as the example involving cutting geometric shapes mentioned earlier. Figure 5 provides a concise classification of the various task types, based on the level of cognitive activation they promote, as per the TAG.

Figure 5

Task Classification Based on their Demands and Cognitive Processes Required to be Solved

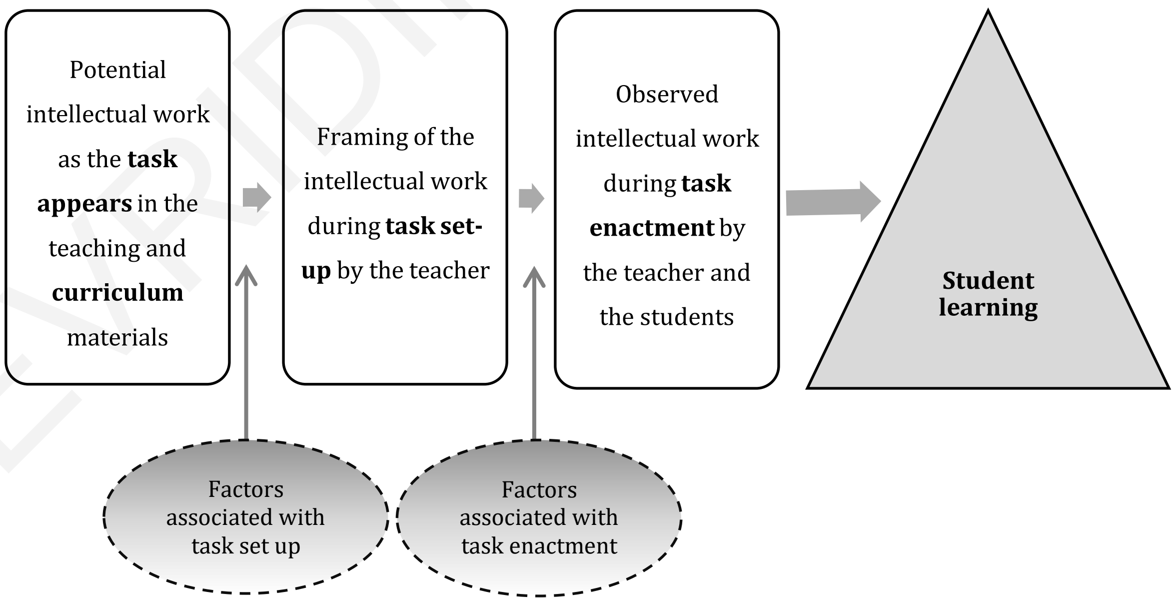


In summary, Figure 5 delineates the hierarchy of cognitive demands and the required depth of understanding for various task types (which is increased as progressing from lower to higher steps). While all task categories contribute to a comprehensive learning experience, with the exception of the last two at the lower steps, it is the engagement with more mathematically challenging tasks that proves vital for attaining a profound conceptual understanding (Sullivan, 2011; Boston, 2009).

Research has consistently demonstrated a connection between the teachers' decisions and pedagogical actions and the ways in which the mathematically challenging tasks are finally implemented during lesson enactment (e.g., Boston & Smith, 2011; Henningsen & Stein, 1997; Smith et al., 2008; Stylianides & Stylianides, 2020; Warshauer, 2015). To better understand how tasks unfold in mathematics classrooms, Smith and Stein (1998) devised the Mathematical Task Framework (MTF). Figure 6 offers a visual representation of this framework, illustrating the key components and relationships that define the MTF, and showcasing how different elements interact to shape the learning experience in mathematics classrooms.

Figure 6

The Mathematical Task Framework and Factors Associated with Maintenance or Decline of the Task's Challenge (adapted from Stein & Lane, 1998)



The Mathematical Tasks Framework (MTF), as depicted in Figure 6, outlines a three-phase progression that each task undergoes in a mathematics classroom. Initially, the task is presented in written curriculum or teaching materials, showcasing its potential to engage students in intellectual work (first rectangle). Following this, the teacher introduces the task in the classroom, setting the stage for the expected level of student intellectual engagement (second rectangle). Finally, the task comes to life as both the teacher and students interact with it during the lesson (third rectangle), culminating in students' actual intellectual work (Doyle, 1988; Tekkumru-Kisa et al., 2019).

The evolution of a task through its three phases significantly impacts cognitive activation and student learning outcomes (see triangle, Figure 6; e.g., Hiebert & Wearne, 1993; Desimone et al., 2013; Stein & Lane, 1996). Yet, the MTF emphasizes that what ultimately determines student learning is how the task unfolds during teaching[8]. In fact, high-quality teaching involves regularly exposing students to high-level tasks, while consistently maintaining their demands when enacting them (Smith & Stein, 2023). Teachers may start with a well-chosen, high-level task, but it is their responsibility to sustain its mathematical rigor throughout the lesson; these tasks are the most difficult to implement as intended. As Tekkumru-Kisa and colleagues (2020) nicely put it, tasks are "in motion," dynamically interacting with the level and nature of student thinking, underscoring the fluidity and adaptability required in effective mathematics instruction.

A wealth of research highlights the challenges teachers often encounter in preserving the mathematical rigor of tasks during teaching, even if they are provided with mathematically challenging tasks (e.g., Charalambous & Hill, 2012; Boston & Smith, 2011; Henningsen & Stein, 1997; Stein et al., 1996; Stein & Smith, 1998; Stigler & Hiebert, 2004; Weiss & Pasley, 2004). As a task progresses through the three phases of implementation explained earlier (see Figure 6), the complexity often diminishes as a result of the teacher's practices and interactions with students. Therefore, simply providing curriculum materials that contain engaging and mathematically challenging tasks does not, in itself, guarantee their enactment at their designer's intended level of challenge (Choppin, 2011; Sullivan et al., 2009).

In sum, challenging tasks are a vehicle for cognitive activation, as they create an environment where students are encouraged to engage deeply, think critically, and learn meaningfully. The TAG and the MTF serve as two vital roadmaps for teachers,

[8] All three phases, but particularly the implementation phase, contribute to student learning. Observational evidence highlights the importance of sustaining task demand in the setup phase to help students develop a shared language for discussing the task's context and mathematics, and also foster richer concluding whole-class discussions (Jackson et al., 2013).

guiding the journey of a mathematical task from its inception in curricular materials to its enactment in the classroom. It is within this dynamic and interactive process that the true essence of student learning is realized. Teachers must not only select tasks with high cognitive potential but also skillfully navigate the ebb and flow of classroom interactions to maintain the task's intellectual integrity. This delicate balance of preparation and improvisation, of structure and fluidity, is what transforms a mathematical task from a mere exercise to a profound learning experience. Managing the complexity of the task and maintaining its high demands, inherently connected to cognitive activation, requires careful consideration of various factors that could impact the levels of task design, set-up, and enactment. The following sections delve into such critical factors.

2.2.4 Factors Supporting Teachers Maintaining Cognitive Activation

Maintaining high cognitive demands in a task is facilitated by a dynamic interplay of factors that span the phases of task selection, setting, and implementation, as well as the fluid transitions between them. These factors can be categorized into five groups[9]: (a) *Selection and design of appropriate challenging tasks*; (b) *Teacher decisions and practices during teaching*; (c) *Teacher's Mathematical Knowledge for Teaching*; (d) *Classroom (socio-mathematical) norms*; and (e) *Student-related factors*. Each of these critical elements plays a vital role in sustaining the intellectual rigor of mathematical tasks, and they are explored in detail next.

Unsurprisingly, the initial set of factors pertains to *selecting or designing appropriate, challenging tasks* (e.g., Chapman, 2013; Smith & Stein, 1998). These tasks should align with learning objectives, promote deep conceptual understanding and establish links between mathematical ideas and real- world applications. They should offer an appropriate level of challenge to stimulate mathematical thinking and encourage communication and reasoning. Additionally, these tasks need to be contextually relevant to pique student interest and should be based on a foundation of prior knowledge that students possess, enabling them to engage effectively. The characteristics of mathematically challenging tasks, which aid teachers in their selection or creation, have been previously detailed and will not be further elaborated here.

[9] This list is not exhaustive and serves as a starting point for understanding the complex factors involved in maintaining cognitive activation. The same applies to the next section.

The second set of factors revolves around the *teaching decisions and practices*, including ingrained habits and tendencies, that a teacher consciously or unconsciously employs during teaching. These factors significantly influence the setup and enactment of the task (Stein & Smith, 1998). Examples of teacher actions that can positively impact the task include: facilitating students' thinking and reasoning or clarifying the task without directly providing solutions or demonstrating methods; encouraging students to find their own solutions rather than instructing them on how to solve the task (i.e., holding back from telling); actively listening to and interpreting students' arguments and language before intervening, and building on their ideas; offering balanced levels of support and scaffolding, enabling students to monitor their progress and reflect on the process; allotting adequate time for productive struggle with the task, while managing individual and collaborative work periods effectively; leveraging students' prior knowledge and frequently making conceptual connections; consistently prompting students to explain, represent, interpret, reason, justify, and evaluate through strategic questioning, commenting, and feedback; introducing "just-in-time" questions and prompts to problematize students; thoughtfully selecting and sequencing solutions for whole-class discussions; fostering a classroom environment that encourages risk-taking and holds students accountable for documenting their thoughts and strategies (e.g., Akcil-Okan & Tekkumru-Kisa, 2021; Boaler & Staples, 2008; Cheeseman et al., 2013; Cheeseman et al., 2016; Henningsen, & Stein, 1997; Ingram et al., 2016; Ponte & Quaresma, 2016; Pettersen & Nortvedt, 2017; Roche & Clarke, 2015; Roche et al., 2013; Smith & Stein, 1998; Sullivan et al., 2013; Warshauer, 2015). Also, Smith and Stein's (2011) work outlined five strategic practices that teachers can use to facilitate productive classroom discussions around mathematical tasks. These practices include anticipating student responses to tasks, monitoring students' work, selecting particular students to present their solutions, sequencing the student responses to be shared, and connecting different students' responses to highlight the key mathematical ideas. The aforementioned practices reveal that much of what happens during teaching is largely within the teachers' control (Smith & Stein, 2023).

Third, task unfolding in mathematics classrooms is profoundly shaped by teachers' Mathematical Knowledge for Teaching (MKT), which encompasses their understanding of subject matter, Pedagogical Content Knowledge, and insights into their students' learning processes. A robust MKT enables teachers to effectively engage students in cognitively demanding activities, fostering a deeper understanding of mathematical concepts (Charalambous, 2010; Hill et al., 2005). To create a mastery-focused learning environment, teachers must guide students towards a deep understanding of mathematics, connecting new information to prior knowledge and highlighting its practical applications (Ames, 1992). This requires a comprehensive grasp of the subject matter, including awareness of multiple solution approaches and the ability to utilize a variety of teaching strategies, representations, and explanations. Teachers need to discern when, why, and how specific teaching practices are relevant to each teaching situation, adjusting their approach to suit the unique dynamics of their classroom. In fact, teachers must understand how students' ideas develop; recognize alternative perspectives and solutions; identify potential student difficulties; and draw connections between student ideas and important mathematical concepts (Baumert et al., 2010; Hill et al., 2005; Stein et al., 2007; Sullivan & Mornane, 2014). It is essential for teachers to be well-versed in the pedagogies associated with high-demand tasks, including understanding the nature of challenging tasks; identifying, selecting, and designing such tasks; analyzing their cognitive demands; connecting the tasks to students' understandings, interests, and experiences; maintaining the demands (Chapman, 2013). Being prepared to effectively implement such strategies ensures that the cognitive demands of tasks are maintained, ultimately contributing to rich student sense-making of mathematics and of doing mathematics (Baumert et al., 2010).

All these teacher practices contribute to the development of a learning culture and the establishment of classroom socio-mathematical norms that bolster students' engagement with challenging tasks. Experimental empirical evidence from classrooms shows that the teacher and students mutually construct socio-mathematical norms (defined as epistemological mathematical approaches established and monitored in social situations) and expectations about what counts as good mathematical knowledge, reasoning, and argumentation (Yackel & Cobb, 1996). In classroom cultures where learning is oriented toward sensemaking, negotiating understanding, accepting errors, and celebrating difference, students value and monitor their commitment to classroom norms that support the cognitive activity of their class, without needing constant reminders about those norms (ibid.). The teacher plays a pivotal role in fostering a conducive classroom culture for the implementation of challenging tasks, by establishing explicit socio-mathematical norms. This includes defining acceptable modes of communication, highlighting valued types of responses, and setting expectations for embracing mathematical challenge (Clarke et al., 2014a; Roche & Clarke, 2015; Sullivan, 2011).

Finally, student-related factors play a crucial role in determining the level of effort and engagement in a task. Student persistence, defined as the determination to understand and solve a task, is vital for delving into significant mathematical concepts that may not be instantly clear but are within reach of their current understanding (Clarke et al., 2014a; Sullivan & Mornane, 2014). This persistence is bolstered when the teacher intervenes judiciously, stepping in only when absolutely necessary to clarify potential misconceptions, assist students who are stuck, or present additional challenges to those who have completed the task. Also, students' prior knowledge, motivation, and attitudes towards the lesson content or the subject matter at hand play a significant role in shaping their learning experience and outcomes (Doyle, 1983).

Building on the previous discussion, utilizing challenging tasks to cognitively engage students sets high expectations for teachers. They are required to sustain the task's complexity while aiding students in their perseverance—something that can prove to be quite challenging (Sullivan, 2011). The following section delves into the challenges and constraints teachers may face when implementing challenging tasks.

2.2.5 Factors Inhibiting Teachers from Maintaining Cognitive Activation

Previous studies have pinpointed a range of factors and conditions that contribute to the erosion of rigor in challenging mathematical tasks. These include (a) *the selection and design of tasks that fail to meet the necessary level of challenge;* (b) *ineffective teaching decisions, practices, and actions by teachers;* (c) *unproductive teacher beliefs, attitudes, and dispositions regarding the mathematical challenge and the pedagogy of mathematics;* (d) *negative past teaching experiences of teachers with the curriculum;* (e) *gaps in teachers' Mathematical Knowledge for Teaching (MKT);* (f) *student behaviors and attitudes;* (g) *insufficient student accountability for high-level outcomes and processes;* (h) *classroom management issues*; and (i) *constraints of time*. Each of these factors will be explored in further detail below.

The selection and design of tasks are pivotal in influencing students' cognitive engagement. Teachers often struggle to assess the challenge presented by a task or to choose tasks that demand higher-level thinking (Boston, 2013; Smith & Stein, 2023). Commonly, teachers may not scrutinize mathematical tasks based on the depth of thought they require from students. For instance, tasks are frequently grouped by their mathematical content or surface characteristics, such as whether they are word problems or include tables or other representations (Arbaugh & Brown, 2005; Remillard & Bryans, 2004). Additionally, there is a tendency to rely on textbook tasks, aiming to cover the content and skills they are obliged to within the academic year (Boston, 2013). This approach may overlook how different types of tasks can shape students' cognitive processes. Sometimes, the issue is that teachers may not fully understand the impact of task types on student thinking and learning. In other instances, if the chosen task is not well-suited to the student group—perhaps due to excessive difficulty or unclear expectations and instructions—students may lack the interest, motivation, or background knowledge necessary to engage with the task, leading to a decline in high-level cognitive activity (Smith & Stein, 1998).

Selecting challenging tasks for teaching does not inherently ensure that students will engage in high-level thinking and reasoning (Boston, 2013). For instance, teachers may inadvertently undermine these tasks by over-explaining them, thus limiting opportunities for students to explore the problems independently (Smith & Stein, 1998). Furthermore, students may seek step-by-step guidance, prompting teachers to provide detailed instructions that takes over the thinking replace student-led problem-solving (Henningsen & Stein, 1997). The task's intended challenge can also be diluted if a teacher decomposes it into less-demanding subtasks or shifts the focus from conceptual understanding to merely obtaining the correct answers and following procedures (Arbaugh et al., 2006). Also, teachers' reluctance to offer critical feedback on students' work and the lack of whole-class discussion to discuss and compare students' solutions reduces cognitive activation (Stylianides & Stylianides, 2020). In such scenarios, tasks that are meant to be challenging and thought-provoking are reduced to routine exercises.

Teachers' beliefs, attitudes, and dispositions regarding the mathematical challenge and the pedagogy of mathematics can also act as barriers to stimulating students' cognitive engagement (Boston, 2013). Teachers accustomed to traditional, teacher-led instruction and routine tasks may find it unsettling to facilitate an environment where students grapple with ambiguous, challenging problems. Notably, even teachers who favor cognitively demanding tasks may inadvertently lower the level of challenge when implementing them (Stein et al., 1996). Recent findings by Russo and colleagues (2019) highlight two significant hurdles linked to teachers' beliefs that impede the promotion of high-level thinking in students. Firstly, some teachers hold the view that challenging tasks are suitable only for the more advanced students, based on the assumption that intelligence is static and that such tasks are too challenging for those perceived as less advanced for whom the task is beyond their ability or control. Secondly, teachers may perceive certain tasks as too abstract or devoid of meaningful and relevant context, making them seemingly inaccessible to students. These beliefs, including those shaped by teachers' own experiences as learners, significantly influence the teaching strategies they employ and their commitment to engaging students in rigorous mathematical tasks (Bobis et al., 2016).

Moreover, teachers' previous teaching experiences with curriculum materials often lead to a simplification of task complexity (Choppin, 2011). Studies have shown that teachers' preferences for certain traditional curriculum resources shape their teaching methods (Remillard & Bryans, 2004). Their approach to the curriculum can result in the conventional use of innovative materials, which in turn affects the learning opportunities available to students. Novice teachers, in particular, may adopt a 'piloting stance,' adhering strictly to the new unfamiliar curriculum by incorporating all its resources and tasks as they are presented in their lesson (ibid.). Furthermore, if prior enactments of challenging tasks have previously required extensive teacher support for student success, teachers might modify these tasks to ensure completion, inadvertently diminishing their complexity (Boston, 2013).

Furthermore, research has explored how various types of teacher knowledge are linked to the reduction in task complexity. A significant body of research—mainly qualitative in nature—has highlighted the connection between teachers' MKT and the cognitive demands of tasks in practice (e.g., Charalambous, 2010). Teachers with limited MKT may not fully grasp the tasks, their demands, and their potential, leading to feeling of insecurity and a reluctance to use them (Boston, 2013). Gaps in MKT can hinder teachers' abilities to effectively use questioning, sequence tasks, draw connections, employ representations, and foster student reasoning, generalization, and justification (Hathcock et al., 2015; Ponte & Quaresma, 2016). Additionally, while a deep understanding of the subject matter (Content Knowledge) is crucial for aligning teaching materials with the curriculum and choosing mathematically challenging tasks, it alone does not guarantee insightful instruction (Baumert et al., 2010). Insufficient Content Knowledge can limit teachers' capacity to facilitate students' understanding through appropriate explanations and representations (e.g., Ball, 1990; Borko et al., 1992; Stein et al., 1990; Putnam et al., 1992). Thus, a robust foundation in both MKT and Content Knowledge is essential for teachers to maintain the Integrity of mathematically challenging tasks and to cultivate an environment where students can engage in and benefit from high-level mathematical thinking.

In addition, the significance of MKT in preserving task complexity has also been emphasized. This knowledge encompasses the ability to select and develop tasks that enhance students' conceptual understanding, stimulate their mathematical thinking, and engage their interest and curiosity, as well as to maximize the learning potential of these tasks (Chapman, 2013). Teachers lacking in MKT may struggle to comprehend the multifaceted nature of challenging tasks, such as their solvability through multiple methods, and to identify, select, and design cognitively demanding tasks. They may also find it difficult to identify and analyze the cognitive demands of tasks, focus on students' understanding, and grasp how their task selection and implementation affect students' mathematical sense-making (Carson, 2010). Furthermore, they might not know which aspects of a task to emphasize, how to manage student work effectively, or what questions to ask and support to provide to challenge students appropriately without taking over their cognitive processes (ibid.).

Student behaviors and attitudes can also impose constraints, particularly when they dovetail with teacher practices. Many students shy away from risk-taking and perseverance in the face of challenging tasks, often seeking explicit procedures or steps to solve the task from the teacher. In turn, teachers may succumb to this pressure, positioning themselves as the sole source of guidance in the classroom (Sullivan, 2011). This dynamic can lead teachers to preemptively simplify tasks during lesson planning, anticipating the need for support, or to reduce task demands in real-time when students do not respond as intended (Charalambous, 2008; Tzur, 2008). Notably, teachers may not recognize their own propensity for this adjustment or may be unaware of alternative strategies to encourage student independence.

The mathematical challenge is also compromised when students are not held accountable for high-level cognitive processes and products (e.g., Akcil-Okan & Tekkumru-Kisa, 2021). For example, the challenge is diminished when they are not required to articulate their reasoning, or when their incomplete or incorrect explanations are not used as springboards for deeper mathematical discussion (Doyle, 1988; Smith & Stein, 1998). Moreover, the suitability of the task to the student group is crucial; tasks that are overly complex, or come with unclear expectations and instructions, can result in a lack of student engagement due to diminished interest, motivation, or inadequate prior knowledge. Ultimately, the essence of mathematical challenge lies not just in the complexity of tasks but in the cultivation of a classroom ethos where students are expected—and feel compelled—to engage critically with mathematical concepts, defend their reasoning, and learn from their misconceptions, thereby fostering a resilient and self-sustaining community of learners.

Additionally, in classrooms where management issues prevail, sustaining student engagement in high-level cognitive activities becomes a challenge (Smith & Stein, 1998). Effective classroom management is crucial for keeping students on-task (Korpershoek et al., 2016; Lipowsky et al., 2009; Seidel & Shavelson, 2007). However, resistance and disruptive behavior are common in many classrooms (Schwarzer & Hallum, 2008), and such disciplinary problems can interfere with teaching, thereby reducing students' mathematical engagement and hindering teachers' efforts to maintain task demands (Henningsen & Stein, 1997). Consequently, when teachers are preoccupied with managing behavior, maintaining a focus on mathematics becomes increasingly difficult. It is essential to recognize that teachers must strike a delicate balance between fostering a positive classroom climate and upholding the rigor of mathematical challenges (Anagnostopoulos et al., 2020).

Moreover, time constraints also play a pivotal role, particularly when students are allotted insufficient time to delve into the challenging aspects of a task, or conversely, too much time, leading to aimless exploration or off-task behavior. Such inappropriate allocation of time often stems from established classroom norms where the emphasis is placed on obtaining the correct answers and covering the curriculum (teaching the 'correct concepts'), rather than on constructively addressing student errors to promote deeper understanding and conceptual learning (Smith & Stein, 1998). The urgency to cover the curriculum acts as a contextual pressure that can inadvertently deprioritize the importance of engaging with the material at a meaningful level.

2.2.6 The Need for and the Role of Professional Learning and Development in Supporting Teacher Efforts to Maintain Cognitive Activation

Research suggests that PLD tailored to strengthen specific areas of teachers' Pedagogical Content Knowledge (PCK) can be instrumental in equipping them with the understanding necessary to sustain high cognitive demand in their teaching (Borko, 2004). PLD programs that concentrate on the planning, selection, and implementation of mathematically challenging tasks have been shown to reinforce teachers' skills and practices, thereby supporting the preservation of task complexity (Boston & Smith, 2011; Foley et al., 2012; Kang, 2017). This section delves into the necessity and impact of PLD in aiding teachers to navigate the complexities of working with students on mathematically challenging tasks.

Several classroom observation studies showed that in a significant number of lessons, the level of mathematical task challenge was diminished (e.g., Arbaugh & Brown, 2005; Boston & Wilhelm, 2017; Choppin, 2011; Stigler & Hiebert, 2004; Silver et al., 2009; Stein et al., 2008; Stein & Kaufman, 2010). Maintaining or raising the challenge levels lies on teachers' shoulders, who need better preparation and support to fulfill this role effectively (Resnick & Zurawsky, 2006). Mason (2002) argued that teaching experience does not automatically translate to teacher learning; rather, it is the deliberate inquiry and reflection upon one's teaching practice that cultivates professional growth. Especially in today's educational landscape, where most reformed curricula incorporate challenging tasks, it is essential for practicing teachers to receive PLD to effectively implement these tasks into their daily teaching (e.g., Arbaugh & Brown, 2005; Arbaugh et al., 2006; Boston & Smith, 2009, 2011; Boston & Wilhelm, 2017; Marrongelle et al., 2013; Roth McDuffie & Mather, 2006).

A recent literature review by Hsu and Yao (2023) on teacher PLD highlighted the fruitfulness of task-centric PLD programs in facilitating teachers to improve their teaching. Effective PLD programs need to be thorough, focused, systematic, and intensive, encouraging problem-solving and reflective discussions on task challenges and classroom dynamics that influence task implementation (Jackson & Cobb, 2010; Stein et al., 2009). The characteristics of effective PLD will be further discussed in Section 2.5.2.

Various PLD tools and materials have been developed to support teachers while working with their students on challenging tasks. These include, the TAG (Smith & Stein, 1998), the MTF (Stein, & Smith, 1998), Critical Elements for Setting Up Complex Tasks (Jackson et al., 2012), Five Practices for Orchestrating Class Discussions (Smith & Stein, 2011), Thinking Through a Lesson Protocol (Smith et al., 2008), and Narrative cases of mathematics instruction (Smith et al., 2009). The previously mentioned TAG and MTF, for instance, have served effectively both as research tools for assessing teaching quality and as PLD resources, heightening teachers' awareness of potential task modifications during teaching and maintaining task complexity (Boston & Smith, 2011; Hsu & Yao, 2023; Smith & Stein, 2023).

Some examples of early and influential multiyear PLD initiatives with a focus on such issues were the *Quantitative Understanding: Amplifying Student Achievement and Reasoning* project (e.g., Silver & Stein, 1996; Smith & Stein, 2023), *the Cognitively Guided Instruction* project (e.g., Carpenter & Fennema, 1992; Franke et al., 2001; Knapp & Peterson, 1995), the *Effective Schools Project* (e.g., Boston & Smith, 2011; Smith et al., 2009), and *Beyond Implementation: Focusing on Challenging and Learning* project (e.g., Silver et al., 2007; Silver et al., 2011). These projects contributed to a body of evidence supporting the idea that promoting cognitive activation in classrooms relies on a deep understanding of how students think and learn, and that PLD programs must be designed to support this complex understanding. Along with subsequent research on these issues, they provided more knowledge on how to scaffold teachers to enact cognitively activated teaching. First, they emphasized the importance of implementing a task-centric PLD approach (Boston & Smith, 2011). Second, they stressed the use of powerful tools, materials, or practice artifacts that provide opportunities for inquiry and reflection on the tasks and the work around them, and connections to different aspects of MKT (Silver et al., 2007). Third, they highlighted that a practice-based PLD program should focus teacher attention on key practices for maintaining the task demand and attending to student thinking (Boston & Smith, 2009). Fourth, they indicated that teachers could navigate challenges related to the use of challenging tasks, by reconsidering what it means to make progress with a task or to be capable in mathematics (Bobis et al., 2016; Russo et al., 2019). Fifth, they confirmed the relevance of teacher conceptualizations and practice (Boston 2013; Voss et al., 2013). Finally, they empirically corroborated the hypothesis that such programs could facilitate teacher learning, with respect to developing sensitivities in identifying key elements of cognitive activation, and hence, responding to student thinking and sense-making and interpreting student understanding (Boston, 2012; Choppin, 2011; Heyd-Metzuyanim et al., 2020; Mason, 2002; Ponte & Quaresma, 2016).

Addressing teacher PLD on cognitive activation is critical, given that the field of tasks and their use is still fertile. Tekkumru-Kisa and colleagues (2020) stressed a number of reasons for still keeping the focus of research on tasks in mathematics classrooms, including that focusing on tasks can help researchers understand how teachers can be supported to attend to student thinking. Research gaps in PLD for cognitive activation include understanding teachers' challenges during PLD (Hsu & Yao, 2023), fostering active learning, reflection, peer observation in a supporting learning community with skilled facilitation (Ghousseini & Kazemi, 2023), customizing PLD to address classroom challenges with demanding tasks (Superfine & Superfine, 2023), and creating coherent learning sequences for teachers that "build on prior learning experiences" (Cohen et al., 2022, p.20). The common thread in these gaps is the need for dynamic, interactive, and context-sensitive PLD, that bridges theory with practice, emphasizing continuous, collaborative learning that is deeply rooted in the realities of contemporary teaching.

2.3 Differentiation

2.3.1 Defining Differentiation

Differentiation, as defined by Carol Ann Tomlinson (2017)—a leading expert in this field—is a philosophy and practice of teaching tailored to meet the diverse needs of students. This approach invites teachers to differentiate the lesson content, the learning process, the final product, and the learning environment, aiming to help students reach their full potential by acknowledging and accommodating their differences in readiness, ability, prior knowledge and skills, interests, and learning preferences[10] (Konstantinou-Katzi et al., 2011; 2013; Tomlinson, 2005b; 2008; Tomlinson & Imbeau, 2023; Valiandes et al., 2011; Van Geel et al., 2019). The concept of differentiation in education is encountered in the literature with different terms, such as 'curriculum differentiation' (e.g., Marishane et al., 2015), 'differentiated teaching' (e.g., Kokkinos & Gakis, 2021), 'differentiated instruction' (e.g., Gibbs & McKay, 2021), 'pedagogical differentiation' (e.g., Kanellopoulou & Darra, 2022), and 'differentiated learning' (e.g., Koutselini & Agathangelou, 2009).[11] While overlapping in their goal to cater to diverse learning needs, these terms offer a slightly different lens through which differentiation in education can be understood and applied. Recognizing that each term varies in focus, in this study, 'differentiation' or 'differentiated teaching' are used to focus on the proactive and responsive teaching practices that cater to students' needs.

---

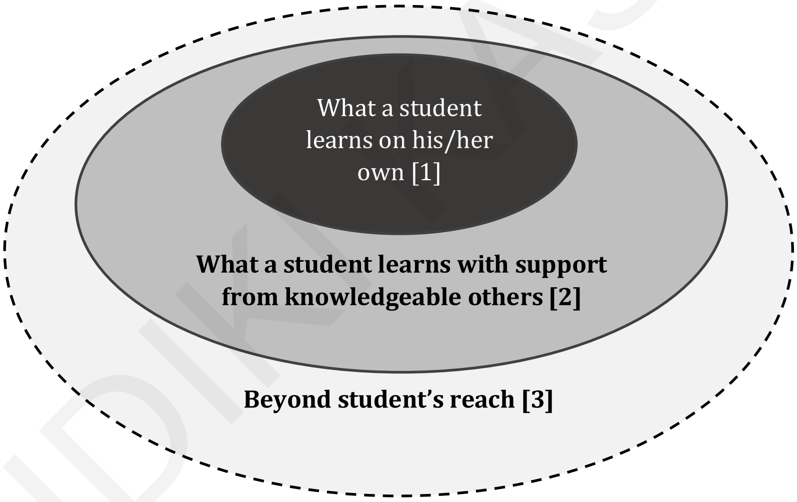
[10] These concepts are revisited and defined in the next section.

[11] Curriculum Differentiation centers on adapting curriculum content; Differentiated Teaching emphasizes varied teaching strategies; Differentiated Instruction focuses on instructional design and delivery; Pedagogical Differentiation merges curriculum and instructional approaches with a broader pedagogical view; and Differentiated Learning shifts focus on the learners' individual styles and preferences.

The core principle of differentiation is to prioritize *the* student at the center of all classroom decisions in contrast to treating *students* as a homogeneous entity (Tomlinson & Imbeau, 2023). Differentiation is described as "a teaching practice with a balanced emphasis on individual students and the class as a whole" (Tomlinson & Imbeau, 2023, p.30). It is grounded in the theory of the Vygotskian *Zone of Proximal Development* (1978), which posits that optimal learning occurs when students are engaged in learning experiences *slightly* beyond their current abilities (i.e., comfort zone, see Figure 7). In the realm of differentiated teaching, this theory places a responsibility on teachers to ensure that each student comprehends and masters the core ideas, processes, and skills of the lesson, by offering appropriate scaffolds.

Figure 7

Schematic Illustration of a Student's Zone of Proximal Development



In particular, students bring a spectrum of differences and zones of proximal development, constructing knowledge in their own way. Figure 7 illustrates student learning variability in three layers: Layer 1 represents tasks and concepts within a student's current knowledge and skill set, manageable without teacher support. Layer 2 covers tasks and concepts *learnable* with appropriate teacher or peer guidance, falling within the student's Zone of Proximal Development. Layer 3 includes tasks beyond the student's immediate grasp, requiring significant scaffolding and support. In practice, this involves dynamic and iterative cycles of *assessing and understanding* the diverse learning needs and experiences of students, while *continually evolving* lesson plans to effectively modify teaching and ensure that each student is effectively connected to essential content (Koutselini, 2001; Tomlinson & Imbeau, 2023). In essence, differentiation is committed to "ensuring that every student experiences the maximum opportunity to learn" (Anthony et al., 2018, p. 117), having unhindered access to an excellent curriculum, effective teaching, appropriate assessment, and necessary support (Tomlinson & Imbeau, 2023).

Delaney (2017) identified two broad ways of understanding and approaching differentiation. The first involves teachers assessing students to tailor their teaching to *each* student's current attainment and needs. This approach views differentiation as a process of *matching* learning targets, tasks, activities, resources, and support with each learner's unique needs, styles, and learning pace (Beltramo, 2017; Stradling & Saunders, 1993). The second approach perceives student differences—such as ideas, solution methods, and contributions—as *valuable resources*, fostering rich classroom discussions during which students share insights, compare, and connect ideas, and freely express questions or confusions to their peers. Teachers anticipate diverse student responses to new concepts, incorporating different explanations and representations.

Both approaches align with Tomlinson's view of differentiation as a proactive and organized, yet flexible, method of adapting teaching methods and resources. Also, both approaches underscore that differentiation is more about qualitative adjustment than quantitative measures. It is not simply about assigning more or less work, but about tailoring the educational experience to the unique needs and abilities of each student. However, a key distinction between these approaches lies in the perception of student diversity. In the first approach, student diversity is often viewed as an impediment, complicating the teaching process (especially the idea of 'matching'), leading some teachers to perceive differentiation as a challenging, if not unattainable, teaching goal; in the second, student diversity is seen as an asset, as an opportunity to enrich the learning experience and as integral to a dynamic and engaging teaching process (Delaney, 2017).

In sum, differentiation is considered a complex teaching skill, incorporating a wide array of teaching practices to address diverse needs (Van Geel et al, 2019). In a differentiated classroom, students gain self-direction and autonomy in their learning and the teacher flexibly coordinates and adjusts all other aspects of teaching that influence student learning, such as teaching time, space, materials, choice, student groupings, learning goals, tasks, communication with/among students, teaching techniques and strategies, resources, support, or pace (Delaney, 2017; Konstantinou-

Katzi et al., 2013; Tomlinson, 2017). In this respect, differentiation is part of teaching quality (Creemers & Kyriakides, 2006; Van Geel et al., 2019), making students feel included in the work of the class and experience success (Parsons et al., 2018; Sullivan, 2011). Tomlinson and Imbeau (2023) nicely summarize the philosophy behind differentiation into a fundamental question that all teachers in contemporary mixed-ability classes should contemplate, pointing to the importance of not treating a classroom as a whole: "What does *this* student need at *this* moment in order to be able to progress with *this* key content—and what do I need to do to make that happen?" (p. 31). The upcoming section discusses the various classroom elements that can be differentiated and the methods for doing so.

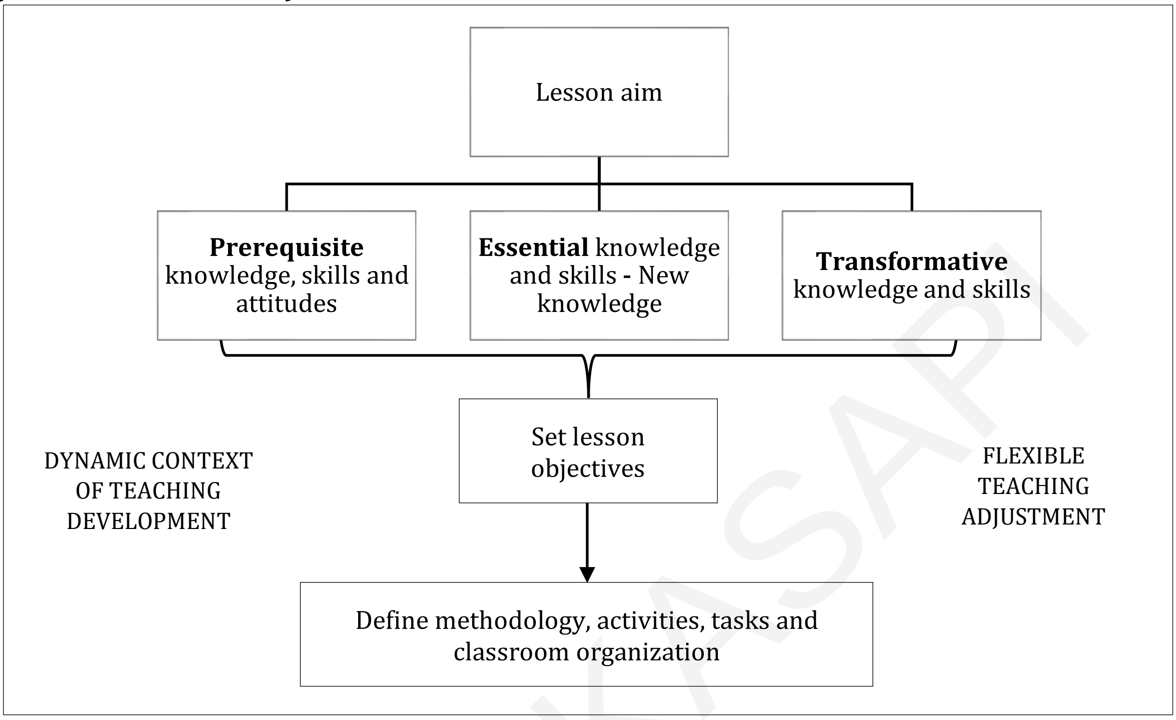
2.3.2 Differentiation as an Integral Part of Teaching

Differentiation plays a crucial role at every stage of the teaching process, from the initial planning of lessons to their actual implementation in the classroom (Valiandes et al., 2017). Although flexible, differentiation requires planning and organization, given the variety of techniques and strategies it offers (Palieraki & Koutrouba, 2021). In educational systems where the curriculum is strongly content-oriented and undifferentiated, Koutselini's model (2006), as depicted in Figure 8, supports teachers to focus on clear and precise learning goals and develop an effective differentiated lesson (Palieraki & Koutrouba, 2021).

Effective differentiation begins with identifying the lesson aim as described in the curriculum and adjusting it based on the pre-assessment of students' knowledge and skill status (Tomlinson et al., 2013). In the second stage, a thorough analysis of the curriculum or the lesson content follows. In particular, the teacher deconstructs the curriculum and material into *fundamental knowledge*, identifying key concepts, information, skills, processes, and strategies that students are expected to learn. This knowledge is then categorized and ranked into three types: *essential (or core), prerequisite, and transformative (or meta-schematic,* Valiandes et al., 2017).

Figure 8

Analyzing the Undifferentiated Curriculum and Planning for differentiation (adapted from Koutselini, 2006)



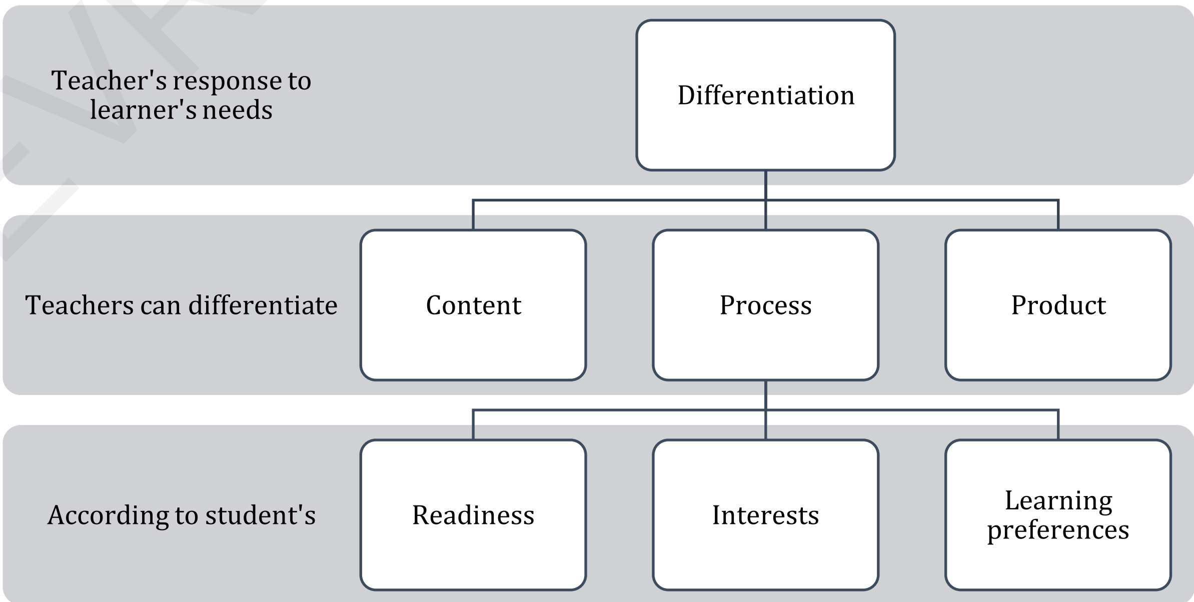
Essential knowledge encompasses new knowledge that *all* students need to learn by the end of the lesson. Prerequisite knowledge refers to the old knowledge that has previously been taught and must be in place for acquiring the essential knowledge. When this kind of knowledge is not in place the teacher must plan for "teaching backward" those students who lack old knowledge (Tomlinson et al., 2013). Transformative knowledge extends beyond the curriculum's basic expectations; it requires higher cognitive thinking, targeting more advanced students, such as gifted students or early finishers (Valiandes et al., 2017). An example of applying this framework, as described by Valiandes et al. (2017), involves a lesson on the distributive property of division (this is the lesson aim according to Figure 8). In this case, the essential knowledge would be the ability to perform divisions with a single-digit divisor using various strategies. The prerequisite knowledge might include refreshing how to perform multiplication and division as inverse operations. For transformative knowledge, several students might progress to creating and solving their own division problems.

In the third stage of planning differentiation, the teacher specifies the learning objectives and outcomes and translates these knowledge categories into ranked tasks or activities, organized from simple to complicated (Pham, 2011). Differentiated teaching offers a variety of techniques (such as entry and exit slips, learning braces, etc.) and strategies (hierarchical lesson activities, anchor activities, flexible grouping, etc.) that can be employed by teachers. While all students work with the same essential understandings (Tomlinson et al., 2013), there are varied levels of complexity and support. The learning environment is organized to balance autonomous work, collaborative activities, and whole-class interactions (Valiandes & Koutselini, 2009; Konstantinou-Katzi et al., 2013). Importantly, as shown in Figure 8, flexibility during lesson enactment is crucial. Teachers need to be prepared to adjust their plans and approaches in response to students' formative assessment and dynamics observed during the lesson (Valiandes, 2015). In all, differentiation is not just a preparatory step but an integral part of the teaching process, actively shaping the way lessons are delivered and experienced by students.

Tomlinson's (1999) seminal work on differentiation highlights that teachers can differentiate at least four key elements in the classroom: content (the 'what'), process (the 'how'), product (the form of 'evidence'), and affect/learning environment ('how the classroom works and feels'). Presented below in Figure 9 is a visual representation of Tomlinson's model of differentiation, illustrating the dynamic ways in which teachers can tailor classroom content, processes, products, and learning environments to meet the diverse needs and abilities of their students.

Figure 9.

A Modified Version of Tomlinson's (1999) Model of Differentiation



Differentiation by content acknowledges that differentiated teaching often involves altering the methods or materials through which students access information or adjusting the content itself (i.e., the material and how it is presented to students, Taylor, 2015). While maintaining the same core learning objective(s) for all students, this approach recognizes that some students might need to focus on prerequisite content to build foundational skills before moving forward, whereas advanced students may be ready to tackle more complex, transformational topics ahead of their classmates (Palieraki & Koutrouba, 2021). Differentiation by content allows teachers to address the diverse range of abilities and knowledge levels within a classroom, ensuring that all students continue to progress in their learning. An example of differentiating by content is the one explained above with the prerequisite and transformative knowledge for division by a single-digit divisor.

Differentiation by process, on the other hand, pertains to the diverse activities, techniques, and strategies students use to engage with and make sense of the content (Tomlinson, 2017). For instance, when teaching division with a single-digit divisor, one group of students might use diagrams, drawings, or representations to conceptualize how division works. Another group might use physical objects like counters or blocks. This hands-on approach allows these students to physically divide objects into groups, offering a tangible understanding of how division operates. A third group could apply division in real-life scenarios or through problem-solving activities to understand the practical utility of division. All three groups of students will come to realize how division works.

Differentiation by product involves diverse ways for students to showcase their learning, understanding, and skills (Tomlinson, 2005b). It serves as an authentic, comprehensive assessment that requires students to apply, extend, or transfer their acquired knowledge and abilities over a period of time. For example, when teaching division with a single-digit divisor, a student group might create a short story or word problems involving single-digit division scenarios, while a second student group could create a poster or infographic that visually explains the division process, including examples with single-digit divisors. For those who enjoy hands-on projects, constructing a physical model that involves division with single-digit divisors could be an engaging way to demonstrate understanding. In each case, the teacher would define and communicate the standards for what makes a qualitative problem story, poster, or model, ensuring that students meet the learning objectives in their own unique way

(Tomlinson & Imbeau, 2023). Also, the teacher monitors their progress on the products and provides students with feedback (Palieraki & Koutrouba, 2021).

Finally, *differentiation by learning environment* entails adjusting the classroom setting, its physical layout, and classroom management, to accommodate various student needs and ensure a safe, respectful, supportive, and positive learning atmosphere (Kronberg et al., 1997; Tomlinson & Imbeau, 2023). Such an approach might involve creating tasks that reflect the varied cultural and familial backgrounds of students, accompanied by well-defined routines and guidelines for independent or group work (Valiandes & Neophytou, 2018). Furthermore, a well-designed classroom promotes learning through the use of vibrant colors, visual cues, readily accessible materials, student-centered layouts, spaces for individual work and collaboration, and furniture arrangements that facilitate peer interaction (Gibson & Hasbrouck, 2008). Moreover, acknowledging the wide range of student learning preferences is crucial: while some students may excel in a quiet, static environment, others may benefit more from a dynamic setting where physical movement is integral to their learning process. Tomlinson and Allan (2000) described the ideal learning environment as a "good cafeteria" offering both essentials and a variety of choices to suit individual preferences (p.26).

For instance, in teaching division with a single-digit divisor, differentiation by learning environment can be exemplified as follows: The classroom can feature various workstations to suit different learning styles: a quiet area with division worksheets for individual work, a collaborative space for group problem-solving, and a computer station for interactive division games. This arrangement allows students to choose their preferred learning environment, enhancing their comfort and effectiveness in mastering division while fostering positive learning attitudes and motivation.

Thus, in a differentiated classroom, students are provided with multiple avenues to receive information, process ideas, and express their learning through a range of products (Tomlinson, 2017). However, it is important to recognize that differentiating every aspect—content, process, product, and learning environment—in each lesson is neither mandatory nor always feasible. Differentiation should be strategically tailored to meet students' specific readiness levels, interests, or learning preferences in each situation. It is the intentional, teacher-driven practice aimed at moving beyond generic teaching that caters to the average (i.e., 'teaching-to-the-middle' or 'one-size-fits-all' teaching, Bondie et al., 2019; Tomlinson, 2015). Santangelo and Tomlinson (2012) further elaborate on these concepts by offering detailed descriptions of student readiness, interest, and learning preferences, as briefly detailed below.

Readiness refers to a student's current level of preparedness or proximity to the knowledge and skills being taught. It is distinct from intellectual ability and aligns more closely with the Vygotskian perspective, which emphasizes the role of a supportive environment in advancing student learning. To optimize learning, students should engage with tasks and materials that are appropriately challenging—neither too difficult nor too easy—matching their current level of proficiency with the concept being taught. This approach ensures that the learning experiences are suitably tailored to foster growth and understanding.

Interest relates to the specific topics or methods that ignite a student's curiosity, motivation, and engagement in learning. This interest often stems from the student's cultural context, personal experiences, individual questions, and inherent strengths. Engaging students in areas of personal interest can significantly enhance their educational experience. It encourages them to exert more effort, persist longer in challenging tasks, think more critically, and achieve a deeper understanding of the subject matter.

Learning preferences refer to the most natural and efficient ways in which students most effectively and comfortably learn. These preferences encompass various factors, such as preferred group settings (whether they learn best in whole class environments, small groups, pairs, or individually), cognitive styles (such as favoring inductive or deductive reasoning, linear or nonlinear thinking, reflective or action-oriented approaches, and collective or individualistic learning tendencies), and environmental considerations (including preferred noise levels, activity levels, and the physical arrangement and choices of furniture in the learning space). In the framework proposed by Tomlinson, student differences are viewed not as static traits inherent to the students themselves, but as dynamic and changeable attributes (Smets et al., 2020).

This section illuminates the essence of differentiation in education, integrating key concepts from Tomlinson's and Koutselini's frameworks. Koutselini's model, particularly, emphasizes the significance of clear, precise learning goals and the analysis of knowledge, supporting teachers in developing effective, student-centered lessons. Alongside, Tomlinson's approach underlines customizing content, processes, products, and learning environments according to students' readiness, interests, and learning preferences, acknowledging the dynamic nature of these attributes. The combined insights from these researchers advocate for a responsive, inclusive educational practice.

2.3.3 Common Misconceptions about Differentiation

Despite significant work that has been developed in differentiation, practitioners often hold several misconceptions. These misconceptions are counterproductive to differentiation and can lead to practices that do not serve all students effectively (Gibbs & McKay, 2021 Tomlinson et al., 2008). Table 2 outlines some common misunderstandings about differentiated teaching along with the corresponding clarifications that debunk these myths.

Table 2

Debunking Common Misconceptions About Differentiation (adapted from Tomlinson et al., 2008, p.4-5)

What Differentiation Is Not	What Differentiation Is
Just for students at the (extreme) ends (e.g., special education or gifted students)	For every student
Something extra in the curriculum	At the core of effective planning and teaching
An approach that mollycoddles students—makes them dependent	Supporting students in achieving at a level higher than they thought possible
Incompatible with standards	A vehicle for ensuring students' success with standards
Use of certain teaching strategies	Use of flexible approaches to space, time, materials, groupings, and instruction
Tracking in the regular classroom	The antithesis of tracking
Assigning students to cross-class groups based on assessment data	Within a classroom
All or mostly based on a particular approach to multiple intelligences	Systematic attention to readiness, interest, and learning preferences
All or mostly based on learning preferences	Systematic attention to readiness, interest, and learning preferences
Synonymous with student choice	A balance of teacher choice and student choice
Individualization	Focused on individuals, small groups, and the class as a whole
More tasks, books, or questions for some students and fewer for others	Varied avenues to the same essential understandings
Something a teacher does because it is the thing to do	Something a teacher does in response to particular needs of particular human beings
It happens all day every day	It happens when there is a need for it
Something a teacher does on the spot when it becomes evident that a lesson is not working for some students (reactive or improvisational)	Something a teacher plans before a lesson based on assessment evidence of student needs (proactive)

This table[12] can be used to dispel misconceptions about differentiation for teachers who may not be fully acquainted with the concept. Common misconceptions include the belief that differentiation is a passing trend that will eventually fade away; means crafting individualized lesson plans and teaching for each student in a class of more than twenty students; requires setting different learning goals for different students; creates chaos and a loss of classroom control; is simply another term for creating homogenous groups; involves only minor adjustments to lesson plans, such as providing marginally more complex tasks for advanced students or slightly easier ones for those struggling; is a set of predetermined teaching strategies, like a formulaic recipe; is a method to compensate for student deficits or disabilities; or is exclusively for students at the extreme ends of the learning spectrum, such as those with disabilities or those who are gifted (Birnie, 2015; Gaitas et al., 2022; Gibbs & McKay, 2021; Tomlinson, 2017). Furthermore, there is a bias among teachers to focus on differentiation by output—expecting different outcomes from different students—and by task, assigning different tasks such as drill-oriented activities for less advanced students and problem-solving for more advanced ones, rather than by engaging in varied processes, adjusting the learning pace, or having targeted discussions with students about their learning (Stradling & Saunders, 1993).

Table 2 emphasizes that differentiation is about providing appropriate levels of support to meet the needs of each student, which may change depending on content, time, and context. Differentiation is not about individualizing every aspect of teaching, but rather about engaging students with the core curriculum through varying levels of complexity and support systems. It emphasizes that neither doing less of what they find difficult helps struggling students, nor does doing more of what they already understand benefit advanced learners. The work is adjusted for different student groups to offer diverse paths to key concepts and objectives (Konstantinou-Katzi et al., 2013; Schwab et al., 2019). Differentiation values flexibility, such as through flexible grouping patterns, ensuring that students are not confined to working only with peers of similar ability but instead experience different collaborative settings based on readiness, interest, and learning styles. Besides, attention to learning preference can be helpful but should not overlook other needs, like readiness. Teachers should sometimes assign tasks to students to help them move forward, while at other times,

[12] This table is designed to illuminate and rectify some of the most prevalent misunderstandings about differentiation. It is not intended to be comprehensive, but rather to shed light on and address widespread misconceptions.

allow student choice. Differentiation is a responsive and proactive approach that becomes impactful when informed by continuous assessment data toward the teaching goals.

These misunderstandings can often lead teachers to view differentiation as an overwhelming and unrealistic approach, particularly when they feel torn between the pressure to cover the curriculum and the desire to cater to diverse student needs (Erotocritou-Stavrou & Koutselini, 2016). The issue often stems from a technical analysis of the term 'differentiation' rather than an approach centered on delivering a meaningful, concept-driven curriculum accessible to all learners (Hertberg-Davis, 2009). The forthcoming section delves into the necessity of differentiated teaching and its beneficial impact on student achievement across various domains.

2.3.4 Why Differentiating Teaching

Education systems globally are not adequately serving many students, as suggested by indicators like attendance, achievement, and dropout rates (Tomlinson et al., 2008; Valiandes et al., 2011). One in five students lacks basic skills in key areas like mathematics, with disadvantaged backgrounds doubling the risk of underperformance, despite equal classroom time (OECD, 2017). In contemporary classrooms, student populations are becoming increasingly diverse in academic, cultural, emotional, behavioral, and social aspects, along with learning preferences, interests, readiness to learn, motivation, personal experiences, and life circumstances (Delaney, 2017; Gibbs & McKay, 2021; Hertberg-Davis & Brighton, 2006; Tomlinson, 2015). This diversity includes a range of exceptionalities recognized in special education, as well as students with varied life experiences, such as migration. Notably, despite the relatively uniform age range within a single typical classroom, the range of student achievements may cover a wide spectrum (Hertberg-Davis & Brighton, 2006; Tomlinson, 2005a). Teachers observe a broad spectrum of learning capabilities: some students face challenges, others exceed grade-level expectations, and many fall somewhere in between (Delaney, 2017). In their daily lessons, teachers face the reality that students assimilate content at different speeds—some with ease, others requiring more time and assistance. The concept of differentiation arises from the need to address the wide range of differences and complexities evident in any classroom (Ollerton, 2009).

In the past few decades, a growing number of experimental and quasi-experimental studies has been conducted, investigating the effectiveness of several aspects of differentiated teaching, after scholarly calls to empirically validate the theoretical claims of differentiation (e.g., Hall, 2002). The majority of the investigations on the effectiveness of differentiation on student outcomes have mainly focused on specific student groups—gifted or disabled (e.g., Geisler et al., 2009; Jones et al., 2012; Lawrence-Brown, 2004; Tieso, 2005); on a small number of students in a specific area of a subject (e.g., reading comprehension strategies: Baumgartner et al., 2003; reading ability: Jones et al., 2012); on teachers' perceived practices and attitudes towards differentiation (e.g., Johnsen, 2003; Smit & Humpert, 2012; Santangelo & Tomlinson, 2012; Tomlinson, 2001); or on students' perceptions towards differentiation and their teacher's differentiation practices (e.g., Abell et al., 2011; Kronborg et al., 2008).

Following the initial research efforts, a significant shift occurred, with studies consistently demonstrating small to moderate positive effects of differentiation on students-as-a-whole' cognitive and affective outcomes across various educational settings and subjects (Deunk et al., 2018, $d = +0.146$; Smale-Jacobse et al., 2019, $d= +0.509$). Evidence of this impact has been found in language arts and literacy (Firmender et al., 2013; Goddard et al., 2015; Guay et al., 2017; Valiandes, 2015; Valiandes et al., 2011), mathematics (Bal, 2016; Goddard et al., 2015; Kim, 2005; Muthoni & Mbugua, 2014; Prast et al., 2018), and science (Simpkins et al., 2009). These beneficial outcomes have been validated at different educational levels, including pre-primary (Kotob & Jbaili, 2020), primary (Baumgartner et al., 2003; Geisler et al., 2009; Goddard et al., 2015; Tieso, 2005; Valiandes, 2015), secondary (Baumgartner et al., 2003; Muthoni & Mbugua, 2014; Simpkins et al., 2009), and tertiary (Konstantinou-Katzi et al., 2012) education. The positive impacts were noted for both practicing (Valiandes, 2015) and prospective teachers (Tulbure, 2011).

The compendium of evidence underscores a pivotal truth: differentiated teaching is not an educational trend, but an empirically validated teaching approach that provides a valid response to the inherent diversity within the classrooms. It stands as a beacon for an inclusive, equitable approach to teaching—a means to bridge the widening chasm of educational achievement and address the multifaceted nature of student disengagement. In light of these findings, differentiation emerges as a critical pedagogical practice, integral to the evolution of teaching and the fulfillment of education's promise to serve every learner. Research also indicated factors and conditions to ensure the effectiveness of differentiation, which are presented below.

2.3.5 Factors Supporting Teachers in Differentiating Their Teaching

The literature review revealed several factors that can support teachers in differentiating their teaching to reach as many students as possible[13]. Those favorable factors can be grouped into eight categories, and are briefly presented below: (a) having sufficient knowledge about the students and the subject matter, known as Pedagogical Content Knowledge; (b) holding positive beliefs about their students' abilities and learning, and the value of differentiation itself; (c) good lesson planning; (d) knowing and employing different strategies for differentiation; (e) knowing and applying various differentiation techniques[14]; (f) assessing student learning consistently and thoroughly; (g) establishing a supportive classroom environment; and (h) receiving appropriate support from people outside the classroom.

A certain type of teacher knowledge is critical for effective differentiation: Pedagogical Content Knowledge (PCK). High PCK allows teachers to provide a quality curriculum and materials, foster conceptual understanding and knowledge transfer, facilitate high-level discussions, help students make connections, and support the creation of quality work (Tomlinson, 2008; Van Geel et al., 2019). Teachers with robust PCK can align content with diverse learning styles and levels, appreciating students' prior knowledge and conceptual frameworks (Carolan & Guinn, 2007). Understanding students encompasses their abilities, learning challenges, pedagogical needs, interests, peer relationships, and motivation. Such insight enables teachers to tailor instruction effectively (Van Geel et al., 2019).

Of equal importance are teachers' positive beliefs about their students' abilities and learning, and the value of differentiation itself. Research indicates that constructivist beliefs about teaching in mixed-ability classrooms are key to implementing differentiated instruction effectively (Aftab, 2015; Dijkstra et al., 2017; Kotob & Jbaili, 2020; Pozas et al., 2020; Robinson et al., 2014; Van Geel et al., 2022). By primarily using teacher self-reported data, these studies reveal that a positive mindset

[13] This outlined set of factors, supporting or hindering differentiation, is not exhaustive but is part of a broader mosaic of elements influencing the successful implementation of differentiation.

[14] A strategy is a detailed plan of action designed to achieve a particular purpose, whereas a technique is a method or tool for carrying out a specific activity or task that needs skill (Cambridge Dictionary, n.d.).

and attitude towards differentiation correlates with effective planning and implementation. Teachers with such attitudes believe in their pivotal role in each student's success and are committed to achieving it (Tomlinson et al., 2008).

Aside from teachers' conceptualizations and beliefs, a third group of factors is related to good lesson planning. Effective differentiation involves understanding and analyzing the content's prerequisite, core, and transformative knowledge; challenging each student appropriately; using engaging tasks and questions that offer choice; and considering students' task preferences. These proactive steps help teachers tailor their teaching to support all students in mastering the content (Koutselini, 2006; Tomlinson, 2005a).

Effective differentiation is facilitated through a range of research-based teaching strategies (Brimijoin, 2005; Jordan et al., 2009; Tomlinson, 2005a). For instance, structuring lesson tasks hierarchically allows students to progress from familiar to new or even transformative concepts, prioritizing tasks from simple to complex (Valiandes et al., 2017). This structure necessitates clear lesson aims and the flexibility for students to work asynchronously, accommodating different paces of learning. *Tiered activities* are another way for differentiating 'horizontally': all students will come away with key understandings while working on adjusted versions or resources of the same task, varied by level, structure, pace, complexity, or creativity, intended for on-level, struggling or advanced students (Allen & Turville, 2010). Besides tiered activities, the teacher can use some anchor activities, to maximize the teaching time for (group of) students, allowing them to review, practice, or extend learning aligned with lesson objectives. Flexible grouping is another strategy, forming temporary and varying groups during the lesson, based on students' learning preferences, interests, and readiness, ensuring dynamic and responsive classroom engagement (e.g., Guay et al., 2017).

Literature on differentiation also outlines numerous techniques that cater to diverse student needs. One such technique involves the use of entry and exit cards, which are quick, informal methods for formative assessment of students' readiness, knowledge, and skills in a particular area (Tomlinson et al., 2013). Entry cards help identify students' misconceptions or areas where prerequisite knowledge may be lacking, while exit cards gauge their learning and ability to apply new knowledge, indicating if further review is necessary. These cards, which may contain open or closed-ended questions, inform the teaching process for both the current and subsequent lessons. Another differentiation technique is the provision of learning braces, such as bookmarks, formula sheets, cardboards, or posters, which serve as handy references for students as they tackle specific tasks or when they need a refresher on prior knowledge or skills (Palieraki & Koutrouba, 2021). Finally, 'Think-Tac-Toe', a variation of tic-tac-toe, is an engaging technique that presents students with a grid of tasks varying in content, process, product, or tailored to individual needs (Dotger & Causton-Theoharis, 2010). Students may be asked to complete a line of tasks (e.g., in a nine-square grid), allowing advanced learners to delve deeper and others to consolidate their understanding. These practices, along with the strategic use of diverse materials, flexible pacing, and fostering of independent, personalized learning, are central to the effective implementation of differentiated teaching.

One of the sine-qua-non elements of effective differentiation is the consistent and thorough pre- and ongoing assessment of student learning (Tomlinson et al., 2008). Pre-assessment determines students' starting points, and ongoing assessment tracks progress, clarifying misunderstandings and guiding adjustments in teaching methods and materials. Assessment functions as a tool for refining teaching strategies, allowing for targeted instruction that meets students' needs, and facilitating necessary revisions to tasks or time allocations (McQuarrie & McRae, 2010). Students thrive when engaged with materials that present the right level of challenge, when their interests are connected to learning outcomes, and when their learning preferences are considered (Beecher & Sweeny, 2008; Santamaria, 2009; Sullivan, 2011; Tomlinson, 2005a). In this respect, teachers must systematically observe and actively monitor their students' work to ensure learning is optimized.

Establishing a supportive classroom environment is pivotal for differentiation, fostering a space where diversity is respected and achievements are celebrated (Brimijoin, 2005). In such settings, students feel valued and understood, hold high expectations for themselves, and trust in their teacher's support for their individual and collective success (Tomlinson et al., 2008). A sense of community within the classroom, where trust and perseverance in learning are cultivated, has also been linked to higher student achievement (Goddard et al., 2015). A positive classroom community emerges when learning experiences are shared, students are heard with respect, and time is allotted for work with constructive feedback rather than just final grades. It is also crucial that students perceive the learning as relevant to their needs and that classroom routines bolster their educational journey. Through the accumulation of consistent, small, and positive interactions, students take greater ownership of their learning, leading to increased self-belief and a drive for excellence (Tomlinson, 2008).

Support from outside the classroom—from parents and school stakeholders—is crucial for teachers to differentiate teaching effectively (Beecher & Sweeny, 2008; Dijkstra et al., 2017; Dixon et al., 2014). For example, in a longitudinal PLD program studied by Beecher and Sweeny (2008), schools that exhibited a commitment to student success, parental involvement, and principals knowledgeable in differentiation saw enhanced student learning outcomes and a narrowed achievement gap. The subsequent section will discuss the factors that impede the implementation of differentiated instruction.

2.3.6 Factors Inhibiting Teachers from Differentiating Their Teaching

Teachers face a range of challenges when it comes to incorporating differentiation in their work. The factors that inhibit the effective implementation of differentiated teaching can be grouped into eight categories: (a) insufficient teacher content knowledge and pedagogical content knowledge; (b) counterproductive teacher conceptualizations, beliefs, and attitudes about learning and differentiation; (c) teacher lack of differentiation skills, techniques, tools, ideas, and strategies; (d) challenges in recognizing and reflecting on students' unique learning needs and readiness; (e) relying on a one-size-fits-all curriculum and struggling in finding and utilizing additional suitable resources; (f) limited time for preparation and planning; and (g) lack of administrative support for differentiation efforts.

A significant obstacle in employing differentiated teaching methods is teachers' insufficient subject-matter knowledge and Pedagogical Content Knowledge (PCK) (Van Geel et al., 2022; van Tassel-Baska & Stambaugh, 2005). Without a profound grasp of the content (CK), teachers may struggle to pinpoint the core concepts, ideas, and skills that students need to acquire, and effectively use the curriculum (Van Geel et al., 2019). This understanding is crucial for adapting lessons effectively, determining the sequence of content from basic to complex, and organizing tasks from simple to advanced levels (Brighton et al., 2005; Hedrick, 2012; Tomlinson et al., 2003). Mastery of PCK is essential for deconstructing and prioritizing lesson content effectively (Erotocritou-Stavrou & Koutselini, 2016; Van Geel et al., 2019). Without such knowledge, and especially the knowledge about students, teachers will face challenges in anticipating the difficulties their students will encounter during teaching; identifying the level and needs of their students; and analyzing student work, through monitoring their work and questioning (Van Geel et al., 2019).

Teachers' counterproductive conceptualizations, beliefs, and attitudes toward learning and differentiation present significant challenges. These entrenched beliefs are often resistant to change (Erotocritou-Stavrou & Koutselini, 2016; van Tassel-Baska & Stambaugh, 2005). Some teachers may lack prior experience or a clear understanding of differentiation (Lortie, 1977; Valiandes & Koutselini, 2009), while others may not recognize its necessity or even reject the responsibility of implementing it (Dixon et al., 2014). There are instances where teachers may resist assessing or adapting materials and methods to differentiate (Tomlinson et al., 2003) or may attempt to differentiate with minimal effort (Dijkstra et al., 2017). There is also a tendency to focus more on less advanced students, neglecting the needs of more advanced ones (Marishane et al., 2015; Ritzema et al., 2016). Moreover, the fear of losing control in the classroom, particularly among those with limited management skills, can deter teachers from providing students with autonomy (van Tassel-Baska & Stambaugh, 2005). These negative perceptions form formidable obstacles, hindering the implementation and effectiveness of differentiated teaching (Aftab, 2015; Dijkstra et al., 2017; Tomlinson, 2017).

Although teachers often intend to employ differentiation, they may lack the necessary pedagogical repertoire—comprising skills, tools, techniques, ideas, and strategies for student-centered teaching, like those described in the previous section (Adami, 2004; Hardré & Sullivan, 2008; van Tassel-Baska & Stambaugh, 2005). This gap makes it difficult for teachers to create lessons that accommodate individual student readiness, interests, and learning preferences, which is essential for successful differentiation (Erotocritou-Stavrou & Koutselini, 2016).

Additionally, teachers often encounter obstacles in identifying and reflecting on the distinct learning needs and readiness levels of their students, particularly those who are highly advanced. They also struggle to gauge the cognitive demands associated with specific curricular activities (Altıntaş & Özdemir, 2015; Brighton et al., 2005; Dijkstra et al., 2017; Erotocritou-Stavrou & Koutselini, 2016). One reason for this challenge is the lack of rich and actionable data on student performance and progress, which is crucial for informing targeted teaching approaches (Van Geel et al., 2019).

Incorporating differentiation into teaching is further complicated by curricula and textbooks that often lack provisions for varied learning needs, leaving teachers to make potentially unsuitable modifications (Haggarty & Pepin, 2002; Van Geel et al., 2019). Teachers frequently express the need for a structured curriculum that progressively builds upon prerequisite, core, and transformative knowledge, adaptable across different classes (Erotocritou-Stavrou & Koutselini, 2016). Without such a framework, teachers must individually tailor materials to meet diverse student requirements. Supportive curricula with clear content descriptions and task guidelines could aid teachers in adjusting content coverage and pacing (Harris, 2012). Yet, differentiation becomes exceptionally challenging when the curriculum is weak (i.e., a curriculum that does not provide the students with the necessary tasks and support in order to master knowledge; hence, students do not benefit from the content they are taught), leading to multiple variations of insufficient content delivery (Hedrick, 2012). Teachers also face the demanding task of sourcing or creating additional resources to fulfill their students' specific needs, which is a complex endeavor in itself (van Tassel-Baska & Stambaugh, 2005). A significant concern among teachers is identifying "what learning materials are suitable for which learning objectives, and how do [the teacher] determine[s] what each student is going to do" (related to assessment, Dijkstra et al., 2017, p.161).

Finally, a common issue for teachers is the lack of time necessary to effectively plan and implement differentiated teaching, even if they are equipped with differentiation tools (Van Geel et al., 2022). The process of assessing student needs, adapting teaching methods, creating tailored materials, and collaborating with colleagues demands a significant investment of time (Aftab, 2015; Bondie et al., 2019; van Tassel-Baska & Stambaugh, 2005). When coupled with large class size, inadequate parental and administrative support, and insufficient PLD in differentiation (see Section 2.3.7 below), these time constraints can severely hinder the practical application of differentiated teaching (Aftab, 2015; Garolan & Guinn, 2007; Resnick, 1987; Turner & Solis, 2017; Van Geel et al., 2019). This underscores the Interconnected nature of the challenges teachers face in executing differentiation strategies (Van Geel et al., 2019).

This section has outlined the significant challenges that complicate the understanding and implementation of differentiation in real classroom settings. The upcoming section delves into the role of PLD in equipping teachers with the knowledge and skills necessary to successfully differentiate teaching.

2.3.7 The Need for and the Role of Professional Learning Development in Differentiating Teaching

PLD programs focusing on differentiation have highlighted multiple benefits for teachers. In particular, PLD plays a vital role in raising teachers' awareness of and skills in recognizing and addressing student readiness and individual needs; it equips them with practices and strategies for planning and dynamically adjusting teaching methods and curricula (Dixon et al., 2014; Erotocritou-Stavrou & Koutselini, 2016; Van Tassel-Baska et al., 2008). Furthermore, PLD has been shown to boost teacher effectiveness and confidence in implementing differentiation (Gheyssens et al., 2020) and correct misunderstandings related to differentiation (Erotocritou-Stavrou & Koutselini, 2016). Also, teachers engaging in such PLD programs not only learn various practices but also actively experiment with them in their classrooms (Prast et al., 2018). The experience of PLD interventions is highly valued by teachers; it not only increases their implementation of differentiation strategies but also reinforces their commitment to continue these practices, especially upon observing the beneficial effects on their students (Valiandes & Neophytou, 2018). In essence, PLD emerges as a transformative tool in empowering teachers to differentiate their teaching and address the multifaceted needs of their students.

The research underscores the need for and importance of PLD programs for both prospective (e.g., Dack, 2018) and practicing teachers (e.g., Dixon et al., 2014; Nicolae, 2014) to improve their conceptualization and practice in terms of differentiation. Yet, starting from teacher preparation programs, initial prospective teacher PLD often touches on differentiation only superficially and theoretically (Dixon et al., 2014). Therefore, many teachers—especially novices or less experienced—feel unprepared for differentiation (Tomlinson, 2001). The PLD offered to practicing teachers on differentiation is often lacking or is based on "traditional, top-down, one shot, lecture approach seminars" (Valiandes et al., 2017, p.47). These glimpses at differentiation are not enough for putting it into practice. Thus, even highly experienced teachers are often not convinced to embrace differentiation; do not know how to implement it; and face challenges in introducing changes into their practice in terms of differentiation, remaining incapable of responding to student diversity (Bondie et al., 2019; de Jager, 2017; Marishane et al., 2015)

Tomlinson (2005a) argued that PLD relating to differentiation should move from "training via mass inoculation" (p. 11) to PLD that is reflective; informed from current research trends; diagnostic (i.e., ensures that teachers develop the intended skills); connective; application-oriented; problem-focused; quality-concerned; collaborative; supportive; sustained; and differentiated itself. These PLD characteristics were confirmed by a number of studies that employed effective teacher PD programs (e.g., Prast et al., 2018; Valiandes et al., 2017; Van Tassel-Baska et al., 2008). Researchers also stressed the need for monitoring classroom implementation to track the frequency and quality of differentiated teaching during the PLD program (e.g., Dixon et al., 2014; Valiandes et al., 2017; van Tassel-Baska et al., 2008). PLD in differentiation should not only introduce the concept but also enable teachers to practice these strategies and receive feedback; foster reflection; educate teachers on the nuances of differentiation; encourage peer observation, feedback, and collaboration on lesson development; and guide teachers in analyzing learning goals, assessing student needs, and adjusting teaching accordingly (Brimijoin, 2005; Dixon et al., 2014). Teachers reported that next to gaining experience and sufficient time for differentiation, an important factor for developing differentiation skills seems to lie in developing communities of practice[15] (Van Geel et al., 2022). This structured collaborative environment could enable teachers to share their learning, resources, experiences, and expertise, and address real classroom challenges, enhancing their ability to implement differentiated teaching (Puzio et al., 2015). These views are also in line with the features of effective PLD that will be described in Section 2.5.2.

The literature review on PLD in support of differentiated teaching highlights several observations and shortcomings. Firstly, a common limitation in PLD research is the reliance on self-reported data from teachers and students, rather than direct observation of teaching practices (e.g., Gheyssens et al., 2020; Prast et al., 2018). Secondly, PLD initiatives typically provide limited opportunities for teachers to develop fundamental understandings of differentiation, often lacking specific techniques and strategies for its implementation (Slade et al., 2006; Santamaria, 2009).

[15] A community of practice refers to a group of practitioners who regularly interact to pursue a shared interest or collective goal (Wenger, 2002). They collectively develop a pool of resources including experiences, stories, tools, terminology, and routines. These resources are instrumental in supporting, facilitating, and interpreting the practices they share (Puzio et al., 2015).

Thirdly, most studies on teachers' conceptualizations and practices of differentiation did not examine any evolution in these areas (Santana, 2020). Fourthly, teachers often hold multiple interpretations of differentiation, and these interpretations significantly impact its implementation across different educational levels and subjects (Bondie et al., 2019; Lalvani, 2013). Notably, shifts in conceptual understanding can inspire teachers to modify their teaching methods, and conversely, changes in teaching practices can lead to new conceptual insights (Erotocritou-Stavrou & Koutselini, 2016; Santangelo & Tomlinson, 2012). Finally, there is a noted deficiency in the number and quality of PLD programs available for practicing teachers (de Jager, 2017).

A recent international comparative study on the development and impact of differentiation across diverse contexts and its effect on students' academic engagement (Maulana et al., 2023) highlighted key insights into PLD for differentiated teaching and the need for more research. Their findings indicated generally low quality of differentiation across various national contexts and a fluctuating use of differentiation over time. This points to the necessity of systematic, structural, and longitudinal monitoring for effective practice. These scholars also emphasized the importance of understanding the challenges teachers face in implementing differentiation and developing support mechanisms for its consistent integration into teaching. Additionally, they recommended equipping teachers with adequate conceptualizations, knowledge, and skills for differentiation through ongoing PLD.

This analysis of PLD in relation to differentiated teaching offers a nuanced understanding of the current state and challenges in this area. It underscores the significant benefits of PLD programs in enhancing teachers' abilities to understand and implement differentiated teaching strategies. However, it also reveals critical gaps, such as the superficial coverage of differentiation in teacher PLD programs and the inconsistency in the quality and application of differentiation across various educational contexts. The existing research emphasizes the need for more in-depth, practical, and sustained PLD initiatives, highlighting the importance of continuous support, reflective practices, and communities of practice. Moreover, it points to the necessity of systematic monitoring and research to further refine and improve teachers' practice of differentiated teaching. This discussion, therefore, not only provides a detailed overview of the current landscape but also sets a direction for future advancements and research in the field of differentiated teaching and teacher development.

Building on these insights, the next section brings together cognitive activation (considered in the previous section) and differentiation (considered in this section), delving deeper into the crucial balance between these two constructs. It elaborates on the importance of differentiating teaching without reducing task complexity, aiming to seamlessly integrate these two pivotal axes in educational practice.

2.4 Differentiating Teaching without Reducing Task Complexity

2.4.1 Ambitious Mathematics Teaching

Underpinned by the belief that mathematically challenging tasks are important for *all* students, ambitious mathematics teaching involves skilled ways of eliciting, supporting, and extending the thinking of all students so that they learn important mathematics (Anthony et al., 2015). Multiple scholars define ambitious teaching as "teaching that *deliberately aims to get all kinds of students*—across ethnic, racial, class, and gender categories—not only to acquire, but also to understand and use knowledge" (Lampert & Graziani, 2009, p. 492), and to "deepen understanding of ideas as well as their engagement in the solving of complex problems, rather than the more commonplace emphasis on activities and procedural talk" (McDonald et al., 2013, p. 385). Attending to the learning of all students, expecting them to learn complex ideas and skills, and be engaged in cognitively demanding work is an *intellectually and socially ambitious goal*, which leads to new definitions of teaching, that is 'ambitious teaching' (Cohen, 2011; Lampert et al., 2010).

A teacher who teaches ambitiously aims high, seeing the curriculum as a dynamic tool rather than a set of standards and tasks, and builds scaffolds that support all students reach those heights (this is the concept of "teaching up", Tomlinson & Imbeau, 2023). This view supports that teaching mathematics for all does not mean simply teaching students who are very capable in mathematics, or their future plans include becoming mathematicians or scientists, since as Brändström (2005) states, "not all students aim to become mathematical experts, but they should at least have the opportunity" (p.12). It also means teaching those students who might be indifferent about this subject matter or who might even hate it.

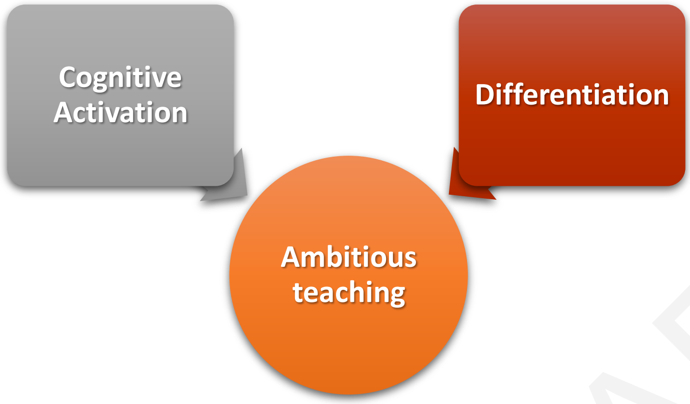
Ambitious mathematics teaching, as captured by the abovementioned definitions, encompasses the notions of involving students in *complex mathematical work* (i.e., inquiry processes and challenging tasks) and *equity* (i.e., responding and valuing *all* students' contributions, needs, understanding or errors in an equitable and responsive way, Anthony et al., 2015; Jackson & Cobb, 2010). This means that the teacher should make task complexity accessible to *all* students, rather than reducing it. This kind of teaching requires expertise in identifying, selecting, using, and modifying challenging tasks to adapt them to classroom heterogeneity.

However, given the diversity in contemporary mixed-ability classrooms, the descriptor "ambitious [mathematics teaching]" and the determiner "all [students]" can render this kind of teaching appear like an arduous task when it is not decomposed into its components which can help teachers better understand and materialize this kind of teaching (Anthony et al., 2015). The situation becomes even more perplexing in lessons in which mathematically challenging tasks are observed to be *too* difficult for some students and *too* simplistic for others, leading to situations in which some students become overly dependent on teacher guidance, while others remain unengaged (e.g., Mellroth et al., 2021). Research on ambitious teaching highlights that to address student learning gaps effectively, each and every student must have access to high-quality teaching, *mathematically challenging* tasks, as well as *differentiated scaffolds* to facilitate ongoing student success at progressively higher levels (NCTM, 2014).

In this respect, bringing *cognitive activation* and *differentiation* together is a way to work towards ambitious mathematics teaching (see Figure 10). Both notions work synergistically: on the one hand, working on challenging tasks requires teachers to constantly adjust a number of teaching variables (e.g., the content or the process, etc.) in order to respond to what the students are saying or doing (Lampert & Graziani, 2009); on the other hand, the value of differentiation lies in presenting a challenging stimulus to all students, with different tiers adjusting the complexity level (Tomlinson, 2015). The *interplay* of cognitive activation and differentiation lies at the heart of ambitious teaching suggesting that successful ambitious teaching depends on *effectively* combining/balancing mathematical challenge with appropriate scaffolds. While significant research has focused independently on cognitive activation and differentiated teaching, only a relatively small number of recent studies (e.g., Charalambous et al., 2022; 2023a; 2023b; Delaney & Gurhy, 2019; Mellroth et al., 2021; Psycharis et al., 2019) have begun to explicitly explore the intersection of these two areas—an emerging focus within the realm of ambitious mathematics teaching.

Figure 10.

Ambitious teaching as the Interplay of Cognitive Activation and Differentiation



Previous research points to a positive relationship between 'ambitious', 'reform-oriented' or 'inquiry-oriented' mathematics teaching and student performance (e.g., Blazar, 2015; Kane & Staiger, 2012; Kazemi et al., 2009). Noteworthy, the meta-analysis of quasi-experimental and experimental studies of Seidel and Shavelson (2007) illustrated that domain-specific processing (i.e., student's cognitive engagement in higher order thinking in mathematics) along with the social context (i.e., the degree to which the teacher establishes a differentiated social learning environment) and appropriate time allocation on working with a task have the highest effect sizes on student learning. Indeed, continuous, and appropriate adaptation of the level of cognitive demand of a challenging task in response to individual student or group of students is a main condition that helps to attend to the learning needs of students and leads to higher levels of understanding and mathematical proficiency (Gallagher et al., 2020). In addition, striving for flexible time allocation (e.g., allowing more time for students who need it to work on the task or increasing the challenge for more capable students or early finishers) can benefit learners and learning (Tomlinson et al., 2008).

Classroom research (e.g., Anagnostopoulos et al., 2020; Fauskanger & Bjuland, 2019; Kinser-Traut & Turner, 2020; Lampert et al., 2010) shows that ambitious mathematics teaching is doable; yet, shifting from the traditional "show and tell" form of teaching to teaching ambitiously so that students and teachers co-construct mathematics knowledge is a daunting task for teachers (Anthony et al., 2015). Implementing differentiation strategies in classrooms, where students are free to engage at their own level but are also challenged, requires knowledge, skills, and intent by the teacher. However, many teachers may have never been educated to really understand how students differ mathematically; learn effective ways of differentiating their teaching; identify appropriate tasks that are accessible and challenge students with varied needs; and structure, sequence, and coordinate these tasks appropriately (Little et al., 2009). Below, we discuss the resources and actions that facilitate ambitious teaching.

2.4.2 Resources and Teachers' Actions that Facilitate Ambitious Teaching

Several factors supporting ambitious teaching have been documented. These factors can be categorized into seven critical clusters: (a) *teachers' high Mathematical Knowledge for Teaching*; (b) *teachers' productive personal theories and beliefs about ambitious teaching*; (c) *teachers' teaching actions and practices that promote ambitious teaching*; (d) *complementing and using challenging tasks with enablers and extenders*; (e) *effectively orchestrating whole-class interactions*; (f) *anticipating classroom events*; and (g) *establishing classroom norms that embrace struggle*. These factors bear resemblance to those previously identified as catalysts for differentiated teaching. Nonetheless, they originate from a separate corpus of research, which consistently yields analogous results in the context of ambitious mathematics teaching. This parallelism in outcomes across the domains signifies a notable coherence in the findings.

Research underscores the pivotal role of robust *Mathematical Knowledge for Teaching* (MKT) in fostering ambitious teaching. A line of study highlights that teachers with strong content knowledge are more likely to utilize teaching materials effectively and differentiate their teaching to challenge students in a disciplinary-appropriate way (Murawski, 2019). This deep knowledge is crucial not only for planning lessons with achievable goals, connected to prerequisite knowledge, but also for adjusting teaching dynamically to align with students' Zone of Proximal Development. However, there is more than content knowledge for teaching ambitiously. MKT is an essential knowledge required for effective teaching, encompassing not just PCK, but also a specialized understanding of mathematical content that is necessary for teaching (Ball et al., 2005). This includes knowledge of why and how certain mathematical procedures work; providing grade-appropriate definitions and comprehensive explanations; using accurate representations; accessing student thinking equitably; and understanding common student errors (Hill et al., 2008). Studies have shown a positive relationship between various components of MKT and teachers' ability to develop and implement an ambitious vision of teaching (e.g., Charalambous et al., 2012; Kim et al., 2020;

Munter & Wilhelm, 2021). In particular, a standard deviation in MKT is associated with .25 standard deviation difference in ambitious mathematics teaching in classrooms where this kind of teaching is consistently implemented at high levels (Kelcey et al., 2019). In sum, MKT emerges as a foundational element for teaching ambitiously.

Another strand of research highlights a crucial link between *teachers' personal theories of teaching and learning* and their ability to engage in ambitious teaching. When these personal theories align with the principles of ambitious teaching, teachers are more likely to adopt such methods (Thompson et al., 2013). Teachers committed to ambitious teaching often view themselves as individuals deeply invested in understanding their students' thinking processes (Wæge & Fauskanger, 2021). Anagnostopoulos and colleagues (2020) emphasize that these beliefs are not static; they are often challenged and evolve as teachers negotiate the tensions and contradictions between their existing and emerging practices. This evolution can be particularly evident when teachers reevaluate their roles and responsibilities, as well as those of their students, within classroom interactions (Hunter, 2008).

In addition, some *key teaching actions and practices* enhance ambitious teaching, which can be further grouped into two subgroups. The first subgroup of practices relates to actions that a teacher can undertake during *lesson planning*. These include establishing clear learning goals; selecting or designing mathematically challenging tasks; analyzing, classifying, sequencing, and modifying tasks; anticipating students' mathematical approaches and potential errors; preparing various resources; and formulating questions or prompts to elicit, support, and extend student thinking (Kazemi et al., 2009; Psycharis et al., 2019; Webb, 2018). The second subgroup pertains to teaching actions during *lesson enactment*. This includes launching mathematically challenging tasks in ways that help students understand the task context and identify what they are asked to do, without prescribing specific methods; asking students to 'think-pair-share'; engaging students in exploration, connection, and reflection; monitoring students' engagement with the tasks during autonomous work and deciding when and how to intervene; orchestrating and facilitating whole-class interactions; addressing student errors; providing prompts and representations to support students' thinking; pressing students for explanations; and posing questions to elicit, support, and extend their thinking and reasoning (ibid.).

The research proposes two certain tools that facilitate the implementation of challenging tasks during student autonomous work: *enablers and extenders* (cf.

Sullivan et al., 2006; 2009). *Enablers, or enabling prompts*, are provided to students who struggle with the core task to help them proceed, possibly by simplifying steps or numbers or varying representation forms. *Extenders, or extending prompts*, are offered in the same context as the original task to enhance the thinking of students who complete the task early, such as by encouraging abstraction, generalization, or some aspect of proof related to solution completeness or strategy legitimacy (Sullivan et al., 2016a). The use of enablers serves a dual purpose: It offers teachers an alternative to telling the students what to do, and it instills in students the expectation to construct knowledge independently (Sullivan et al., 2009). Similarly, extenders communicate the message to students who have the potential for doing more that they can keep thinking and learning (ibid.). Crucially, as Sullivan's group highlights, teachers must allocate sufficient time for students to productively engage with tasks and actively monitor their progress. This monitoring helps determine when and how to appropriately intervene with an enabler or extender. Such decisions require a deep knowledge of student needs. For instance, premature provision of an enabler can reduce cognitive demands, as it may steal some of the student's thinking process. Likewise, extenders should not be given to early finishers or more advanced students without first confirming their understanding of the key mathematical ideas and successful task completion, as doing so does not ensure learning. Additionally, enablers must be mathematically correct and not substitute for student thinking, while extenders should relate to or expand upon the main task's challenge, requiring the teacher to understand the task, its demands, and the embedded mathematics (Charalambous et al., 2023a).

Another pivotal element in ambitious teaching is the *effective orchestration of whole-class interactions*, ensuring all students contribute to and understand key mathematical ideas and solutions. Recent research (Psycharis et al., 2019; Stylianides & Stylianides, 2014) has identified teaching practices that can balance challenge and differentiation in orchestrating whole-class interactions. These practices include building on students' contributions and ideas by revoicing, rephrasing, or reformulating student contributions to focus class attention on them or on learning milestones; asking students to explain the thinking underlying their contributions, while encouraging others to listen and respond (e.g., by expressing and justifying their agreement or disagreement); scaffolding through simplification (e.g., using a familiar situation) and extension (e.g., comparing different approaches, solutions, or representations); and valuing student contributions. A crucial action during this phase is soliciting diverse or even conflicting contributions, including difficulties and high-reasoning responses, from several students, thereby creating space for expression. Teachers should make the challenge accessible to all by recording solutions and ideas, inviting them to make connections among these solutions, questioning proposed ideas, and fostering an inclusive classroom environment (e.g., encouraging participation from typically silent students). In all, students should have opportunities to act as resources in the classroom, by interacting with and questioning one another as part of the discourse community, supporting the mathematics learning of all students in the class (Delaney & Gurhy, 2019; National Council of Teachers of Mathematics [NCTM], 2014).

Though teaching inherently involves unpredictability and uncertainty, the *anticipation of certain classroom events* can significantly enhance the enactment of ambitious teaching strategies (Stylianides & Stylianides, 2014). For instance, prior to a lesson, teachers can consider common difficulties or alternative perspectives students might encounter when engaging with the mathematical concepts being taught. Additionally, foreseeing typical patterns in students' solutions and responses to tasks—and planning ways to address them—can be beneficial. For example, thinking ahead about which questions to ask or which models and materials to use to elicit and scaffold students' thinking can enable teachers to be better prepared. Such preparedness helps in maintaining the rigor of the lesson without diminishing its demands.

Establishing classroom norms and cultivating a culture that embraces struggle are key factors in facilitating differentiation and engagement with challenging tasks (Delaney & Gurhy, 2019; Russo & Hopkins, 2019). Such norms become more apparent when students are encouraged to share and critically discuss their ideas with classmates, rather than solely presenting their solutions visually. Furthermore, openly discussing socio-mathematical norms and the nature of learning mathematics with the entire class helps in establishing this culture. For example, conversations that emphasize how learning mathematics requires effort and time, rather than innate talent, are beneficial. In this environment, students take responsibility for their own learning and are given the autonomy to decide whether and how to use the resources, and materials.

The following section explores the challenges teachers face when endeavoring to teach ambitiously, with a specific focus on simultaneously addressing cognitive activation and differentiation.

2.4.3 Challenges in Teaching Ambitiously

Teaching ambitiously is inherently complex (Kunter et al., 2008; Little et al., 2009). As Kazemi and colleagues (2009) observed, it is not a straightforward task that depends solely on knowing the correct actions; rather, it involves navigating various challenges and uncertainties. These challenges can be categorized into nine distinct clusters: (a) *lack of knowledge about students in relation to the content*; (b) *unproductive beliefs about ambitious teaching and learning*; (c) *difficulties in selecting, analyzing, and modifying challenging tasks during planning*; (d) *challenges in sequencing tasks in ascending challenge levels during planning*; I *difficulties in actively monitoring student progress and managing unexpected events during student autonomous work*; (f) *difficulties in balancing challenge and support while orchestrating whole-class interactions*; (g) *issues related to handling classroom heterogeneity*; and (h) *teaching time constraints*. In a similar line to the aforementioned discussion, a different body of literature, distinct from that used in cognitive activation and differentiation, is employed to explore the challenges inherent in ambitious teaching. The consistency observed across these research domains underscores a significant coherence in findings.

One of the primary challenges in ambitious teaching is related to *teachers' limited knowledge of students and content knowledge* (CK). Often, teachers lack the extensive CK necessary for ambitious teaching (Askey, 2001). Teachers with weaker CK may shy away from challenging tasks or fail to adequately address students' questions and difficulties (Russo & Hopkins, 2019). Such teachers might not fully understand the curriculum above and below their current grade level and are often unaware of alternative ideas or the prerequisite knowledge students need to achieve the lesson's goals (Clarke et al., 2014b). Furthermore, a lack of understanding of student thinking and the task significantly impacts the selection, modification, and implementation of challenging tasks (Sullivan et al., 2009; 2014). Consequently, teachers may struggle to understand or identify student readiness levels; anticipate student difficulties; and support and evaluate students' progress effectively (Clarke et al., 2014b; Stylianides & Stylianides, 2014).

Another significant challenge in accomplishing ambitious teaching is the presence of *unproductive beliefs* about this approach among teachers. For instance, many elementary teachers view "some of their students as incapable of engaging in rigorous mathematical activities" (Jackson et al., 2014, p.2). Consequently, when they perceive students to be struggling, they often reduce the mathematical challenge of the tasks. Interestingly, teaching experience might inversely relate to teachers' beliefs regarding ambitious teaching (Spillane et al., 2018). Specifically, as this study showed, more experienced teachers tend to be less oriented towards ambitious teaching on average than their less experienced counterparts, with every additional year of teaching experience associated with a .02 standard deviation shift towards less ambitious orientations.

Furthermore, many teachers encounter challenges in effectively *selecting, analyzing, and modifying tasks during the planning stage*. As a result, they often struggle to choose tasks that are suitably challenging based on their students' abilities (Little et al., 2009). Teachers frequently fail to consider their students' levels and the intrinsic nature of the tasks (Clarke et al., 2014b). Consequently, tasks that are open-ended and connected to real-life experiences are often overlooked. Teachers also tend to struggle with varying or differentiating tasks to meet diverse student needs or with incorporating enablers or extenders (cf. Charalambous et al., 2023a). Hence, when teachers attempt to modify a task by subdividing it into smaller subtasks, following an ideal-typical solution path and expecting less advanced students to engage only with the simpler parts of it, they inadvertently exclude these students from rich conceptual learning opportunities (Büscher, 2019).

In addition, *sequencing tasks in ascending challenge levels during planning* is a key yet difficult aspect of ambitious teaching. The sequence of tasks and the allocation of resources play significant roles in the development of student mathematical thinking and understanding. Teachers need to first ensure that students possess the necessary prerequisite knowledge before progressing to a sequence of tasks that aligns with students' readiness (Sullivan et al., 2011). Two effective approaches to task sequencing are highlighted in the literature. The first approach suggests creating sequences that begin with simpler tasks and progress to more challenging ones (Tzur, 2008). The second approach advocates starting with a challenging task and then effectively differentiating learning with task variations, including enablers and extenders (Sullivan et al., 2016b). However, many teachers find it particularly challenging to design lessons with task sequences that both cognitively activate students and meet their diverse needs (Stosich, 2016).

Moreover, teachers often encounter challenges in *actively monitoring students' progress and managing unforeseen classroom events* during student autonomous work.

Stylianides and Stylianides (2014) highlighted this complexity, noting that “the multiplicity of student contributions in ambitious teaching, as well as the expectation for the teacher to build on students’ ideas during the lesson, can increase the demands for improvisation and in-the-moment decision making” (p. 375). Novel student ideas or contributions that significantly deviate from lesson goals necessitate teachers’ proficiency in actively monitoring student work; rapidly diagnosing and responding to students’ thinking and contributions; flexibly grouping students; and effectively organizing and allocating resources (Clarke et al., 2014b; Kazemi et al., 2009). Managing significant deviations from the lesson plan introduces a high degree of uncertainty about how to proceed, leading to overly guiding students (Clarke et al., 2014b).

Orchestrating whole-class interactions is also one of the major challenges that accompanies the complex and uncertain ambitious teaching (Anthony et al., 2015) because the outcome cannot be anticipated (Dooley, 2009). During these discussions, teachers are tasked with maintaining the focus on mathematical ideas while simultaneously building on, honoring, and responding to students’ diverse thinking, abilities, paces, and learning preferences (Engle & Conant, 2002; Sullivan et al., 2006). The decisions regarding which approaches or solutions to share, ’he order in which they will be shared, and the types of questions that facilitate connections between different strategies and ideas are particularly challenging. Often, whole-class discussions devolve into ‘show-and-tell sessions’, in which the contribution of each student’s solution or idea to the overall understanding remains unclear (NCTM, 2014). Teachers thus face the dilemma of presenting only the correct solutions, potentially neglecting to showcase a range of solutions, including incorrect ones (Sherin, 2001).

Classroom heterogeneity presents additional challenges for implementing ambitious teaching (Clarke et al., 2014b; Stosich, 2016). For instance, teachers have noted that some tasks are overly challenging or lengthy for less advanced or slower students; that solutions or conclusions are reached at varying speeds by different students; and that less advanced or less confident students often struggle to start tasks independently. Additionally, many students show a lack of persistence in solving challenging tasks; not all real-life situations are age-appropriate; and students frequently focus more on finding the correct result rather than understanding solution methods (Clarke & Roche, 2009). Sullivan and colleagues (2006) studied the impact of student heterogeneity on teaching in upper elementary mathematics classes in

Australian schools. They found that even with skilled teachers and detailed plans, challenges persisted: some students found tasks too difficult, leading to off-task behavior; others required extra prompts or support to stay engaged; and a few finished tasks quickly and needed additional challenge. These observations underscore the difficulties posed by the wide range of student abilities, contributing to frustration for both teachers and students.

Finally, *teaching time constraints* significantly hinder teachers' ability to create optimal ambitious learning opportunities (Russo & Hopkins, 2019). Trying to balance between providing students with ample time to persist in solving challenging tasks and the students' varied learning paces, teachers often face the dilemma of choosing between providing ample time for thinking and engaging with the task; peer or teacher sharing; and directly instructing students on task solutions. When implementing ambitious teaching, teachers often wonder, "Do we follow the curriculum or follow the students?" (Horn & Garner, 2022, p.10). This is especially difficult when trying to balance the mathematical challenge and support to students who either complete tasks quickly or struggle to start (Anthony et al., 2015).

The next section elaborates on the available research on designing and implementing PLD programs to support teachers while teaching ambitiously, highlighting open issues.

2.4.4 Supporting Teachers in Teaching Ambitiously: Research Findings and Open Issues

In recent years, researchers have recognized the challenges of engaging students with diverse abilities and readiness levels in challenging work. Studies have focused on aiding practicing teachers in integrating mathematically challenging tasks into lessons and helping students effectively engage with these tasks (e.g., Hunter, 2008; Pfister et al., 2015; Sullivan et al., 2015). Table 3 presents a comparative overview of various studies on the effectiveness of PLD for practicing teachers with respect to *ambitious* mathematics teaching.

To understand what literature suggests regarding the effectiveness of PLD programs focusing on ambitious mathematics teaching, studies conducted from 2000 onwards were reviewed, utilizing multiple databases including SCOPUS, ERIC, Web of Science, and Google Scholar. This starting point was selected as the concept of ambitious teaching was prominently discussed in the educational discourse in the early

21st century, with the influence of scholars in mathematics education like Lampert (2001) and Ball (2003). The search employed specific keywords combined with Boolean operators to refine the results: ("ambitious teaching" OR "ambitious instruction" OR "teach* ambitious*") AND "mathematics" AND ("professional development" OR "intervention" OR "professional learning and development") AND ("student learning" OR "student performance") AND ("teaching quality" OR "instructional quality" OR "quality teaching"). Only English language studies were included. Initially, 24 studies were identified, but the selection was refined based on certain exclusion criteria. Exclusions were made for studies without interventions, those focusing solely on prospective teachers, or those concerned with intervention adaptations rather than the teacher or student outcomes. Additional exclusions were applied to studies that lacked an examination of the interventions' effectiveness. After applying these criteria, 19 studies remained that directly addressed the effectiveness of PLD programs. This review was not intended to be exhaustive; instead, it aimed to provide a targeted exploration of key findings and unresolved issues in evaluating PLD programs focused on ambitious mathematics teaching.

These 19 studies, detailed in Table 3, span a range of grade levels, from kindergarten to secondary education. They also concentrate on two main areas: outcomes related to the teacher (their learning, professional noticing, MKT, and teaching practice) and outcomes related to the students (as depicted in the last two columns of Table 3)—this criterion was used to group and present the studies in Table 3. Some studies focus on one of these areas, while only a few studies encompass both, reflecting an understanding of the interplay between teacher learning, teaching practices, and student learning outcomes. Notably, most of these studies focus more on teacher and teaching practice (totaling 19 studies) rather than student outcomes (a total of six studies). This prioritization is reasonable since it allows for a detailed exploration of potential changes in teachers and their teaching, before examining the impact on student learning. This sequence ensures a thorough understanding of if and how improvements in teaching practices can subsequently influence student outcomes.

Table 3.

Comparative Overview of Studies on the PLD of Practicing Teachers with Respect to Ambitious Mathematics Teaching

Authors	**Type of study**	**Number of participants**	**Grade level**	**The focus of the intervention**	**The Intervention**	**Duration (duration of focal data)**	**Outcomes examined**	**Results**	
								Teachers and teaching practice	*Student outcomes*
Gibbons et al. (2017)	• Qualitative • Case study approach	Five teachers	Elementary (Grades 4-5)	Nurture a vision of AMT[1], with practical tools and practices for school-wide adoption.	One-on-one coaching in teachers' classrooms; Implementation of Teacher Time Outs (TTOs)[2] into math lab sessions in a school-wide PLD	Three years (a math lab session during the second year of the intervention)	Examined how the TTOs facilitated teachers' learning and contributed to the development of a school-wide professional community *over time*	Preliminary data on supporting teachers' *collective* learning and understanding of teaching as a complex enterprise; and developing a professional school-wide community	-
Gibbons & Okun (2023)	• Qualitative • Discourse analysis of TTOs	Not given. Involved teams of 18-20 teachers (each school year), a coach, and the school principal	Kindergarten through Grade 5	Consider broad AMT categories, such as eliciting student thinking, facilitating effective mathematics discussions among students, etc.	A full-day (≈6 hours) job-embedded structure with one-on-one coaching; individualized teacher support was offered through TTOs in real-time classroom teaching	Three years (analysis of 360 TTOs)	Examined how the TTO routine supported teachers' professional learning and their understanding and ability to enact AMT *over time*	• Positive findings in terms of focus (towards more complex teaching aspects). • Fostered *collective* inquiry into practice, deepening teachers' AMT understanding and enactment.	-
Wæge & Fauskanger (2021)	• Qualitative • Coding and development of analytical memos based on video analysis of TTOs *within*	14 teachers	Elementary (Grades 1–7)	Support teachers' learning of AMT practices, focusing on problem-solving, eliciting and responding to student	12 sessions, with nine of them focusing on rehearsals of teaching in a cycle of enactment and investigation. Used TTOs while	Two years (18 videotaped rehearsals, each lasting maximum 30 minutes)	Examined the patterns of use of TTOs in rehearsals and how TTOs enable teachers to learn core AMT practices (changes in the	• Teachers worked simultaneously on multiple practices; *collectively* learned to use AMT practices effectively and flexibly, adapted	-

	and across rehearsals			thinking; aiming toward a mathematical goal; and using representations	teachers taught an activity to their colleagues who acted as students		focus and discourse of TTOs)	to student inputs; and developed a shared understanding of AMT.	
Wæge & Fauskanger (2023)	• Qualitative • Coding and development of analytical memos based on video analysis of TTOs *within and across rehearsals*	14 teachers	Elementary (Grade 7)	Support teachers' learning of AMT practices, focusing on problem-solving, eliciting and responding to student thinking; aiming toward a mathematical goal; and using representations	12 sessions, with nine of them focusing on rehearsals of teaching in a cycle of enactment and investigation. Used TTOs while teachers taught an activity to their colleagues who acted as students	Two years (18 videotaped rehearsals, each lasting maximum 30 minutes)	Examine how TTOs support teachers' *collective* learning of AMT and the development of their pedagogical judgment	• Positive findings of teacher *collective* learning of AMT practices and the enhancement of their pedagogical reasoning and in-the-moment decision-making during teaching.	-
Jakopovic (2021)	• Qualitative • Case study • Narrative inquiry	One teacher	Elementary (Grade 1)	Develop teacher's professional noticing of AMT practices, such as developing mathematical goals, planning and adapting mathematical tasks, and examining student thinking	Consisted of iterative coaching cycles, in which they planned, co-taught, and debriefed lessons together	Not specified (two coaching cycles)	Examined the *shift* in the teacher's professional noticing of AMT practices as influenced by the coaching cycles *over time*	Preliminary evidence that iterative coaching with a focus on developing sustained teaching goals and a gradual release model of coaching can gradually shift teachers' professional noticing of AMT	-
Fauskanger & Bjuland (2019)	• Qualitative • Conventional content analysis of the PLD discussions	Seven teachers	Secondary (Mostly taught Grades 5–7)	Develop AMT concepts and practices for teaching multiplicative properties	A school-based PLD project with cycles of enactment and reflection with TTOs; it ran over four semesters, with three four-	Two years (three four-hour sessions from the fourth PLD cycle)	Examined the learning opportunities provided to teachers in understanding and implementing	Positive evidence of change in *collective* teacher learning of AMT; Teacher struggles with AMT practices provided opportunities for	-

					hour cycles each semester		AMT practices throughout the PLD	learning various AMT components	
Witherspoon et al. (2021)	• Mixed methods • Case-comparison • Hierarchical linear growth • Logistic regression analysis • Survival analysis	Four coach-teacher pairs	Elementary and secondary (Grades 4-8)	Implement AMT practices; understand when and why they should enact certain high-leverage practices; focus on maintaining cognitive demand and attending to student thinking	A coaching cycle included goal and task selection; a pre-observation conference; lesson observation by the coach; a post-observation conference	Two years (both years; each coaching pair was met three times in the first year, and twice in the second year)	Assessed the teaching quality, specifically the ability to conduct AMT; Examined how the coach-teacher interactions were linked to teacher gains in enacting AMT	Coach-teacher pairs discussing when and why implementing certain practices, and allowing greater teacher input, have greater gains in AMT lessons.	-
Leong et al. (2021)	• Qualitative • Case analysis of four participating schools • Thematic analysis	Over 100 teachers	Secondary (Grades 7-10)	Focused on scaling up an AMT practice, mathematics problem-solving (MPS), among teachers as a regular classroom element	A long-term approach of PLD with various phases, including learning about MPS, teaching MPS, embedding MPS into the curriculum, and refining the MPS approach	Thirteen years (nine years)	Examined the capacity of schools *at the end of the program* to sustain MPS teaching, the implementation of the MPS, and the support structures within schools	• Varying results among schools • Positive trend in sustained capacity for MPS teaching; adaptations of the MPS; and development of supportive structures within three of the schools.	-
Charalambous et al. (2023a)	• Qualitative • Multiple case study	Four teachers	Elementary (Grades 1, 2, and 6)	Use enablers and extenders as a tool for achieving AMT, in the context of designing and using these tools in lesson planning and enactment	A video-club PLD program, focusing on cognitive activation and differentiation in teaching. Teachers were engaged in guided discussion and	Over six months (nine 2.5-hour video-club sessions; around 120 hours of PLD)	Explored how teachers worked with enablers and extenders during lesson planning and enactment and the challenges they faced in this process; implicitly compared	• Positive changes in teachers' experimentation with these tools, while reporting on and codifying the challenges they faced. • Notable *variation* in teachers' experimentation	-

					reflection upon their practice.		lessons before and after being introduced to enablers and extenders		
Anthony et al. (2018)	• Qualitative • Case study • Thematic analysis	One teacher, Tina	Elementary (a class of 8–10-year-old students)	Support teachers in developing ambitious teaching, pedagogical vision, and practices, particularly in settings serving diverse students	A whole-school PLD initiative emphasizing equitable and culturally responsive pedagogies, discourse-rich mathematical inquiry, and the use of challenging collaborative tasks.	Does not specify the exact duration but indicates it covers the first year of the intervention.	Examined changes in Tina's views on students' capabilities and her engagement with AMT practices, especially regarding students who struggle with mathematics	Tina is presented as a case of no considerable change; she maintained largely unchallenged deficit framings of students and struggled to incorporate new teaching practices effectively.	-
Garrett et al. (2019)	• Meta-analysis of randomized experiments of interventions directed at teaching practice • Random- and mixed-effects models	The number of teachers varied across the included studies	Kindergarten to Grade 12	Focused on how teaching practice responds to PLD, particularly in terms of specific aspects of classroom practice and the effects of different intervention features	Various interventions to improve classroom practices for supporting student learning. The interventions encompassed PLD program and coaching for teachers	Duration varied across the included studies	Examined the extent to which teaching practice is responsive to intervention (changes in classroom practice as measured through classroom observations)	• Classroom practice is responsive to intervention. • Interventions directed toward classroom practice have positive impacts on average. • Substantial heterogeneity in the effects	-
Sun et al. (2014)	• Quantitative • Hierarchical linear models	89 teachers	Secondary (does not specify the grade level)	Support teachers learning to improve student learning, and incorporate reform-oriented	PLD and coaching for teachers, with a specific emphasis on improving MKT and teaching practices	Two years (used data collected from both years)	Examined changes in teachers' MKT and their teaching practices, exploring the influence of	• Positive link between teaching practice change and access to close colleagues' expertise. • No significant link between teachers'	-

				standards and curricular materials into teaching.	through peer interactions and coaching.		teachers' networks, and coaches' expertise in increasing teachers' learning through interactions with close colleagues	MKT change and access to close colleagues' MKT. • No significant effect of coaches' expertise on teaching practice.	
Lindvall et al. (2023)	• Quantitative • Repeated measures	52 teachers	Elementary and secondary (Grades 1–9)	Improve teachers' teaching practices in mathematics; and student mathematics achievement.	A national-scale PLD program comprising modules based on collegial meetings under the guidance of a trained coach	One year (PLD sessions comprise around 60 hours; analysis of 174 video-recorded mathematics lessons during the PLD year)	Examined the effects of the PLD on teaching quality *during* the program, rather than comparing teaching before and after the PLD	• Negligible and non-significant change in teaching quality *throughout* the program	-
Shumway et al. (2020)	• Mixed methods • Descriptive statistics • Linear mixed-effects models • Case studies (open and axial coding of 11 student interviews)	Five teachers and 75 students	Elementary (Grade 2)	Develop subitizing and number system knowledge of teachers; enhance teaching "talk moves" to facilitate and orchestrate high-quality and purposeful discussions with students	Four PLD sessions before, during, and after teachings (including learning about number sense, rehearsals; provision of materials; and debriefing of teaching in groups and individually)	Nine weeks (Implementation of NSK treatment in the classroom three days per week; nine implementations of NSK treatment per teacher)	Examined variations and shifts in students' Number Sense Knowledge (NSK) outcomes *after* participating in the NSK treatment	-	• Significant improvements, especially for students with the lowest pretest scores • Students' views on the NKS treatment were overall positive
Dash et al. (2012)	• Quantitative • Randomized controlled trial design	79 teachers and 1438 students	Elementary (Grade 5)	Improve teachers' PCK and teaching practices (using representations, drawing	Three online PLD courses in fractions, algebraic thinking, and measurement;	Three semesters (one week of orientation and six weeks of course content per semester; 4-	Examined the effects of the intervention on teachers' PCK, pedagogical practices, and	Significant gains in PCK and teaching practices for the teachers in the experimental group	No meaningful differences in students' achievement

				generalizations, attending to students' understanding)	one course per semester	6 hours of work per week per teacher)	their students' achievement		
Kraft & Hill (2020)	• Quantitative • Randomized field trial	142 teachers	Elementary and Secondary (Grades 3-8)	Support teachers in implementing Common Core-aligned mathematics teaching through web-based coaching	A two-day summer PLD for teachers, followed by ongoing coaching	Two years (including the implementation year and a follow-up year)	Assessed the impact of the Mathematical Quality of Teaching (MQI) Coaching on teachers' ability to analyze teaching and their teaching practice	Significant and sustained effects on teachers' ability to analyze teaching and their teaching practice	No corresponding increases in students' mathematics test scores.
Lindvall et al. (2022)	• Quantitative • Non-equivalent groups design • Linear regression	• 162 teachers and 3618 students in Grade 4 • 191 teachers and 3884 students in Grade 8	Elementary and secondary (Grades 4 and 8)	Improve teachers' teaching practices in mathematics, and student mathematics achievement	A national-scale PLD program comprising modules based on collegial meetings under the guidance of a trained coach	One year (the full year included covering two modules)	Examined the final effects of the PLD program on teachers' teaching practices and students' mathematics achievement	A small but statistically significant positive effect of the PLD program on teachers' teaching practices	No significant effect on students' mathematics achievement
Jacob et al. (2017)	• Quantitative • Randomized control trial design	• 105 teachers and 1523 students	Elementary (Grades 4 and 5)	Improve teachers' MKT; enhance their ability to elicit student thinking and reasoning; and develop effective classroom teaching strategies	Involved working with tasks and various strategies; student work analysis; teaching demonstration; addressing classroom issues; and reflections on teaching.	Three years (over 40 contact hours per year; a week-long summer school and four to six in-person days during the school year)	Examined the impact of the program on teachers' MKT; their teaching practices; and student achievement	Limited evidence of positive impact on teachers' MKT	No effects on teaching practice or student outcomes
Hill et al. (2018)	• Mixed	105 teachers	Elementary (Grades 4-5)	Enhance teachers' MKT	A scaled-up PLD program	Three years (district PLD	Evaluate the impact on	• No significant impacts on	No significant impacts on

	• Cluster randomized trial • Qualitative case study	and teaching, and student participation in mathematical thinking and reasoning; and improve student test score outcomes	delivered, including three four-day summer institutes and in-person sessions during the school year	sessions once a month during the school year; totaling 19 days of PLD)	teachers' teaching practices and student outcomes; identify valid reasons for null results	teachers' teaching practice • Slight improvement in some aspects of teaching but overall, null, or negligible results.	student outcomes

Notes.
1. AMT: Ambitious Mathematics Teaching
2. A Teacher Time Out (TTO) refers to a moment during teaching enactment when the ongoing activities are deliberately paused. This pause is utilized to allow teachers and teacher educators to pose questions, reflect collaboratively, and consider their teaching decisions before resuming the teaching process (Gibbons et al., 2017).

Various trends are evident in the Table in terms of the research findings. In particular, most *qualitative* studies generally show a positive trend in supporting teachers' *collective* learning *throughout* the PLD intervention. Key findings include a *shift* in understanding of teaching as a complex endeavor (Gibbons et al., 2017); in focusing on more complex aspects of teaching; in enhanced inquiry into practice (Gibbons & Okun, 2023); in understanding and learning to use ambitious mathematics teaching practices effectively and flexibly (Fauskanger & Bjuland, 2019; Wæge & Fauskanger, 2021); in improving pedagogical reasoning and decision-making during teaching (Wæge & Fauskanger, 2023); and in developing professional noticing of ambitious mathematics teaching (Jakopovic, 2021).

Despite the insights yielded from these studies, they also have some limitations which point to research gaps in understanding the impact of PLD programs. These include the reliance on participant feedback at the end of the intervention for evidence of PLD effectiveness (Gibbons et al., 2017), and the focus on teacher *collective* learning without considering any differences withing the PLD group (Wæge & Fauskanger, 2021; 2023). Despite analyzing data within and across the PLD sessions, by analyzing teacher learning as a unified whole, researchers may miss individual or group variances in teacher learning. Moreover, the predominant use of data from *simulated* teaching scenarios instead of real classroom settings (Gibbons et al., 2017; Gibbons & Okun, 2023; Fauskanger & Bjuland, 2019; Jakopovic, 2021; Wæge & Fauskanger, 2021; 2023) limits insights into the true complexities and dynamics of teaching environments and the direct impact of PLD on actual teaching practices.

Another series of studies employing *qualitative or mixed methodology* generally indicate a positive impact on teachers' learning and teaching practice, yet they reveal a *noticeable variation*. The studies highlight a significant pattern suggesting that not all teachers (Anthony et al., 2018; Charalambous et al., 2023a; Witherspoon et al., 2021) or schools (Leong et al., 2021) respond in the same way, or achieve the same outcomes, after participating in the same PLD program. For example, Anthony and colleagues' (2018) portrayal of a teacher as a case of no considerable change in terms of her practice and beliefs, and Charalambous and associates' (2022) cases of four teachers with notable variation and challenges in their experimentation with enablers and extenders, reflect different levels of uptake of the concepts and ambitious mathematics teaching. Such findings that emerge from real classroom environments reinforce the aforementioned conjecture that teachers may experience different outcomes or react differently to the PLD interventions and practices promoted through these programs. However, the reliance on *self-reported data* by the teacher (Anthony et al., 2018) and the focus more on the experimentation rather than on explicitly documenting changes in teachers practice (Charalambous et al., 2023a) make it difficult to vividly depict the impact of this intervention on teaching practices. These observations highlight a broader issue in PLD research stressing the necessity for *clearer articulation of change between lessons over the course of the PLD program*.

Interestingly, the *quantitative* studies in Table 3 illustrate *mixed findings* of the potential of PLD to improve teacher practice. Specifically, one study supports significant gains in teacher learning (Kraft & Hill, 2020); some studies report significant improvements in teaching practice (e.g., Dash et al., 2012; Kraft & Hill, 2020), while others show minimal or even null impact (e.g., Hill et al., 2018; Lindvall et al., 2022; 2023). Similar contradictions are identified in the effect of PLD programs on teachers' knowledge, with some studies suggesting significant improvement of teachers' PCK (Dash et al., 2012) and limited impact on their MKT (Jacob et al., 2017). The studies also generally show that PLD programs do not significantly impact student achievement in mathematics (e.g., Dash et al., 2012; Hill et al., 2018; Jacob et al., 2017; Kraft & Hill, 2020; Lindvall et al., 2022). The only exception is the research by Shumway et al. (2020), focusing solely on student outcomes; this study revealed an average gain of 21 percentage points across all classes, suggesting some potential for positive impact in specific contexts or program designs. However, interesting patterns and variations were identified among students, especially between those with low and high pretest scores, indicating the possibilities of teacher effects. This supports the argument that there may be differences *within* the PLD groups regarding what they learn and how they apply what they learn.

However, as in qualitative studies, the focus of these studies on the *collective* change of teachers *at the end* of the PLD program, rather than on examining differences in the learning of different teacher groups during the PLD program, presents a limitation. This approach may overlook the nuances of how different teachers or groups of teachers assimilate and implement the learned practices, and how different factors and challenges may impact the uptake and application of PLD content. By not examining this individual or group variance, the studies may miss valuable insights into the *differential effectiveness of the program and potential areas for personalized or contextual adaptations*. Those conclusions and the scarcity of mixed-method research suggests a need for more studies. These studies should not only explore how teachers collectively learn and improve their teaching practice but also investigate how different individuals or groups of teachers learn and improve. Finally, there is a need to examine changes throughout PLD programs, rather than merely compare initial and final products of such programs; doing so will allow gaining a deeper understanding of the learning process and its effects on teaching effectiveness.

The present research seeks to address these gaps, following researchers' urge to better explore and understand the differential teacher learning and change in terms of teaching practice to reduce variation in implementation and student learning (Witherspoon et al., 2021). Hill et al. (2018) emphasize the necessity for designing PLD programs that closely correspond to the real-world needs and circumstances of both teachers and students. This entails the adoption of a differentiated approach for teachers, mirroring the principle of differentiation used for their students: being sensitive to teachers' unique and diverse challenges and skills (Hill et al., 2018; Shumway et al., 2020).

Two studies by Lindvall's team (2022; 2023) presented in Table 3 stand out and warrant individual attention. Both studies used data from the same PLD program but differed in their approach and findings. Lindvall et al. (2022) questioned the effectiveness of a year-long PLD program in enhancing the ambitious teaching quality of nearly 200 teachers *after the completion of the program*. They noted a small but significant effect on teaching practice based on teacher questionnaires (i.e., .3 standard deviations). In contrast to their earlier study, Lindvall et al. (2023) focused on the timing and progression of teaching changes *during PLD* by analyzing 174 lessons from 52 teachers (including 3-4 videotaped lessons per teacher). Repeated measures revealed that teaching quality *on average* did not increase over the PLD. Although the approach of analyzing lessons throughout the program, instead of only at its conclusion, marks a significant departure from the other quantitative studies listed in the table, this average result might mask systematic variations among teachers, with possible differential impact on different teachers, based on teacher characteristics, initial status and teaching practice at the commencement of the program, or contextual circumstances faced by each teacher (Lindvall et al., 2023). An additional limitation concerns the fact that these studies did not examine any changes in teachers' conceptualizations and teaching practice, let alone the interconnection between the two.

Such exploration of conceptualization and teaching practice is undertaken in the multiple case-study of Horn and Garner (2022). By closely documenting eight cases, the authors delve into the conceptual challenges and learning processes teachers undergo to implement ambitious teaching practice in their classrooms. This approach not only contextualizes teacher learning within real-world teaching environments but also highlights the diverse factors influencing teachers' PLD and the implementation of innovative teaching practices. These scholars suggest that future research might delve into more specific aspects of teacher learning, particularly in relation to the sociocultural dynamics of the classroom and the PLD environment. Additionally, the book's focus on ambitious and equitable teaching practices suggests that further research could explore how these practices are implemented across different educational contexts and their impacts on teacher development. The complementarity of this book with the results of the quantitative studies listed above underlines the importance of conducting mixed-method studies to explore the effectiveness of PLD programs from different perspectives.

As briefly discussed in the current section, the features of the PLD design play a key role in supporting teachers as they learn new ways to conceptualize and implement ambitious teaching. The next section will focus on defining contemporary views of teacher PLD and identifying the elements of an effective PLD program.

2.5 Teacher Professional Learning and Development

2.5.1 Defining Teacher Professional Learning and Development

Professional Learning and Development (PLD) captures the structured, facilitated learning activities, actions, interactions, and experiences designed to enhance teachers' learning, knowledge, abilities, skills, attitudes, beliefs, and teaching practice, and ultimately benefit student learning outcomes (Darling-Hammond et al., 2017; Desimone, 2009; Sims et al., 2021; Vangrieken et al., 2017; Yoon, et al., 2007).[16] Traditionally, teachers, researchers, school administrators, teacher educators, and other relevant stakeholders have held a rather narrow view of PLD, a perspective that, in most cases, continues to prevail (Guskey, 2000; Smith & Gillespie, 2023). When asked, many teachers and stakeholders describe PLD as off-site single or sporadic

[16] It is important to note that PLD is broader than career development. The latter refers to the professional career cycle of a teacher and is defined as the actions undertaken by a teacher to advance to higher job responsibilities or to transition to a new position within the same organization (Villegas-Reimers, 2003).

gatherings, such as meetings, events, lectures, conferences, seminars, or workshops, which are typically condensed into a few hours or days throughout the school year and focus on different educational topics (Borko, 2004).

Indeed, teachers' experiences of PLD are frequently limited to "a series of unrelated, short-term workshops and presentations with little follow-up or guidance for implementation" (Guskey, 2000, p. 15). Furthermore, teachers are usually not involved in planning these meetings, and the content presented often fails to resonate with their specific classroom or school contexts (Bayar, 2014). This situation is exacerbated by policies that mandate teachers to accumulate a certain number of PLD hours annually to maintain their evaluation or employment, reinforcing the perception that PLD is merely an obligation, quantifiable in hours, rather than an integral part of a professional teaching culture that fosters continuous learning for everyone—students, teachers, and school administrators alike (Darling-Hammond et al., 2017; Guskey, 2000). As Guskey (2000) aptly puts it, teachers 'tend to think of PLD in terms of "How can I get my hours?" rather than "What do I need to improve my practice, and how can I achieve it?" (p. 15).[17]

Over the past two decades, the educational research community has made significant strides in clarifying and enriching our understanding of teacher PLD and its processes (Borko, 2004; Desimone, 2009; Evans, 2014; Smith & Gillespie, 2023). The concept of PLD has expanded to include a variety of activities, from formal, scattered, structured seminars or workshops conducted on in-service days (i.e., *the traditional PLD model*) to *ongoing* PLD opportunities within a school program or a local context, which includes study circles, sharing groups, or inquiry groups comprised of teachers from the same school or district (i.e., *the job-embedded PLD model*) (Smith & Gillespie, 2023). The latter type of PLD focuses on developing teacher learning (including the what, when, and how a new teaching skill can be effective) and changing teaching practice customized to the teacher's educational environment, through the use of artifacts of practice (e.g., student work, videoclips from teaching, etc.), actual experimentation in their classrooms, and critical reflection on their practice (Ball & Cohen, 1999; Smith & Gillespie, 2023). Arguably, the most impactful PLD experience occurs within a teacher's own classroom through self-analysis of and reflection upon

[17] Although fragmented, this model of PLD can be appropriate and effective when teachers need to acquire information about new programs or policies. Its effectiveness is further enhanced when it includes follow-up activities or meetings that support the implementation of new ideas (Guskey, 2000).

their practice; this non-traditional PLD model illustrates the dynamic and multifaceted nature of teacher learning and improvement (Avidov-Ungar, 2016).

As already explained in the previous section, research indicates significant variation in the effectiveness of PLD initiatives in achieving their intended goals (Darling-Hammond et al., 2017; Fullan, 2007; Sims et al., 2021). For example, a study by The New Teacher Project (TNTP, 2015) in four U.S. districts serving predominantly low socioeconomic status students revealed that despite substantial annual investments in teacher PLD—nearly $18,000 per teacher—there was minimal change in teachers' practices as measured by teacher evaluations. These evaluations either remained constant or declined over two to three years. This finding does not suggest reducing PLD investments. Instead, authors recommend that school systems should guide teachers on how to improve—and develop awareness that they have room to improve—and create conditions conducive to growth.

Concurrently, an increasing number of studies on PLD demonstrates that well-designed and effective PLD can lead to positive changes in teaching practice (Darling-Hammond et al., 2017). This leads to key questions: What characteristics define effective PLD models? What conditions are necessary to support and enhance teacher learning (Clarke & Hollingsworth, 2002)? Recent research underscores the importance of addressing these questions, considering the mixed results regarding the impact of PLD efforts on teacher learning (Hill et al., 2013). We address these questions in the following section.

2.5.2 Key Features of Effective Professional Learning and Development

Scholars have extensively investigated the features of PLD, focusing on the content, the context, and the design of teachers' learning experiences and their impact on teachers' learning and teaching practice (e.g., Darling-Hammond et al., 2017; Desimone, 2009; Guskey, 2003; Merchie et al., 2016; Sims et al., 2021; Timperley, 2008; Wei et al., 2009). In the early 2000s, the educational research community endeavored to establish a set of criteria to evaluate the effectiveness of PLD (Desimone, 2011). Guskey (2003) analyzed 13 existing lists delineating the characteristics of effective PLD and concluded that the criteria for PLD effectiveness varied widely, with research arguments often being inconsistent or contradictory. For instance, some researchers defined effectiveness based on teachers' self-reports of PLD features, while others sought scholarly consensus about PLD. Consequently, these lists were seldom grounded in research evidence that confirmed the significance of the criteria, and the connection between the identified characteristics and specific measures of teacher or student outcomes was largely unexplored. Guskey (2003) posited that the research community needs to achieve consensus on effectiveness criteria, complete with clear descriptions of factors related to the PLD context.

During the last two decades, more systematic efforts were made to record the characteristics of effective PLD programs. For instance, Wei and associates (2009) reviewed several PLD policies and practices in the USA and abroad. They identified several common features of PLD practices in countries that perform well in international studies like PISA and TIMSS. According to this review, teachers in these top-ranked countries have access to *extensive opportunities* for in-service PLD activities, *deeply embedded in their contexts*. These activities are not only ongoing but also *span a relatively lengthy period*. Teachers receive allocated *in-school time* for professional learning and *collaboration with colleagues*. Additionally, they are *supported by school administrators* in participating in decisions about curriculum and teaching practices. While these findings do not establish causal links between these features and student achievement in high-achieving countries, the frequency of these characteristics in research conducted in these nations suggests a potential connection between the opportunities for teacher PLD and the resultant quality of teaching and learning.

Longitudinal research has shown that the *features* of PLD programs, such as active and collaborative learning, play a significant role in enhancing teacher learning and teaching practice, rather than the *structure or specific types of activities* included in these programs (e.g., Desimone et al., 2002; Garet et al., 2001; Porter et al., 2000). Numerous eminent scholars have developed conceptual frameworks that encompass the key or most commonly identified features of effective PLD (cf. Darling-Hammond et al., 2017; Desimone, 2009; Merchie et al., 2016; Timperley, 2008). These seminal works with rigorous designs have empirically validated the effectiveness of different PLD programs, thereby contributing to a growing consensus about the key features of effective PD.

Table 4 provides a comparative view of the various key features of effective PLD according to four seminal and influential conceptual frameworks. These frameworks emerged from *meta-analyses or meta-syntheses* of quantitative and/or qualitative data.

A checkmark indicates whether a particular researcher or research team has included that feature in their framework.

Table 4

Review of Conceptual Frameworks of Effective PLD Features

Key Features of Effective PLD	Timperley (2008)	Desimone (2009)	Merchie et al. (2016)	Darling-Hammond et al. (2017)
• *Extended duration, with multiple opportunities to revisit and experiment with new practices*	✓	✓	✓	✓
• *Opportunities for participating in professional learning communities*	✓	✓	✓	✓
• *Integration of pedagogical content knowledge and skills*	✓	✓	✓	✓
• *Coherent and evidence-based; aligned with effective teaching principles and worthwhile content*	✓	✓	✓	
• *Involvement of knowledgeable and high-quality experts to plan and facilitate PLD*	✓		✓	✓
• *Active learning by reflecting on experience*		✓		✓
• *Focus on developing teacher professional inquiry and self-regulatory learning skills to increase ownership*	✓		✓	
• *Engage teachers in discussing how their existing ideas differ from the promoted ones*	✓			
• *Focus on the links between teaching practices and student outcomes*	✓			
• *Momentum maintenance and long-termed improvement*	✓			
• *Active leadership to maintain teachers' interest and ongoing learning*	✓			
• *School or site based; incorporated into teachers' daily practice*			✓	
• *Use of models and modelling, such as written cases of teaching, lesson plans, and observations, or curriculum materials*				✓

The table organizes the characteristics of effective PLD based on their frequency across four theoretical frameworks, as denoted by checkmarks. Some features included in the conceptual frameworks were merged or embedded within others due to their similarity or explicit interrelation. The top seven features, widely recognized by at least two of the frameworks—and hence, by several individual studies since, as already noted, these frameworks were developed based on meta-analyses and meta-syntheses of studies—appear to attract wide recognition for their essential role in teacher PLD.

Notably, research suggests that while successful PLD programs typically integrate most of the features contributing to their effectiveness, they rarely incorporate all of them simultaneously (Darling-Hammond et al., 2017).

The current study leverages such consensus to inform the design of the teacher PLD program (see Chapter 3). Toward this end, the aforementioned seven features which were most frequently discussed and found to be most impactful on teacher practice were systematically organized into three interconnected categories, as illustrated in Figure 11.

Figure 11.

A Synthesis of Key Features of Effective Teacher Professional Learning and Development

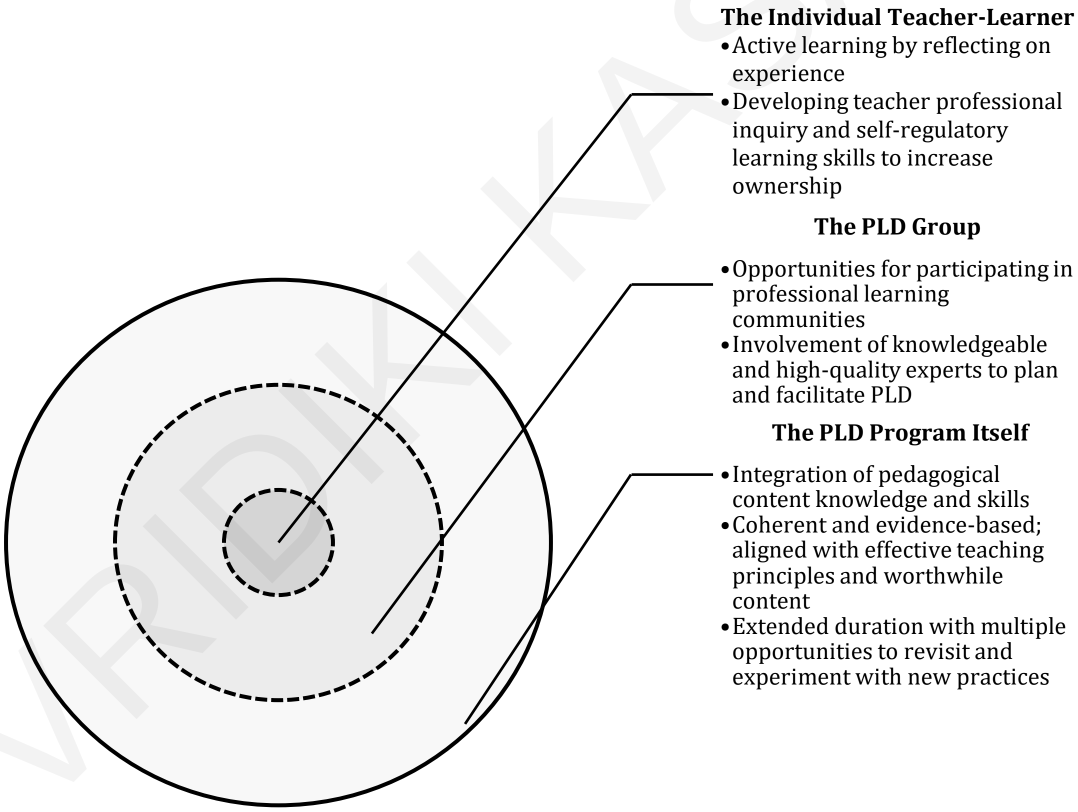


Figure 11 categorizes the seven critical features of PLD into three broad, nested categories or layers, as illustrated by the concentric circles. At its core is the individual teacher-learner (Layer 1), surrounded by the members of the PLD group in which the teacher participates (Layer 2), which, in turn, is influenced by the characteristics of the PLD program itself (Layer 3). Although these layers are depicted as distinct, the dotted lines illustrate that they are interrelated, with some features being central while others being embedded or connected. The subsequent subsection delves into the seven core PLD features, foundational to this conceptual framework, starting from the innermost layer and expanding outwards.

2.5.2.1 The Individual Teacher–Learner

Two features of PLD associated with individual teacher-learners have been identified as positively influencing their learning by making it more relevant and meaningful to them: active learning by reflecting on experience; and developing teacher professional inquiry and self-regulatory learning skills to increase ownership.

Active learning by reflecting on experience. Active learning in PLD is characterized by an inquiry-based approach, wherein continuous examination of practice and reflection on both professional and academic knowledge are central (Creemers et al., 2013; Merchie et al., 2016). This approach tends to encounter less resistance from teachers, as it positions them as co-creators of knowledge rather than passive consumers (Desimone, 2009; Merchie et al., 2016). Teachers' experiences act as a resource while they participate in cycles of challenging existing practices, acquiring new skills, implementing changes, working collaboratively to solve problems, as well as working on hands-on activities that directly relate to actual teaching and student learning (Desimone, 2009; Timperley, 2008). This approach mirrors the conditions under which students learn, emphasizing the construction and personalization of their own learning experiences, promoting a deeper understanding and a more meaningful implementation of new concepts and strategies in their teaching practice.

Active learning provides frequent opportunities for teachers to model and practice new skills, and critically examine and experiment with certain teaching practices, while facilitating both individual and collective guided reflection on their teaching experiences (Sims et al., 2021; Van Driel & Berry, 2012). Focused and deliberate reflection on practice serves as one of the two key mediating processes (the other being the experimentation with new practices) that facilitate the transition of changes in one domain (such as teacher knowledge or conceptualizations) to another (such as teaching practice) (Clarke & Hollingsworth, 2002). The interplay between enactive and reflective processes highlights the dynamic and non-linear nature of teachers' work. This balance involves continuous oscillation between classroom practice and reflective analysis, oriented around specific axes of teaching and learning

(Creemers et al., 2013). This intricate balancing act underscores the professionalism inherent in teachers' work.

Developing teacher professional inquiry and self-regulatory learning skills to increase ownership. Teachers benefit from directing their own learning paths, developing self-regulatory skills to identify both their students' needs and their own learning requirements (European Commission, 2013; Darling-Hammond et al., 2017; Trotter, 2006). This self-awareness approach enables teachers to actively seek feedback and effectively track their progress towards desired teaching outcomes (Timperley, 2008). Ownership and acceptance of PLD outcomes are enhanced when teachers voluntarily participate in PLD; are deeply involved in almost all phases of PLD development and implementation; exchange ideas with colleagues and develop shared values (European Commission, 2013; Jürimäe et al., 2014; Trotter, 2006). As Cordingley (2015) nicely puts it, "The process of determining with a partner how to tackle new approaches and coming together regularly to offer each other an ear and moral support is an effective catalyst for ownership of professional learning" (p. 5).

2.5.2.2 The Professional Learning and Development Group

Two key features have been identified that illustrate the dynamics of learning communities and the nature of the activities involved: opportunities for participating in professional learning communities; and involvement of knowledgeable and high-quality experts to plan and facilitate PLD.

Opportunities for participating in professional learning communities. Wenger and colleagues (2002) emphasized the significance of developing learning communities among individuals with common professional roles, such as teachers working in similar grades, subjects, or schools. These communities provide a supportive and trusting environment where teachers can freely exchange ideas and advice, collaborate on problem-solving, share and address common challenges, explore new teaching methods and tools; observe each other's teaching and give feedback (Merchie et al., 2016). In this collaborative atmosphere, teachers should engage in meaningful discussions and "collective critical reflections" on their practices (Creemers et al., 2013, p. 51; Darling-Hammond et al., 2017; Desimone, 2011; Villegas-Reimers, 2013). Over time, such interactions lead to the development of shared knowledge, practices, and a collective sense of identity among teachers.

Further underscoring the value of these communities, Timperley (2008) identifies collaborative professional learning communities as pivotal in facilitating teacher change, helping teachers overcome initial obstacles, and igniting their motivation for transformative change. This process can manifest in various patterns, including changes in practice but not beliefs, changes in beliefs but not practice, and changes in both practice and beliefs, indicating a multifaceted impact on teacher development (Tam, 2015).

Involvement of knowledgeable and high-quality experts to plan and facilitate PLD. While the group may have the potential to develop knowledge internally, external expertise is often essential in challenging existing conceptions, broadening horizons, assessing teacher needs, fostering the development of new knowledge and skills, developing deeper insights into their subject and student learning, and effectively implementing novel curricula (Cordingley et al., 2007; Darling-Hammond et al., 2017; Gallagher et al., 2017; Guskey & Yoon, 2009; Roth et al., 2011; Sims et al., 2021; Timperley, 2008).

Creemers and colleagues (2013) emphasize the importance of teacher educators in evaluating teachers' developmental needs and collecting data from their contexts and daily workplaces. They acknowledge that different groups of teachers prioritize different areas for improvement and are at various stages of professional growth, factors that can significantly influence the effectiveness of PLD. This situation necessitates the calibration of PLD programs to align with teachers' profiles, including their specific needs and their unique contexts (e.g., Agathangelou et al., 2024). From this perspective, teacher educators are encouraged "to think about the scope and sequence of teacher education experiences in the same way and with the same care that they develop scope and sequence guides for students from kindergarten to twelfth grade" (Creemers et al., 2013, p. 54). Such an approach involves carefully ranking or grouping teaching skills and incorporating a variety of PLD strategies into the program.

Teachers engaged in PLD efforts frequently find the complexities of teaching and the assimilation of new learning overwhelming without the support, challenge, and feedback provided by an expert in the field (Timperley et al., 2007). Feedback and guided reflection (discussed earlier), while distinct, synergistically function to guide teachers toward achieving expert levels of practice (Darling-Hammond et al., 2017). Teachers benefit from receiving frequent, rich, and constructive feedback after experimenting with a particular teaching practice—delivered in a manner that respects the professionalism of teachers and acknowledges their efforts towards change—by teacher educators, who read their lesson plans or observe their lessons and subsequently facilitate reflective discussions (Campbell & Malkus, 2011; Ingvarson et al., 2005; Shabani et al., 2010; Sims et al., 2021).

2.5.2.3 The Professional Learning and Development Program Itself

This category encompasses three overarching features of PLD, which, while not directly targeting the individual learner or the PLD group, are crucial when designing any PLD effort: integration of pedagogical content knowledge and skills; coherent and evidence-based; and extended duration, with multiple opportunities to revisit and experiment with new practices.

Focus on pedagogical content knowledge and skills. PLD programs that prioritize the most relevant and worthwhile content and pedagogical expertise and skills are crucial in enhancing teacher effectiveness (Creemers et al., 2013; Darling-Hammond et al., 2017; Desimone, 2011; Sims et al., 2021; Timperley, 2008). Content and pedagogy-focused PLD can support teachers to develop knowledge about mathematics content, students' mathematical thinking and work, and effective teaching practices (Darling-Hammond et al., 2009; Polly et al., 2014; Timperley, 2008). This process requires a focus on deep rather than shallow, superficial learning. Deep learning provides teachers with the necessary experiences and conceptual understandings to examine and reflect on their current beliefs, ideas, and practices, thereby improving their teaching (Hill et al., 2002). Such PLD ensures that teacher learning is directly applicable to the classroom setting and also extends beyond merely presenting them with models of expert teaching (Van Driel & Berry, 2012). Providing various examples or artifacts of practice, such as lesson plans, student work, worthwhile tasks, videotaped lessons, narratives, or curriculum materials, help teachers develop "a clear vision of what best practices look like" (Darling-Hammond et al., 2017, p. v; Sims et al., 2021).

Coherent and evidence-based. PLD must be a coherent endeavor, characterized by a clear vision and well-defined, meaningful goals, rather than a series of random and fragmented events, to achieve effectiveness (Guskey, 2000). It should also be *systemic* and aligned with broader trends in educational policy and research, effective teaching principles, and worthwhile content, ensuring that the PLD is relevant and up to date (ibid.). 'Systemic' implies that PLD is not an isolated endeavor but is deeply intertwined with the broader educational context. This means that critical elements such as the organization of the educational system, current trends in educational research, prevailing educational policies, and the involvement and support of schools and stakeholders are all integral to and exert a significant influence on teacher learning. Non-systemic PLD processes often encounter challenges due to their lack of clarity, misleading content, or misalignment with a coherent set of teaching practices (Darling-Hammond et al., 2017). As a result, teachers may struggle to understand how new innovations integrate with previously implemented strategies, leading to difficulties in adopting new approaches within a system that may inadvertently place barriers rather than support their work (Allen & Penuel, 2015).

Extended duration, with multiple opportunities to revisit and experiment with new practices. Research indicates that conceptual and teaching practice change needs PLD programs to be adequately lengthy (Desimone, 2009). According to Desimone (2009) this includes both the overall period over which the activity extends (for example, spanning an entire semester) and the total hours dedicated to the activity. She continues that while research has not pinpointed a precise "critical threshold" regarding duration, there is evidence favoring programs that are distributed across a semester (or intensive workshops with subsequent follow-up) and involve at least 20 hours of engagement.

Effective PLD should be a continuous process, enabling teachers to capitalize on numerous learning opportunities to enhance their teaching skills. Research indicates that it often takes a year or two for teachers to fully grasp and internalize how their existing conceptions and practices differ from those advocated by the PLD program, ultimately leading to meaningful changes (Kennedy, 2019; Timperley, 2008). This transformation requires substantial time, especially considering that many of the teachers' pre-existing beliefs and methods are challenged during PLD. In contrast, isolated workshops, sporadically held throughout the school year, tend to reinforce the notion that PLD is a separate entity, disconnected from the everyday responsibilities of teaching. Such an approach is less effective as it fails to integrate PLD into the continuous professional journey of teachers.

The synthesis of the four conceptual frameworks into three interconnected categories—individual teacher-learners, the PLD group, and the PLD program itself—provides concrete guidelines for studying and understanding the complex dynamics of PLD, emphasizing its multi-layered nature. Figure 11 suggests that PLD is a nuanced process that requires a deep understanding of the individual needs of teachers, the collective dynamics of professional learning communities, and the overarching goals and methodologies of the PLD program. Moving forward, these insights informed the development and implementation of a PLD program centered on issues of cognitive activation, differentiation, and their interplay (see Chapter 3). Since this program was based on video clubs, in the next section we elaborate upon this form of PLD, while also discussing how this specific PLD form incorporates the seven aforementioned features.

2.6 Video Clubs: A Model of Teacher Professional Learning and Development

Over the past decades, the landscape of PLD for teachers has undergone significant evolution. Moving beyond the traditional, less effective one-off workshop model, contemporary research has illuminated more impactful approaches to teacher learning (Villegas-Reimers, 2003, refer to Section 2.5.1 for details). This shift in perspective has given rise to innovative PLD models, such as lesson study (e.g., Fernandez & Yoshida, 2004), coaching (e.g., Desimone & Pak, 2017), and video-clubs (e.g., Sherin & Han, 2004), providing teachers with a diverse array of opportunities to enhance their professional competencies and knowledge. While many of these models are not novel, having been in use for decades, their renewed applications within the contemporary context of PLD have led to wider acceptance and recognition of their effectiveness in fostering professional growth (e.g., Kraft & Hill, 2020; Suh et al., 2021; Taras et al., 2022).

Table 5 presents a comparative overview of three major PLD models that have been extensively used in research: lesson study, coaching, and video-clubs. As models of PLD for teachers, lesson study, coaching, and video clubs share key elements, including collaboration, reflective practice, a focus on student learning, and practical application in real teaching scenarios. While all these models were found to be effective in developing teacher learning through the analysis of videotaped or lived lessons (e.g., Blazar & Kraft, 2015; van Es & Sherin, 2010; Yoshida, 2012), video-clubs appear to offer several benefits compared to the other two models.[18] For example, in lesson study, not every teacher may have the opportunity to teach and receive individualized feedback on their lesson plans (Lewis et al., 2006). Additionally, some teachers may view lesson plans as exhaustive detailed scripts rather than flexible

[18] This study does not assert that video clubs are categorically "better" than coaching and lesson study, as the effectiveness of each approach can vary based on specific educational contexts, goals, and personal preferences of teachers. However, video clubs have unique advantages that might make them more suitable or effective in certain situations.

frameworks, because of the emphasis on the lesson planning phase on the expense of the other lesson study phases (Wolthuis et al., 2020). Furthermore, coaching, while valuable, often lacks the collaborative culture in the sense of a learning community among teachers (Lynch, 2014). Also, its effectiveness heavily relies on the quality and skills of the coach, leading to inconsistent learning experiences and outcomes among teachers (Kraft et al., 2018). Hence, despite their affordances, both lesson study and coaching present unique challenges.

Table 5

Comparative Overview of Three Major PLD Models

Model	***Definition***	***Key Benefits***	***Main Challenges***
Lesson Study	A group of teachers plans, observes live classroom lessons, and refines lessons together to enhance classroom practice (Lewis et al., 2006)	Promotes collaborative learning, encourages collective lesson planning and reflective teaching	Involves one teacher performing the lesson and others observing and collecting data, may overemphasize lesson planning
Coaching	One-on-one or small group guidance to teachers by coaches who observe teachers' lessons and provide feedback to help them improve in day-to-day practice (Lynch, 2014)	Provides tailored support, enhances individual teaching skills, offers direct, personalized, and actionable feedback	Depends on the quality of the coach, promotes individual learning rather than a collaborative learning culture among teachers; is more demanding in terms of the personnel needed
Video Clubs	A small group of teachers meets to collectively review and discuss videoclips from each other's teaching (Sherin & Han, 2004)	Facilitates self- and group reflection, fosters a community of practice, teachers can review lessons multiple times with a different focus each time	Requires openness to peer feedback, lacks opportunities for individual feedback for the whole lesson, depends on the quality of the discussions

Although also having certain limitations (see Table 5 and discussion below), video clubs embody several of the features of effective PLD outlined earlier (see Section 2.5.2). They create a setting where small teacher groups actively engage in a cyclical process of designing, enacting, videotaping, and reviewing their own lessons, transforming their classrooms into dynamic learning environments (Van Es, 2012). This model not only fosters individual reflection but also encourages collective examination of the video-recorded lessons. Therefore, teachers become deeply and actively involved in their professional growth, gaining insights into student thinking, classroom interactions, and problems of practice through shared analysis of video clips (Van Es, 2012).

The opportunity to repeatedly visit and reflect on both personal and peer teaching practices guided by certain foci—with the possibility of analysis of a video-clip from different angles—enhances an ongoing learning and improvement culture (Borko et al., 2011; Sherin & Han, 2004). Extended collaboration allows teachers to delve into and understand teaching-related challenges more deeply (Van Es, 2012). The video-club model has the potential to generate a collaborative professional learning community (Alles et al., 2019), with the appropriate scaffolds from a facilitator (Dobie & Anderson, 2015). Typically, facilitators, who are subject-matter experts, guide productive video-based discussions effectively, concentrating on student thinking and the implications of specific teaching moves (Van Es & Sherin, 2008; 2010). Their role is crucial in maintaining the focus on lesson analysis, rather than teachers themselves, and fostering a positive, non-judgmental learning atmosphere. This environment encourages teachers to openly share their queries and concerns, contributing to a constructive culture where teaching methods and actions are not critiqued but explored for improvement. Video clubs represent a "highly adaptive [PLD model], designed to be readily responsive or adapted to the goals, resources, and circumstances of the local PLD context" (Koellner & Jacobs, 2015, p. 51). This PLD model thus offers a versatile platform for addressing a wide range of teaching and learning topics, accommodating diverse educational needs and interests.

Video clubs, while beneficial for teacher PLD, encounter some limitations. Their primary constraint lies in focusing mainly on selected teaching episodes for group analysis, rather than providing extensive, personalized feedback for entire lessons, potentially leading to a limited perspective on teaching practices. As video clubs often focus on specific instances, they might not fully represent a teacher's overall teaching approach. Also, the effectiveness of video clubs greatly depends on the quality of discussions. Additionally, privacy issues also arise, as teachers might be uncomfortable with filming their classes and peer scrutiny, which could further contribute to the risk of receiving negative feedback, potentially demotivating participants (Beisiegel et al., 2018). In addition, the varied levels of teacher engagement and interest can impact the dynamics and effectiveness of these video-clubs. Finally, maintaining ongoing interest and momentum in video club activities presents another challenge.

To address these limitations inherent in video clubs, a multifaceted approach is required. First, integrating full lesson discussions alongside the viewing of selected clips and providing teachers with individual feedback on both the selected clip and the entire lesson could offer a more comprehensive view of their teaching. Second, skilled facilitation can maintain productive, relevant, and focused discussions (Mitchell et al., 2022). Third, clear communication and consent protocols can alleviate privacy concerns, while fostering a positive, supportive group culture can reduce the risk of negative feedback (Xia et al., 2022). Fourth, involving teachers in setting the agenda and choosing videoclips from their own lessons can increase engagement and address individual needs (Beisiegel et al., 2018). Lastly, holding regular meetings (Beisiegel et al., 2018) and regularly assessing and adjusting the video-club activities based on teacher feedback can help sustain interest (Amador et al., 2023).

Based on the discussion regarding the three models, the current study opted to adopt the video-club approach. All the strategies to mitigate the limitations of video clubs have been carefully considered and incorporated in the development of the PLD video club program of the current study (see Chapter 3). The next parts delve into various aspects of video clubs: their objectives, key features, applicable contexts, and the type of learning teachers can develop through participation.

2.6.2 The Structure and Functioning of Video Clubs

Video clubs, a PLD model originally introduced in the USA during the early 1990s, involve small groups of teachers and one or more facilitators who meet regularly to watch and discuss video excerpts from each other's classrooms, focusing on specific areas of interest (Sherin & Han, 2004). Gaining popularity over the past two decades (e.g., Dobie & Anderson, 2015; Taras et al., 2022), this model centers around the use of video clips from actual teaching sessions (as reflected by the model's name), which serve as a tool for teachers to analyze and reflect on authentic representations of practice (Sherin & van Es, 2009). Unlike real-time classroom observations, video clubs allow teachers to repeatedly view lesson videoclips, enabling them to pause, rewatch, and analyze these clips from various perspectives. This form of PLD supports teachers in gaining insight into their own and their peers' teaching methods, fostering a collaborative space for discussing a wide range of teaching and learning issues.

A video club typically comprises five to six teachers who convene regularly for a series of meetings. Either a group member or a teacher educator can act as the facilitator, ensuring adherence to community norms and guiding the ensuing discussion.[19] Between meetings, teachers record a lesson centered on the previous meeting's focus. Each subsequent meeting involves sharing videoclips from teachers' lessons, sparking cycles of discussion and reflection on their practices. Teachers and facilitators individually view these videotaped lessons, deliberately selecting and bringing to the group specific episodes that highlight notable teaching aspects and can potentially stimulate productive discussions aligned with the video club's focus. These videoclips may showcase various classroom dynamics, such as students' work on tasks, whole-class interactions, or teacher-student engagements, and could include examples of effective teaching, typical classroom dilemmas, and diverse practices (Gaudin & Chaliès, 2015; Van Es & Sherin, 2010). Reflecting on these video recordings, teachers are encouraged to critically assess their teaching, recognizing that no teacher or teaching method is flawless. Having outlined the structure and functioning of video clubs, the next section delves into the specific objectives and intended outcomes of this PLD model, further elucidating how it serves to enhance teaching practices and foster professional growth among teachers.

2.6.3 The Main Goals of the Video-Clubs

The utilization of videos from actual classroom teaching in video clubs offers a valuable window into the examination and understanding of teaching practices. Watching themselves teach and reflect upon their video recorded lessons helps teachers improve their teaching practices (e.g., Borko et al., 2008). Video use in teacher PLD has grown significantly in recent years across various subjects, grade levels, and countries (Gaudin & Chaliès, 2015; Ramos et al., 2022). This increase is due to several factors: technological advancements facilitating the capture and storage of large video data, the limited chances for teachers to observe their colleagues' classrooms, and the unique capacity of videos to capture and allow for repeated, reflective viewing of classroom interactions from multiple perspectives (Brophy, 2004; van Es & Sherin, 2002, 2010). Additionally, videos offer exposure to diverse teaching styles and student populations, making them a valuable tool in teacher PLD (Gaudin & Chaliès, 2015).

A key objective of video clubs is to enhance teachers' abilities to notice, understand, reflect on, and interpret key moments and interactions within classroom

[19] The facilitator can be a teacher-leader, a coach, a district leader, or a researcher. Recent research has shown that the type of group facilitation (i.e., the group being led by an external member or by a video club participant) does not influence teachers noticing (identifying and interpreting) and focus on important mathematics teaching aspects (Mitchell et al., 2022).

settings, moving from a basic to a more in-depth, evidence-based approach, particularly regarding topics discussed in meetings (Rich & Hannafin, 2009; Van Es & Sherin, 2008). Considering the complexity and simultaneous occurrences in classroom environments, the task of identifying and analyzing significant or noteworthy events, and then relating them to broader pedagogical concepts, is quite intricate (Van Es & Sherin, 2010). As Lampert (2001) argued, transitioning from being a teacher to a 'reflective practitioner' can be challenging. However, this process allows teachers to examine the same teaching episodes from multiple perspectives, a depth of analysis less likely when working individually.

The use of video clips in this context is not aimed at judging teaching as good or bad, nor at directly emulating or discarding specific teaching methods. Instead, these clips are utilized as springboards for in-depth discussions and analyses of teaching and learning processes. They also serve to help teachers develop, refine, and sometimes challenge their own theories and understandings about teaching and learning (Borko et al., 2011; Gaudin & Chaliès, 2015). Hence, video analysis provides an opportunity to track changes in teacher learning and teaching development over time (Borko et al., 2014).

Although videos are beneficial in teacher PLD, merely watching and reflecting on classroom footage neither automatically guarantees the development of a professional learning community nor enhances teaching (Alles et al., 2019; Tekkumru-Kisa & Stein, 2015). Teachers, especially novices, may focus on superficial aspects when viewing a videoclip or hastily form conclusions from a brief classroom episode, failing to grasp the entirety of the lesson's context (Alles et al., 2019; Mitchell & Marin, 2015). In addition, teachers often stray from the video club's specific focus, veering into broader topics, while managing these deviations proved to be a significant challenge for facilitators (ibid.). In light of these insights, the following section explores conditions that need to be established for video clubs to be effective, addressing the challenges and maximizing the benefits of this approach in teacher PLD.

2.6.4 Conditions for Maximizing the Effectiveness of Video-Clubs

For the effective implementation of the video-club approach, it is crucial to develop a professional learning community among the participants, including the facilitator. Teachers need to be open to sharing their practices and videotaped lessons, and engaging in reflective discussions that focus on key aspects of teaching and student learning (Brantlinger et al., 2011). This involves creating a culture where teachers learn from their own and their peers' teaching experiences. While initially there may be hesitation to critique peers' practices, with the appropriate facilitator's support, teachers can progress toward providing constructive, non-judgmental feedback. The implementation and adherence to mutually agreed-upon discussion rules are crucial for creating a conducive learning environment for effective conversations about classroom interactions observed in video recordings from teachers' own classrooms (Alles et al., 2019).

Additionally, video-club participants must have a shared learning goal or focus, while the video-recorded lessons must be appropriately utilized to serve this common goal. Given the challenges in doing so outlined in the previous section, these highlight the necessity for guidance in identifying key aspects in video clips and effectively utilizing them for learning (Van Es et al., 2014). The video-club facilitator plays a crucial role in structuring and coordinating the viewing and discussion of videos to maximize their potential in PLD, tailored to the specific learning objectives and the targeted teacher-learner group (Blomberg et al., 2014).

The selection of video clips must be done with care and purpose, integrating them into PLD in ways that encourage teachers to notice and reason about critical incidents, while also allowing time for reflection (Borko et al., 2014; Sherin & Van Es, 2009; Tekkumru-Kisa & Stein, 2015). With familiarity with the content and critical incidents in the selected clips, the facilitator assists teachers in identifying and making "sense of what is captured in video, […] focus on the details of the interactions and draw informed interpretations of teaching and learning" (Van Es et al., 2014, p. 352). Several studies highlight the facilitator's vital role in ensuring that teachers derive benefit from analyzing and discussing "images of teaching" captured in the videos (e.g., Borko et al., 2014; Zhang et al., 2011).

Van Es and colleagues (2014) devised a framework to aid facilitators in engaging teachers in meaningful video-based discussions, identifying four key practices for productive discussions: orienting the group to the video analysis task; sustaining an inquiry stance; maintaining a focus on the video and subject matter (in this study mathematics); and supporting group collaboration. However, facilitators need to be mindful of how they challenge teachers' thinking, as overly pressing could be perceived as disrespectful. This is particularly important when the facilitator is a peer or less experienced than the teachers in the video-club group (Zhang et al., 2011). In summary, the success of video clubs hinges on creating a supportive, collaborative environment, guided by a skilled facilitator, where teachers can effectively engage with and learn from video-based analysis of teaching practices. Transitioning to the next section, we examine the empirical data that illuminates the effectiveness of video clubs and highlights open issues.

2.6.6 Research Evidence on the Video-Club Approach and Open Issues

The effectiveness of video clubs has been well-documented in numerous studies (e.g., Sherin & Han, 2004; Van Es & Sherin, 2008). Historically, video clubs have been used predominantly for PLD in mathematics (e.g., Van Es & Sherin, 2009; 2010), whereas more recently, they have also employed in science education (e.g., Barnhart & van Es, 2020; Luna & Sherin, 2017) and in PLDs integrating mathematics with language and literacy (e.g., Xia et al., 2022). Focusing on mathematics, the topics explored in these studies range widely. They include the development of professional noticing skills (e.g., Mitchell & Marin, 2015; Superfine & Bragelman, 2018; Sherin & Van Es, 2005; van Es & Sherin, 2008) and professional vision (e.g., Sherin, 2001; Sherin & van Es, 2009), analyzing and interpreting student thinking during teaching (e.g., Beisiegel et al., 2018; Sherin & Han, 2004; Sherin & van Es, 2009; Van Es & Sherin, 2010), analyzing classroom discourse (e.g., Brantlinger et al., 2011), planning and orchestrating productive discussions (e.g., Borko et al., 2014), enhancing mathematical problem-solving teaching (e.g., Yap & Leong, 2015), using teaching strategies like enablers and extenders (e.g., Charalambous et al., 2018; 2023a), and familiarizing teachers with observational tools such as the Mathematical Quality for Instruction (e.g., Mitchell et al., 2022). These studies collectively reveal that participation in video clubs can somehow improve teachers' focus on student thinking and classroom dynamics, cultivate their professional vision and noticing, adopt a less evaluative and more interpretive stance toward teaching, and foster innovative methods in analyzing both teaching and student learning.

The bulk of research in video clubs has predominantly focused on prospective teachers, with only a handful of studies delving into the experiences of practicing teachers, emphasizing the skills of noticing (Santagata et al., 2021). Evidence on how these video clubs contribute to changing teachers' practice is currently limited, underscoring the need for more research in this area. Research (e.g., Charalambous et al., 2018; van Es et al., 2017a; van Es & Sherin, 2008) has also started to identify diverse learning paths in actual teaching practice (among prospective teachers). These findings suggest that video club participants represent a diverse group with varied learning trajectories, rather than a homogenous entity, thereby emphasizing the importance of recognizing and catering to these individual differences in future video club research and implementation. To further enhance the understanding of the impact of video clubs on teacher learning and practice, it would be beneficial to explore the different learning paths among practicing teachers, especially since most current documentation of these varied trajectories pertains to prospective teachers.

In addition, while there seems to be increasing interest in exploring the contribution of video clubs to enhancing teacher learning and practice in ambitious mathematics teaching (e.g., Charalambous et al., 2022; 2023a; Sun & Van Es, 2015; Van Es et al., 2017a; 2017b), research still predominantly centers on prospective teachers. Specifically, prior research studies have highlighted the experimentation and learning experiences of both prospective and practicing teachers as they experiment with key tools of this approach, such as enablers and extenders, in the context of video clubs (Charalambous et al., 2022; 2023a). Additionally, research on developing ambitious mathematics teaching for prospective teachers encompasses elements like focusing on student thinking and understanding, managing complex tasks, employing effective questioning techniques, fostering productive classroom discourse, and implementing formative assessments (Sun & Van Es, 2015; Van Es et al., 2017a). Moreover, van Es and their team (2017b) have investigated how exceptional practicing secondary mathematics teachers understand and address equity in their teaching, exploring aspects such as teacher-student dynamics, resource allocation, classroom environment, and the nature of interactions. From the aforementioned research, a few important observations can be drawn: all the studies were qualitative, possibly driven by the typically small number of participants in video clubs; a very limited number focused on practicing elementary teachers; and each study highlighted the diversity in teachers' learning approaches and the varying levels of success they achieve in implementing ambitious mathematics teaching practices.

To date, no video-club study has specifically examined how teachers conceptualize cognitive activation, differentiation, and their interplay, nor how they implement teaching practices that support ambitious teaching. In fact, existing research on video clubs, especially in relation to ambitious mathematics teaching, is still developing. Hence, continued and more comprehensive research work in this area appears to be essential. Such research could not only enhance our understanding of video clubs as a tool for PLD, but also help improve their effectiveness in cultivating dynamic, reflective, and adaptive teaching approaches. Therefore, there is a clear need for further studies to investigate how teachers develop these conceptualizations and change their teaching methods through extended engagement in video-club programs focused on these specific educational aspects.

Having explored the potential of video clubs in PLD, it is crucial to consider how we can effectively evaluate their impact. This shifts the focus to choosing the appropriate model for evaluating PLD programs, by examining multiple levels of outcomes. The next section explains why Kirkpatrick's model (2007) was selected among other available PLD evaluation models to gauge the success of a video-club program focusing on issues of cognitive activation, differentiation, and their interplay.

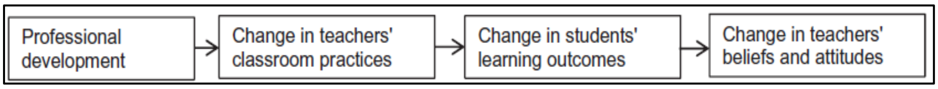
2.7 The Kirkpatrick's Model for Evaluating Teacher Professional Learning and Development Programs

Over the years, several models (see Figure 12) have been developed for theorizing, designing, analyzing, or evaluating PLD (e.g., Clarke & Hollingsworth, 2002; Desimone, 2009; Evans, 2004; Guskey, 2002; Opfer & Pedder, 2011).[20]

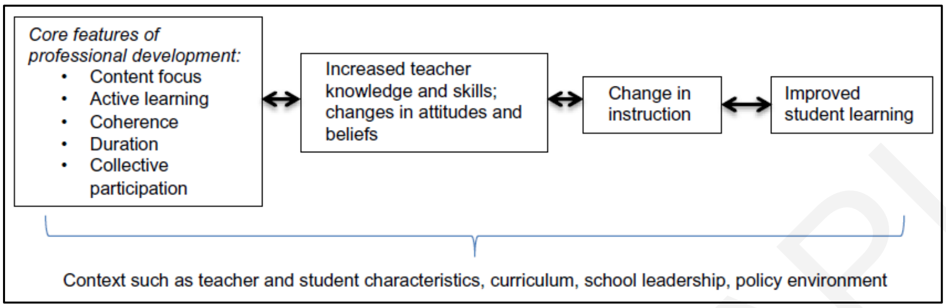
[20] Providing a detailed presentation of this presentation is beyond the scope of this presentation (for a detailed presentation of them see Boylan et al., 2017).

Figure 12.

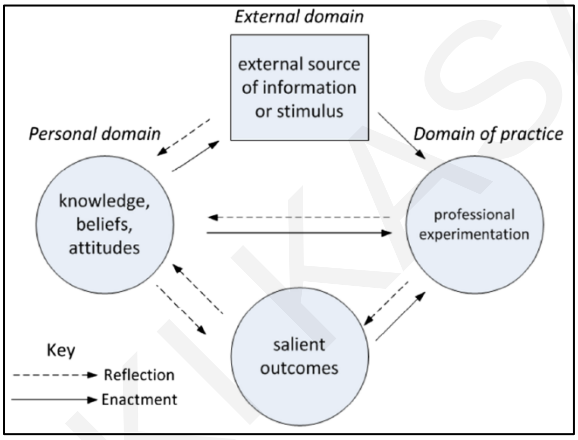
Models of Professional Learning and Development (adopted from Boylan et al., 2017)



Guskey's (2002) model.



Desimone's (2009) model.



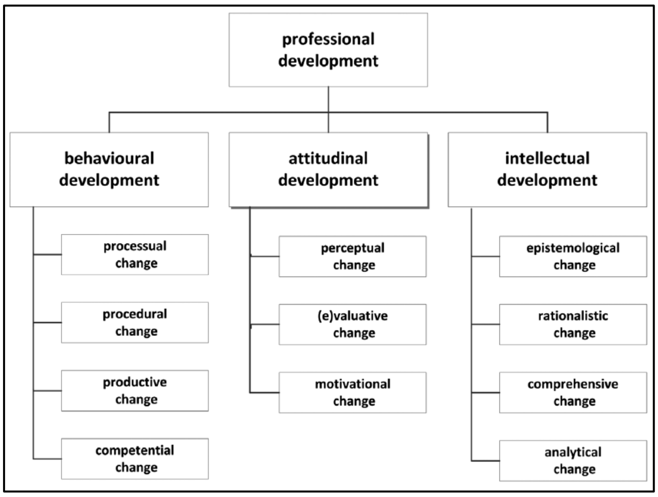
Clarke and Hollingsworth's (2002) model.

Elements: nested systems and subsystems

- learning activity system
- teacher
- school system

Relationships: Orientations between and to systems

Opfer and Pedder's (2011) model.



Evan's (2004) model.

Despite their potential utility, they also exhibit certain limitations (cf. Boylan et al., 2017). For instance, Guskey's (2002) linear model proposes a one-way causal progression in teacher PLD, potentially oversimplifying complex change processes. Building on Guskey's model, Desimone's (2009) non-recursive model allows for variability in the sequence of elements and considers some contextual factors, albeit statically. This model also focuses on measuring teachers' knowledge, which might affect teachers' willingness to participate in a PLD program, especially if received as evaluative. In contrast, Clarke and Hollingsworth's (2002) model, while complex, highlights multiple change pathways through enactment and reflection across four interconnected domains. However, it lacks specificity in connecting these domains, and delineating how changes in one domain directly lead to or mediate changes in another. The model's complexity and lack of parsimony pose challenges for PLD evaluation. Similarly, Opfer and Pedder's (2011) dynamic system-based model emphasizes change through interactions within nested systems but may not adequately focus on individual teacher characteristics and choices, complicating the evaluation process. Lastly, Evans' (2014) model, focusing on individual cognitive processes across behavioral, attitudinal, and intellectual dimensions, presumes a direct link between recognizing a 'better way' and changes in practice. However, it does not fully consider how existing habits, school culture, or resource constraints might hinder the application of new knowledge in the classroom. A common shortfall in these models is the omission of a reaction level evaluation, crucial for assessing teachers' perceptions, emotional responses, motivation, engagement, and alignment with cultural and contextual needs, as well as identifying areas for improvement in program design and delivery.

Given these limitations, the current study utilizes Kirkpatrick's model (2007) as a framework to evaluate a PLD program on ambitious mathematics teaching. Its four-level structure, encompassing reaction, learning, behavior, and results, addresses many of the shortcomings of the aforementioned models. The Kirkpatrick model not only evaluates participants' immediate reactions and learning outcomes, but also assesses the behavioral changes and actual results in practice. This holistic approach makes it well-suited for capturing the multi-faceted impacts of PLD, considering both individual and contextual factors, and providing insights into the overall effectiveness and areas for improvement in program design and delivery.

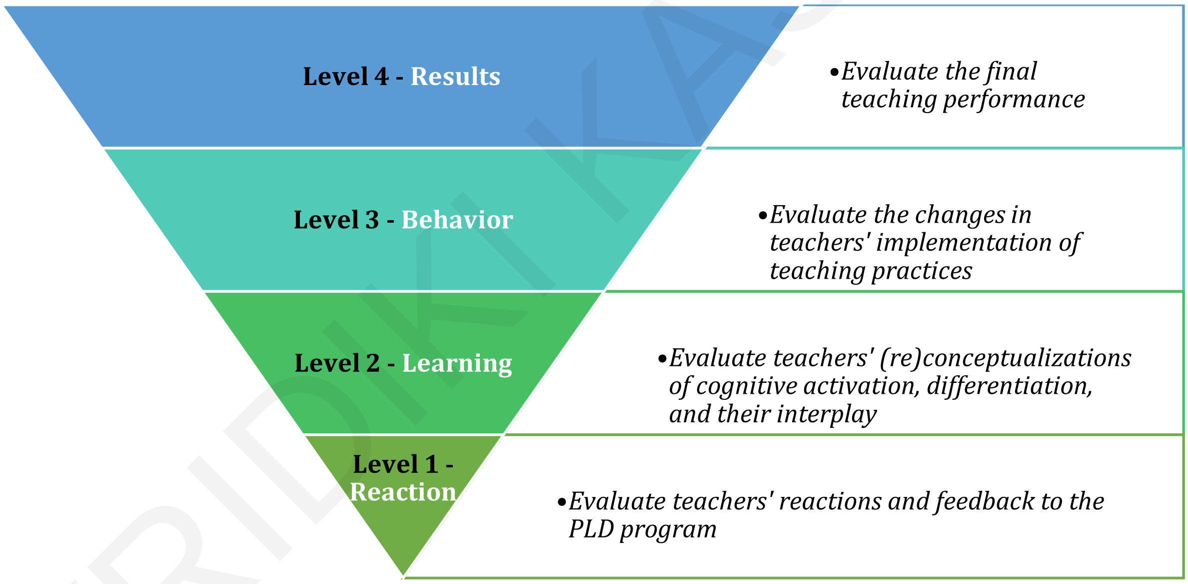
Initially introduced in the business sector in 1959, Kirkpatrick's model stands as a seminal evaluation tool and is versatile in the realm of business and education PLD

(Alsalamah & Callinan, 2021). While predominantly utilized in medical training evaluations, it also shows significant application in other fields, including social sciences (such as evaluation of head teachers, teachers, or students) especially during the past decade (ibid.). Despite encountering various critiques over time and the emergence of newer models, Kirkpatrick's model continues to be widely and increasingly adopted due to its simplicity, flexibility, adaptability, and practicality (Alsalamah & Callinan, 2021; Reio et al., 2017).

As illustrated in Figure 13, the Kirkpatrick Model proposes that a PLD program can be evaluated based on four levels: Reaction, Learning, Behavior, and Results (Kirkpatrick & Kirkpatrick, 2007).

Figure 13.

Schematic Illustration of the Kirkpatrick's Model, adapted to the Study's Focus



(adapted from https://www.valamis.com/hub/kirkpatrick-model)

At Level 1, the focus is on evaluating *the teacher learner's reactions to the PLD program*, by using reaction sheets or focus groups (Kirkpatrick & Kirkpatrick, 2007). This level seeks to investigate how satisfied teachers were with the PLD they received. This initial stage is crucial for multiple reasons. First, it provides an understanding of the perceptions, emotional responses, and engagement of the teacher participants in the PLD. Secondly, neglecting this level could lead teachers to perceive that the teacher educators assume they already know what is needed, implying that their feedback is undervalued for gauging the program's effectiveness (Kirkpatrick & Kirkpatrick, 2007). Focusing on the reaction level can help teacher educators elicit, formatively evaluate, address teachers' individual developmental needs, including the opportunities and constraints present within teachers' workplace (Creemers et al., 2013). Lastly, positive teacher feedback regarding the PLD on ambitious teaching can encourage future participation in PLD, while negative reactions might deter them, affecting their engagement and completion of such programs (Reio et al., 2017). However, positive reactions, while important, do not inherently guarantee that either learning (Level 2) or improved teaching performance (Level 3) have occurred, because teachers' reactions might be based on their enjoyment of the program, highlighting the importance of evaluating these subsequent levels.

Level 2 evaluates *what teachers have learned because of the PLD program and during the PLD program*, by employing mainly checklists or conducting performance appraisal interviews or meetings (Kirkpatrick & Kirkpatrick, 2007). Learning is defined as the extent to which participants have changed attitudes, improved, or increased their understanding of specific concepts and practices as a result of the program (ibid.). In this respect, learning is viewed as a process of conceptual change, highlighting the transformation of foundational understandings (Horn & Garner, 2022). Within the current study, Level 2 pertains to exploring how teachers (re)conceptualize cognitive activation, differentiation, and their interplay throughout their participation in a relevant PLD program. As Kirkpatrick and Kirkpatrick (2007) emphasized, evaluating targeted learning is crucial because without it, no change in behavior can occur. On the contrary, they also argue that learning does not necessarily lead to behavioral changes (Level 3) or improved teaching performance (Level 4). In such cases, learning might have occurred, but the teachers might have encountered some challenges that prevented or discouraged the application of their learning in actual teaching. Therefore, the model recommends also examining teachers' behavioral changes, specifically focusing on the evolution of teachers' teaching practices throughout the program.

Level 3 *evaluates observable changes in teachers' teaching behavior in their classrooms*. This evaluation, conducted mainly through observations or ideally a mix of observations and interviews, and secondarily through surveys, questionnaires, interviews, or focus groups, seeks to answer the pivotal question of what happens when teachers return to their classrooms after the PLD meetings. The true impact of PLD hinges on its practical implementation in the classroom; without the application of learned concepts, the program cannot contribute to organizational goals (Level 4). In addition, recognizing and quantifying changes in teaching performance acts as a reinforcement, encouraging the sustained adoption of new behaviors, while validating the PLD program's effectiveness in enhancing teaching practices (Kirkpatrick & Kirkpatrick, 2007).

Evaluating behavior is inherently more complex, challenging, and time-intensive than assessing reactions and learning (Levels 1 and 2). This complexity stems from the importance of *timing* in the evaluation of Level 3. Ideally, evaluations should be conducted both before, during, and after the program to capture the full extent of behavior change. However, the timing of these evaluations can vary based on the context and the nature of the PLD program. Additionally, the variability in how teachers respond to a PLD program underscores the need for a flexible and repeated approach to Level 3 evaluation. Teachers may exhibit immediate behavioral changes upon returning to their classrooms, while others might take months to adapt, or may not change at all. Additionally, some teachers who initially adopt new behaviors might revert to their old methods after an experimentation period. This diversity in response patterns requires allowing sufficient time for new behaviors to emerge and become established and conducting evaluations at multiple time points. By repeating the evaluation at strategically chosen intervals, a more comprehensive understanding of the effectiveness and sustainability of the behavior changes induced by the PLD program can be achieved. Level 3 evaluation serves as a critical level linking PLD efforts to the final outcomes (Level 4).

The final level assesses *the ultimate effects of the PLD program on key organization metrics, such as teacher teaching performance and student performance* ***after*** *the completion of the PLD program*. Teacher educators grapple with questions such as, "To what extent has teaching quality improved due to the PLD program?" or "What impact has the PLD program had on student performance?". These queries aim to define success from the perspective of stakeholders (Kirkpatrick & Kirkpatrick, 2007, p.197). Within the current study, the focus is on measuring quantifiable *end* results in teaching performance (in Chapter 3, we explain why it was not feasible to also examine student learning). It is important to note that the reinforcement of teacher learning into new teaching habits (Level 3), and the subsequent manifestation of results (Level 4), is a process that unfolds over time. Therefore, whenever feasible, data on teacher performance should be collected both before and after the PLD program.

The four levels of Kirkpatrick's model form a crucial chain of evidence to fully comprehend the success of a PLD program (Reio et al., 2017). Although the four levels follow a hierarchical order, presupposing an association between its levels, as already explained, Kirkpatricks highlight that favorable aspects in one level do not ensure success in later levels, questioning the assumption of a consistent correlation across all levels (Reio et al., 2017). Hence, while there is a relationship between the different levels of evaluation, the nature of this relationship is quite complex and multifaceted (Alsalamah & Callinan, 2021).

The combination of information from different levels provides a more comprehensive picture of the program's effectiveness (Kirkpatrick & Kirkpatrick, 2007). Therefore, it becomes important to examine the information each level provides about the effectiveness of a PLD program; each level might be integral to a comprehensive understanding of whether, how, and why the PLD program succeeded or failed. For instance, focusing solely on the final outcomes (Level 4) might not yield substantial insights. However, observing the trajectory of teaching practices and the challenges a teacher encounters during the experimentation with such practices throughout the PLD program (Level 3) could provide a better understanding of the success level of the PLD initiative.

Kirkpatrick's model can significantly enhance the evaluation of ambitious mathematics teaching programs in PLD. Specifically, the use of this model overcomes the limitations of previous studies (reported in Section 2.4.4), which primarily relied on solely teacher end-of-intervention self-reported data or solely quantitative data. Unlike these earlier methods, Kirkpatrick's model facilitates a mixed-method approach, to provide a holistic analysis and documentation of changes in both collective and individual teacher learning and practice in real classroom settings. Moreover, this multi-level model offers valuable insights into changes in teachers' conceptualizations (Level 2) and teaching practice (Level 3) and can elucidate any interconnections between them.

2.7 Summary

This chapter explores six key areas essential for addressing the research questions: Cognitive Activation, Differentiation, Ambitious Teaching, Professional

Learning and Development (PLD), Video Clubs, and Kirkpatrick's Model. Each section sheds light on significant findings and identifies research gaps that this study seeks to address; these are succinctly summarized next.

Cognitive activation, a pivotal concept in this study, refers to the engagement level and thinking processes students undergo when interacting with mathematical content. Teachers play a crucial role in this process, especially in selecting and implementing suitably mathematically challenging tasks. Classrooms exhibiting high cognitive activation and consistent task demands often see superior learning outcomes across varied educational contexts. Tools like the TAG and MTF are instrumental in categorizing tasks based on cognitive demands and guiding effective implementation. However, maintaining task rigor in classroom settings is challenging. Teachers frequently grapple with preserving the complexity of tasks, risking a reduction in cognitive demands during lesson enactment. Notably, PLD that focuses on choosing and enacting mathematically challenging tasks can bolster teachers' ability to attend to student thinking and sustain task demands. Although different PLD programs have focused on cognitive activation, further research is needed in this area to identify and comprehend the specific challenges teachers face in such PLD initiatives, by exploring how teachers conceptualize cognitive activation and apply their PLD learning in real classrooms.

Differentiation, the second focal point of this study, represents a teaching philosophy and practice aimed at meeting diverse student needs. It involves qualitative adjustments beyond mere variations in workload. As a pedagogical approach, differentiation balances attention between individual student groups and the entire class, ensuring each student engages meaningfully with core content for maximal learning. Despite its acknowledged importance, differentiation is subject to misconceptions, and teachers often struggle to integrate it into their practice. In this context, PLD is indispensable for enhancing teachers' understanding and skills in applying differentiation. Nonetheless, notable research gaps exist in PLD focusing on differentiation: PLD often relies on self-reported data; lacks practical application; omits direct observation and long-term tracking; and does not focus on conceptual shifts in teachers' understanding. Therefore, scholars advocate understanding the challenges in implementing differentiation and call for ongoing PLD to equip teachers with the necessary skills.

The third crucial area, ambitious mathematics teaching, is defined as teaching that engages students of all backgrounds in deepening their understanding and engagement in complex problem-solving. This kind of teaching demands expertise in identifying, selecting, using, and modifying mathematically challenging tasks to suit classroom diversity and make task complexity accessible to all students. It represents a synergy between cognitive activation and differentiation, requiring constant adjustment in teaching variables in response to student interactions, with presenting a challenging stimulus to all students with varied complexity levels. Research highlights ambitious mathematics teaching's positive impact on student performance, emphasizing the critical role of PLD programs in supporting teachers. Key gaps in ambitious teaching literature include reliance on self-reported data and simulated contexts, which may not fully represent classroom complexities as observed lessons would. More studies are therefore needed to explore differential changes in teacher learning and teaching practice throughout PLD programs, emphasizing the importance of mixed-method research. Further research is also required to understand how teachers' perceptions of ambitious teaching evolve through PLD and how these changes relate to their teaching practices.

The fourth theme considered in this chapter pertains to teacher PLD, encompassing structured activities aimed at enhancing teachers' knowledge, skills, attitudes, and practices. Moving beyond the traditional view of sporadic workshops, PLD is increasingly seen as an ongoing process integral to a teacher's work environment. This perspective recognizes the dynamic nature of teacher learning and the importance of contextually relevant PLD. Research shows variations in PLD effectiveness, with some initiatives resulting in minimal change in teaching practices. This inconsistency highlights the need to identify key features of effective PLD. Drawing on meta-analyses and meta-syntheses, in this literature review we identified seven key features of effective teacher PLD and synthesized them into a three-level structure. These features include active learning through reflection on experience and developing teachers' professional inquiry skills (individual teacher-learner level), participation in professional learning communities and involvement of high-quality experts (PLD group level), and a focus on pedagogical content knowledge and skills, coherence, evidence-based approaches, and extended duration (PLD program level). Drawing on these features, this study aims to portray how practicing teachers *initially* conceptualize and implement cognitive activation, differentiation, and their interplay, and the extent to and ways in which their conceptualizations and practices *change* through participation in a PLD program characterized by these seven effective features.

Considering a particular PLD model where teachers analyze video excerpts from their lessons, the fifth theme focuses on Video Clubs. Among existing PLD models, video clubs embody most of the seven aforementioned features of effective PLD, making them a suitable choice for this study. Video clubs have been shown to effectively improve teachers' skills, such as professional noticing, interpreting student thinking, and enhancing classroom discourse. Despite their benefits, research on video clubs has primarily involved prospective teachers, with a limited number of studies focusing on practicing teachers, let alone on ambitious mathematics teaching. Therefore, more in-depth research is needed to understand the impact of video clubs on how practicing teachers conceptualize and implement cognitive activation, differentiation, and their interplay in their practice.

To evaluate the video-club program's impact, we also introduce and discuss Kirkpatrick's model (the sixth theme of the chapter). This model evaluates teachers' reactions to the PLD program (Level 1), what teachers have learned from the PLD program, including changes in conceptualizations of specific ideas (Level 2), observable changes in teachers' teaching behavior (Level 3), and the ultimate effects of the PLD program on key metrics, such as teaching performance and, where possible, student performance (Level 4). This model not only addresses the limitations of previous studies that relied solely on self-reported data or quantitative analysis but allows for a mixed-method approach, offering insights into changes in teachers' conceptualizations and practices, a main focus of this study.

In conclusion, this literature review brought to light several trends and unresolved issues in teacher PLD, particularly focusing on cognitive activation, differentiation, and their interplay, emphasizing—yet, often overlooking—the gradual, non-linear, repetitive, dynamic, and complex nature of teacher learning and practice (Goldsmith et al., 2014). Scholars advocate for a paradigm shift in PLD research, urging for a focus on the *coevolution* of teacher learning and practice (Helsing et al., 2008; Kazemi & Hubbard, 2008). This involves examining both what teachers learn and how their teaching practices evolve, as well as understanding the comparisons between these aspects. Recognizing that the impact of the same PLD may differ significantly among individuals, it is essential to conduct research that addresses both *collective* and *individual* teacher learning and practice trajectories (e.g., Schoenfeld, 2023; Goldsmith et al., 2014; Witherspoon et al., 2021). Understanding the possible gains teachers make from PLD programs is key to developing support strategies that effectively integrate new concepts into existing practice systems (Kennedy, 2016). The following chapter offers a detailed account of the methodological approaches used in this study to address the research gaps identified in the literature review.

CHAPTER 3. METHODOLOGY

This chapter delineates the methodology employed to explore the impact of a PLD program focusing on issues of cognitive activation, differentiation, and their interplay on its participating teachers. The program, leveraging a video-club format and targeted PLD materials, aimed to enhance teacher competencies in these critical areas. A convenient sample of eight Cypriot elementary school teachers with diverse background characteristics participated in the program. A rich dataset comprising of videotaped lessons, lesson plans and ancillary materials, pre- and post-lesson teacher interviews, videotaped video-club sessions, teacher reflection cards, and end-of-program teacher interviews, was collected from this sample. A convergence model of mixed-methods triangulation design facilitated the integration of quantitative evaluations of teaching performance (using descriptive and inferential statistics) with qualitative explorations of conceptual development, teaching behaviors, and encountered challenges (using thematic analysis). Data analysis drew upon Kirkpatrick's model, assessing the program's impact across Reaction, Learning, Behavior, and Results levels first for the whole group and then for different cases. This dual-level analysis aimed to reveal both collective trends and individual learning paths. Despite inherent methodological constraints, measures to enhance the study's trustworthiness were implemented, ensuring the reliability of the findings.

3.1 Research Design

This study employed the *convergence model of mixed methods triangulation design* (cf. Creswell & Plano Clark, 2017) to evaluate the effectiveness of a PLD program on practicing elementary teachers' understanding and implementation of cognitive activation, differentiation, and their interplay. Specifically, it traced the coevolution of their conceptualization, experimentation, and teaching practice during the PLD program, identified the challenges they encountered, and evaluated their teaching performance and reaction to the program. This design facilitated the combination of qualitative and quantitative data collection and analysis methods, with *equal weight* given to both.

The integration of *mixed methods* allowed to address the research questions from multiple complementary angles, ensuring a more comprehensive understanding of the research problem at hand. The quantitative methods yielded valuable insights into the teachers' final teaching experimentation and performance in their culminating lessons, as well as any quantifiable modifications in their experimentation during the PLD program. The qualitative component captured nuanced, context-specific details of the teachers' conceptual reorientations, challenges faced, changes they might have introduced in their teaching practice over time, and their reaction to the program—aspects that might be obscured by quantitative data alone. Hence, the use of mixed methods mitigated the inherent limitations of each data collection method by leveraging the strengths of both quantitative and qualitative approaches.

After the separate analysis of the quantitative and qualitative data, *triangulation* was achieved by comparing and contrasting findings from both approaches. This process identified *convergences* and *divergences* within the data, effectively merging the two datasets into a unified interpretation. Such cross-validation of findings constituted the cornerstone of the research design, contributing towards enhancing the study's rigor, credibility, and validity. The next section provides background information on the research project.

3.2 Background of the EDUCATE Project

Data were drawn from a teacher PLD program of a larger ERASMUS+ European project, namely EDUCATE[21], geared toward helping the teachers improve the quality of their teaching with respect to issues of cognitive activation, differentiation, and their interplay. The main goal of the project was to develop, implement, validate, and refine materials for teachers and teacher educators that address issues of cognitive activation and differentiation in an integrated manner. Even more crucially, the project aimed at educating teachers to use these materials by engaging them in guided reflection around their practice and to scaffold teacher educators to offer solid guidance to teachers in doing so.

Involving organizations from four European countries (i.e., Cyprus, Greece, Ireland, and Portugal), the project unfolded in four phases. In the first phase, European, international, and national policy documents on cognitive activation and differentiation, as well as prior studies on teachers' needs and challenges when having to teach for both goals were reviewed. This top-down approach was complemented by a bottom-up approach involving lesson observations of a maximum variation sample of prospective and practicing elementary or secondary teachers, and pre- and post-lesson interviews with them to identify their actual and perceived needs, as well as challenges when having to engage *all* their students in challenging tasks.

[21] This acronym stands for "**E**nhancing **D**ifferentiated Instr**u**ction and **C**ognitive **A**ctivation in Mathematics Lessons by Supporting **T**eacher L**e**arning" (https://websites.ucy.ac.cy/educate/en/).

In the second phase, building on this needs-assessment analysis, five modules for teachers and five associated modules for teacher educators were developed, which aimed at helping them effectively deal with the most crucial challenges identified. Academics, teacher educators, and teachers from the four participating countries collaborated on developing these materials. Moreover, practicing teachers and school inspectors' feedback on the clarity of the modules, their reasonableness, applicability, and usefulness was solicited, to revise and improve the modules.

After being content-validated and pilot-tested, these modules were implemented with elementary and secondary teachers (prospective and practicing) working in different school settings (including underserved areas) in the four participating countries. During this third phase, recruited teachers worked in video-club settings through iterative cycles of experimentation with ideas included in the EDUCATE materials (which structured and enriched the video-club sessions) and guided reflection upon their practice, facilitated by teacher educators. Teachers collectively worked on analyzing and reflecting upon their practice in light of these ideas, thus setting new goals for improvement. Teacher educators used the respective teacher educator modules to scaffold teachers' work. In the fourth and final phase, the effectiveness of the EDUCATE materials and PLD approach was examined. The project culminated by revising the EDUCATE materials based on the lessons learned from this examination and by producing an e-learning course for teachers and teacher educators.[22]

The current study builds on the rich dataset collected in the third phase of the project. This phase not only yielded valuable insights into the effectiveness of the PLD approach but also delved into a deeper understanding of the participating teachers' conceptualizations and practice. In the following section, information about the sample and sampling procedures employed is provided.

3.3 Sample and Sampling Procedures

In the third phase of the EDUCATE program, 17 practicing Cypriot elementary teachers voluntarily participated in the PLD program, alongside five teacher educators. The teacher participants were split into three video-club groups (Group A was formed by eight teachers; Group B included six teachers; and Group C consisted of three

[22] All the materials used in the PLD sessions are freely accessible and hosted on the online EDUCATE platform, which can be found at http://educate-platform.com/.

teachers) according to the school and/or district they served, due to practical issues (e.g., convenience to meet). This criterion would also give the opportunity to teachers working in the same school or district to collaborate with their colleagues and share any common concerns and challenges they faced.

At least two teacher educators, who were experts in the domains of cognitive activation and/or differentiation, were recruited for each group to act as facilitators of the video-club sessions and provide support, feedback, and guidance to teachers. Specifically, four teacher educators facilitated the first group (including the author), four teacher educators facilitated the second group, while the third group was facilitated by two teacher educators (N=5). Because the EDUCATE project was used as a learning site not only for teachers but for facilitators as well, both seasoned and less experienced facilitators were recruited. The seasoned facilitators utilized this opportunity to further develop their skills, while the less experienced ones used it as a valuable learning opportunity to gain experience and expertise in video-club facilitation.

Sampling was conducted at two stages: initially at the *PLD group level*, followed by the level of *selected cases of teachers*. At the first stage, in which the *group-as-a-whole* was utilized as the unit of analysis, the sample for the present study was drawn from Group A due to the author's role as a teacher educator within the group, in contrast to the other groups in which she had no significant role. The author actively participated in almost all stages of designing and facilitating the video-club sessions as well as collecting data for the first group. This active involvement allowed for a better understanding of the group and the flow of the video-club sessions, including the interactions occurring during the sessions; the relationships that developed among the participants; the contributions and reflections of the teachers involved; as well as their individual traits and backgrounds. Consequently, the author gained a deeper insight into the teacher participants of this group compared to those in the other groups. Although a potential bias may arise due to the author's dual role as a researcher and a teacher educator, certain measures were taken to mitigate it, as will be explained later (i.e., through the use of the research's autobiographical note, see Section 3.8).

Although being a *convenient sample* (Patton, 2015), there was some variation in a number of background characteristics of the recruited teachers (N=8), such as their experience, credentials (additional qualifications), grades at which they were teaching mathematics, and the student populations they served. With respect to the last criterion, about half of the recruited teachers were working in schools serving disadvantaged students and/or diverse student populations. This helped develop deeper insights into the challenges that teachers faced when they sought to implement differentiation to offer high-quality learning that meets all their students' needs and levels. Table 6 summarizes the demographic information of the teachers (all females).

At the second sampling stage, certain *cases of teachers* out of the eight recruited teachers of the video-club group were selected as the unit of analysis. Following *heterogeneous sampling* (Patton, 2015), *three* teacher cases—Kate, Pina, and Michelle (all pseudonyms)—were selected, differing in various characteristics (including credentials; years of teaching experience; grade being taught during the PLD program; prior knowledge and experience on issues of cognitive activation, differentiation, and their interplay; and motivation to participate in EDUCATE, see Table 6). Additionally, quantitative variations were also identified in the evolution of their teaching practice across their lessons (see, Figures 24, 30 and 37 in Chapters 5, 6, and 7, respectively). Hence, these three teacher cases were deliberately selected to represent diverse trajectories of learning and teaching development among the group participants. The goal was to capture the widest range of variation within the PLD group possible and potentially facilitate a better understanding of teacher learning and practice.

Table 6
Demographic Information of the Study Sample

Name[1]	Ariana	Georgia	Kate	Michelle	Nancy	Pina	Souzana	Stella
Teaching experience	13 yrs. (five in upper grades)	24 yrs. (21 in upper grades)	18 yrs. (four in upper grades)	13 yrs. (four in upper grades)	12 yrs. (two in upper grades)	15 yrs. (two in upper grades)	13 yrs. (nine in upper grades)	12 yrs. (six in upper grades)
Credentials	BA in Elementary Education	BA in Elementary Education	BA in Elementary Education	BA in Elementary Education	BA in Elementary Education	BA in Elementary Education	BA in Elementary Education	BA in Elementary Education
	MA in Mathematics Education	MA in Educational Leadership	MA in Language and Cultural Education	-	MA in Art Education	MA in Mathematics Education	-	MA in Mathematics Education
	PhD candidate in Mathematics Education	-	-	-	-	-	PhD in Educational Policy and International Comparative Education	-
Grade taught during the PLD program	Grade 4 (upper elementary grade; students' ages 9-10 yr.)	Grade 6 (upper elementary grade; students' ages 11-12 yr.)	Grade 6 (upper elementary grade; students' ages 11-12 yr.)	Grade 5 (upper elementary grade; students' ages 10-11 yr.)	Grade 5 (upper elementary grade; students' ages 10-11 yr.)	Grade 3 (lower elementary grade; students' ages 9-10 yr.)	Grade 5 (upper elementary grade; students' ages 10-11 yr.)	Grade 5 (upper elementary grade; students' ages 10-11 yr.)
Prior knowledge or experience in the PLD program axes	No experience	No experience	Prior experience in differentiation *(had attended a two-year relevant PLD program)*	No experience	No experience	Prior experience in cognitive activation and challenging tasks *(had attended a relevant PLD program)*	Prior experience in differentiation *(had attended university courses and sporadic seminars)*	No experience
Motivation to participate in the PLD program	Passionate about mathematics teaching; attracted to the focus on cognitive activation and, notably, on differentiation; viewed the extended PLD duration as advantageous.	To enhance her mathematics teaching skills; excited about exploring cognitive activation and differentiation; and considered the program as an excellent PLD opportunity.	To deepen her differentiation knowledge and skills; curiosity to learn what cognitive activation is.	Limited knowledge of the axes and their entailments; recognized the need for PLD on these axes.	Focused on teaching mathematics; typically avoided long PLD programs but was keen on addressing diverse student groups.	Previous positive experience of videotaping teaching in a PLD program focusing on using challenging tasks, conducted by one of the teacher educators.	Valued ongoing education and PLD; to expand her differentiation skills, apply new teaching practices, and get feedback on their implementation.	Keen to stay informed and current in mathematics teaching; viewed the extended duration of the PLD as advantageous.

Note: 1. All teacher names are pseudonyms; cases are presented alphabetically.

3.4 The Intervention

The PLD intervention lasted approximately six months during the school year 2018-2019, commencing in October 2018, and concluding in April 2019. In total, nine 2.5-hour afternoon video-club sessions (during out of school hours) were conducted with the group, with all sessions being videotaped. These video-club sessions were organized every three to four weeks to keep participants focused on the process of learning and teaching improvement. The total duration of these sessions was 22.5 hours.

Teachers engaged in a *cyclical* process of PLD. Almost in between every session, teachers planned, taught, and videotaped a lesson (in any mathematical content) trying to experiment with ideas discussed and tackle the focal challenge of the video-club session that preceded. The teachers had the option to start videotaping lessons right after the end of the first session till the very end of the academic year in which they received the PLD program. In every subsequent session, the teachers, who videotaped a lesson before the session, shared videoclips of their videotaped lessons related to this challenge, identified complexities in this work, and discussed the extent to which the strategies listed in the teacher modules for dealing with the challenge were helpful. Working collaboratively, teachers analyzed records of practice drawn either from the teacher materials or from their videotaped lessons. They also collectively proposed additional ideas, practices, or strategies, for handling the focal challenge, which were documented and shared with the participants via email after each session. Based on teachers' needs and feedback, sessions utilized Modules 1, 2, and 5, which emphasized engaging all students in challenging tasks, lesson planning considerations, and fostering an inclusive classroom culture, respectively. Concepts from autonomous work (Module 3) and whole-class discussions (Module 4) were integrated as needed. The same cycle (i.e., discussing, experimenting, and reflecting upon a challenge) was repeated as teachers progressed from one module to another.

Almost all video-club sessions had a similar structure. Each session began with a video-club component. During this part, the participants watched and discussed excerpts from videotaped lessons included in the EDUCATE materials. From the second session onwards, the videoclips shown during this part, came from the teachers' lessons videotaped before each session. The aim of doing this was to reflect upon one's lesson in terms of how they had dealt with various cognitive activation and/or differentiation issues. After this, several activities from the PLD materials were enacted, to focus on these issues in more depth and detail. At the end of each session, teachers and teacher educators were asked to complete reflection cards and logs, respectively. Providing feedback on the implementation of the PLD program, these logs and reflection cards were used formatively to inform and refine its implementation. Table 7 details the focus, content, and unfolding of the nine video-club sessions, encapsulating the learning experiences of the participants.

Table 7

Overview of the Content and Unfolding of the Video-club Sessions

VCS[1]	Focus	Content and unfolding
VCS1	Discussing cognitive activation and differentiation, with emphasis placed on the factors which may promote or hinder cognitive activation	▪ Brainstorming factors that help/hinder engaging students in mathematical thinking and reasoning. ▪ Identifying features of challenging tasks ▪ Sorting tasks according to their level of challenge ▪ Discussing the importance of challenging tasks ▪ Identifying teacher actions that contribute to presenting and enacting a task at a challenging level
VCS2	Considering issues pertaining to differentiation	▪ Sharing participants' views on differentiation and reflecting upon common teacher perceptions and misconceptions on differentiation ▪ Identifying and discussing different differentiation practices ▪ Discussing how differentiation can be implemented in real-class settings
VCS3	Considering issues pertaining to the interplay of the two axes, with a particular focus on how to organize a whole-class discussion	▪ Compiling a list of criteria and techniques for selecting which students' solutions to share ▪ Discussing video footage in which other teachers were (un)successful in activating students during whole class discussion. ▪ Discussing the challenges faced during this lesson phase and suggesting some solutions for addressing them. ▪ Discussing the relation of the two axes and concluding that cognitive activation and differentiation are mutually supportive
VCS4	Focusing on questioning as a means to achieve both cognitive activation and differentiation	▪ Summarizing strategies for orchestrating whole-class interactions ▪ Focusing on questioning and how it can be used to scaffold students' work at different lesson phases, proposing questions for supporting students at each of these phases. ▪ Discussing aspects to consider while using questioning
VCS5	Modifying challenging tasks to address different student needs and levels	▪ Identifying practices that either promote or hinder cognitive activation and differentiation in illustrative clips of the study participants' attempt to promote both axes through questioning. ▪ Introducing the idea of enablers and extenders ▪ Generating enablers/extenders for given curriculum task
VCS6	Generating a list of steps to plan a lesson aiming to promote cognitive activation and differentiation	▪ Continuing work on enablers and extenders ▪ Generating a list of steps to be followed during lesson planning in order to concurrently address cognitive activation and differentiation. ▪ Working in groups to prepare a new draft plan for a lesson they would teach, in the light of what had been discussed
VCS7	Compiling a list of tools and ideas which could help teachers' lesson planning in analyzing students' prior knowledge and anticipating students' conceptions and solution approaches	▪ Sharing and discussing experiences from utilizing the list of steps generated in VCS6. ▪ Considering the idea of anticipating students' prior conceptions or alternative ideas while planning; discussing and considering alternative ideas and strategies that students may have for particular tasks. ▪ Considering methods of formative assessment to elicit student thinking and understanding

VCS8	Focusing on teachers' enduring challenges pertaining to lesson planning and enactment with respect to cognitive activation and differentiation	▪ Sharing some of the challenges that teachers had faced during their videotaped lessons; initiating a discussion among the participants around issues pertaining to the scaffolding of less advanced students. ▪ Working in groups to collectively prepare a lesson plan on fraction division
VCS9	Creating a classroom culture to foster cognitive activation and differentiation	▪ Surfacing classroom culture issues which foster cognitive activation and differentiation. ▪ Culminating the PD program by raising a broader discussion around things teachers learned, issues they still needed to consider, and features of the PD program they thought to be more/less useful

Notes.
[1] VCS#: Video-club session's ordinal number.

During the planning and implementation of the intervention, an effort was made to consider the features of effective PLD as delineated in the literature review (see Chapter 2, Figure 11). This ensured that the intervention was not only informed by research findings but also tailored to the needs and contexts of the participating teachers, thereby enhancing the potential for meaningful learning growth and changes in their teaching practice. Table 8 outlines the incorporation of these features into the PLD program.

Table 8

Incorporation of features of effective PLD into the EDUCATE Intervention

Feature	How each feature was incorporated into the PLD program
The Individual Teacher-Learner	
Active learning by reflecting on experience	▪ Teachers videotaped their experimentation with various practices; reflected on it; brought up concerns for discussion; and engaged in activities, like task analysis and reviewing practice artifacts (e.g., lesson plans, student solutions, etc.).
Developing teacher professional inquiry and self-regulatory learning skills to increase ownership	▪ The use of tools, materials, or practice artifacts provided opportunities for professional inquiry and for navigating challenges. ▪ The monitoring mechanism through the use of teacher reflection cards, guided PLD program adjustments aiming to increase teacher ownership.
The PLD Group	
Opportunities for participating in professional learning communities	▪ Facilitated by the teacher educators, teachers openly shared practices and videoclips with peers; pursued common goals of enhancing their teaching; and partook in reflective talks on teaching and learning.
Involvement of knowledgeable and high-quality experts to plan and facilitate PLD	▪ Experts on cognitive activation and/or differentiation planned and facilitated the sessions, ensuring a safe environment for sharing videos. They maintained focus on video analysis and mathematics, encouraged inquiry over evaluation, supported collaboration, addressed individual teacher needs, and if requested and given the time constraints, provided personalized feedback on teachers' practice in one-on-one meetings.

The PLD Program Itself	
Integration of pedagogical content knowledge and skills	▪ Research tools, materials, or practice artifacts were connected to different PCK aspects, such as setting learning objectives, analyzing tasks, attending to student thinking, or formatively assessing their prior knowledge.
Coherent and evidence-based; aligned with effective teaching principles and worthwhile content	▪ Teacher materials were developed with insights from literature on cognitive activation and differentiation, incorporating research excerpts, and using a task-centric approach with tools like the TAG or MTF, and frameworks by Tomlinson and Koutselini (see Chapter 2).
Extended duration with multiple opportunities to revisit and experiment with new practices	▪ The intervention lasted nearly an academic year (~6 months, 22.5 hours total), with teachers videotaping their experimentation in four lessons each.

The intervention aimed to incorporate these key features of effective PLD as fully as possible, mindful of the constraints and with a commitment to maintain a supportive and non-pressing environment for participants. Within the context of the program, these features have been largely met, striving for a balance between ambition and practicality.

3.6 Data Collection and Preparation

A rich corpus of different data sources was collected and utilized for the purposes of this study, including (a) *videotaped lessons*, *lesson plans*, and *pre- and post-lesson teacher interviews*; (b) *videotaped video-club sessions*; (c) *teacher reflection cards*; and (d) *end-of-program teacher interviews*. These are outlined below.

Videotaped lessons, lesson plans, and pre- and post-lesson teacher interviews. As a first step, permission for conducting the study was obtained from the Cyprus Educational Research and Evaluation Centre (i.e., the National Institute Review Board). To ensure compliance, written consent was obtained from both teachers and parents or guardians for the videotaping of their child’s classroom during mathematics lessons; oral consent was also obtained from the students. The written consent forms, approved by the National Commissioner for Personal Data Protection, clearly stated the research purposes, data collection processes, and participants' rights, including the voluntary nature of their involvement and the right to withdraw at any time without any consequences. This ensured that all participants were fully aware of and informed about the study, its requirements, and their rights prior to the commencement of the research, including their right to communicate any complaints they had. All collected data were stored securely using pseudonyms and participant numbers to protect their identity and were only used for the stated PLD and research purposes.

Students whose parents had not granted consent were situated outside of the cone of videotaping defined by the camera (set at the rear of the classroom) and the whiteboard at the front of the classroom. The camera adeptly followed the teacher's movements, focusing on specific groups or individual students whenever possible, and utilizing zoom-in capabilities to record student work and teacher-student interactions. This zoom-in feature was also utilized to capture any content displayed on the classroom board. Additionally, the teacher was outfitted with a lavalier microphone to capture clear audio.

Four lessons for each of the eight participating teachers were videotaped (N=32). Both pre- and post-lesson interviews (PRI and POI, respectively) with the teachers were conducted, with a total of 64 interviews either videotaped or audio recorded and transcribed verbatim. Prior to or following each lesson, teachers were asked to provide a copy of their lesson plan if they wished. Given that that doing so was optional, only eight lesson plans were provided; teachers also handed in other ancillary materials (such as student handouts) in 14 instances.

All the videotaped lessons were coded using an observation protocol developed during an earlier phase of the EDUCATE project. This protocol was designed to explore teaching quality in terms of cognitive activation, differentiation, and their interplay and comprised two components. In the first part of the protocol, coders[23] parsed each videotaped lesson into discrete mathematical tasks[24]. For each mathematical task, coders identified specific phases, including Task Launching (TL), Student Autonomous Work (SAW), and Whole-class Interactions (WCI), as applicable.

Although mathematical lessons can have different structures, they are often parsed into these three widely observed phases across various studies (e.g., Charalambous et al., 2023b; Jackson & Cobb, 20 Sullivan et al., 2006; Tomlinson et al., 2017). TL involves activating prior knowledge and/or introducing/posing a mathematical task for students to work on; SAW refers to the phase in which students independently explore/work on the mathematical task, either individually or in small groups, with targeted intervention by the teacher to scaffold or facilitate student work; and WCI involves a phase where students discuss and reason about their solutions to

[23] Training meetings were conducted to train the coders in the rubric's use. In these sessions, coders gained practical experience in coding lessons, discussed open issues and complex codes, and worked towards a consensus and shared code interpretation. The author of this dissertation was one of the trained coders.

[24] According to Stein et al. (1996) 'a mathematical task is defined as a classroom activity, the purpose of which is to focus students' attention on a particular mathematical idea'' (p. 460).

the assigned mathematical task in a plenary setting (cf. Charalambous et al., 2023b). For example, in the teaching model proposed by Sullivan and colleagues (2006), the phases of TL, SAW, and WCI correspond to the model's phases of Launch, Explore, and Summarize/Review, respectively. Tomlinson (2017) describes a lesson characterized by a "*repeated rhythm*" of TL, called "whole-class preparation," followed by SAW, termed an "opportunity for individual or small-group exploration, extension, and production," and ending with WCI, known as "whole-class review and sharing" (p. 9). Recognizing that parsing a lesson could involve multiple occurrences of TL, SAW, and WCI, each occurrence was coded separately as long as they met specific criteria. In particular. these occurrences qualified if they intended to fulfill the goals which were characteristic of a certain phase (described above), regardless of duration. An occurrence denoted a specific instance or distinct time during which students were engaged in either TL, SAW, or WCI. For example, if a lesson on parsing included "five occurrences of SAW" this indicated that there were five separate periods during the lesson during which students worked individually or in groups on a mathematical task.

For each phase occurrence, coders assigned scores ranging from 0 to 3 (see Appendix 1). These scores were used to capture the quality of teaching in relation to three criteria: (a) cognitive activation, (b) differentiation, and (c) the interplay between the two. For any given task that encompassed all three phases, coders assigned scores for up to 35 distinct codes (see Appendix 2). For the majority of the practices (23 out of 35), agreement rates exceeded 70%. Additionally, seven out of 35 codes showed agreement rates between 60% and 70%, and five out of 35 codes had agreement rates ranging from 50% to 60%. In all cases, there were extensive discussions and reconciliations to achieve consensus among the coders. For the purposes of the analysis, these reconciled scores were used. The coding process served to systematically evaluate teaching quality and quantitatively analyze teacher-student engagement with challenging tasks and differentiation in the observed lessons.

Videotaped video-club sessions (VCS). All nine video-club sessions (refer to Table 7) were videotaped using a stationary camera positioned at the front of the room. This camera recorded the participants' activities with PLD materials, their interactions, and their involvement in the discussions, encompassing their comments, suggestions, questions, ideas, concerns, reactions, silences, and so forth, throughout the video-club sessions. All verbal contributions provided by the teachers were transcribed verbatim to ensure accuracy and authenticity in capturing their insights. Instances in which the teacher educators contributed to the discourse of the video-club sessions were concisely summarized and annotated. However, key questions drawn by the teacher educators during these sessions were transcribed verbatim, due to their critical importance in the context and discussion of the video-club session.

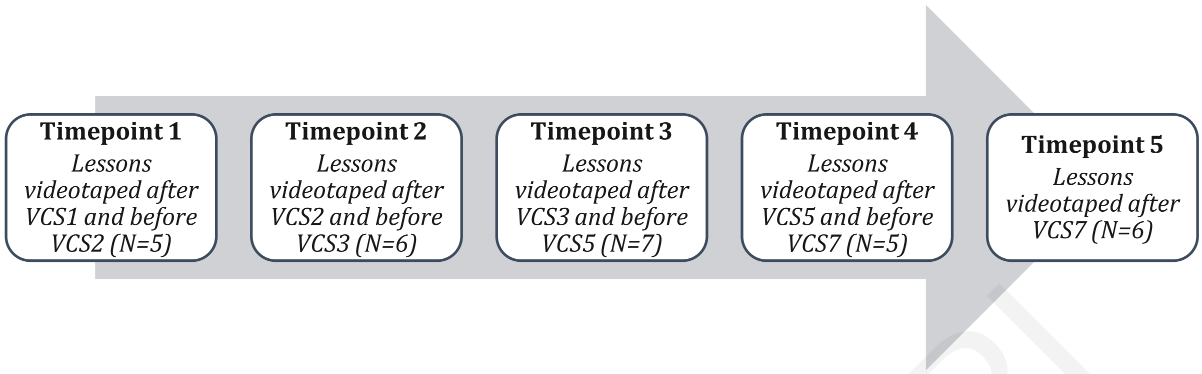
Teacher reflection cards (TRC). At the end of each session, teachers were asked to complete individual reflection cards (totaling N=72), identifying some ideas they had learned during that day's session. Additionally, they were prompted to comment on their intentions to implement any of the discussed ideas to support diverse student groups in tackling challenging tasks—and detail how, if positive. Furthermore, they were encouraged to provide suggestions for upcoming sessions, building on their perceived unresolved issues, concerns, and challenges. This feedback was used to inform the planning of the content and activities of subsequent video-club sessions.

End-of-program teacher interviews (EPI). At the conclusion of the PLD program, individual semi-structured interviews were conducted with the participants to capture their perspectives and experiences regarding the PLD program. These interviews prompted participants to reflect on their involvement in the program, highlighting its strengths, limitations, and the challenges they encountered. Participants were also encouraged to discuss the benefits they derived from the program, substantiating their points with specific examples. Furthermore, they were invited to provide suggestions for potential enhancements to the program. All end-of-program interviews (N=8) were audio-recorded and then transcribed verbatim.

Preparing the dataset for analysis. Teachers were asked to videotape four lessons; although they were encouraged to do so during certain timepoints of the PLD program, to maintain the rapport build with the study participants, they were not obliged to videotape lessons after certain video-club meetings. As a result, although all teachers videotaped the same number of lessons, they differed in when these lessons were videotaped. Therefore, to establish some comparability and enhance the statistical power of the analyses, all coded lessons, along with their interviews and teacher reflection cards, were grouped into five *timepoints* (i.e., measurements). The criterion for grouping the lessons was the focus and content of the session that preceded the videotaping of each lesson (see Figure 14).

Figure 14

Timeline of Lesson Videotaping Across Different Intervention Timepoints



As can be seen in Figure 14, each videotaped lesson was categorized under its respective timepoint, aligned with the content that teachers had been exposed in the preceding video-club session. This alignment ensured that the lessons meaningfully reflected the topics and tools discussed. Teachers were encouraged to experiment with the concepts and strategies introduced in the previous session. Lessons videotaped after VCS1, which focused on cognitive activation topics, were designated as Timepoint 1.[25] During this timepoint, teachers were asked to select a challenging mathematical task from their student textbooks and videotape its enactment. VCS2 concentrated on differentiation, with lessons during Timepoint 2 incorporating differentiation strategies into teaching practices. Subsequent sessions, VCS3 and VCS4, shifted focus towards the synergy between cognitive activation and differentiation. Teachers were motivated to apply the tools discussed in these sessions during Timepoint 3. Lessons for Timepoint 4 were videotaped after VCS5, in which teachers were provided with specific praxis tools to enhance the interplay between cognitive activation and differentiation (cf. Charalambous et al., 2022), and after VCS6, which focused on designing lessons that promote both axes. VCS7 to VCS9 equipped teachers with lesson planning tools aimed at assessing prior knowledge, anticipating student responses, implementing formative assessments, and cultivating a classroom environment that supports both axes; therefore, lessons videotaped after these sessions were clustered under Timepoint 5.

Because some teachers were videotaped twice within a single timepoint, we retained their lesson with the highest performance and excluded the other one from

[25] Because teachers videotaped their first lesson after VCS1, a baseline lesson, which could have been used as a reference point to assess teaching quality before the PLD program commencement, was unavailable.

the quantitative analysis (N=29). The decision to retain only the highest-performing lesson rather than averaging performances was based on two considerations. First, the analysis showed no notable difference (in relative ranking) in the delineation of teachers' practice when using the maximum performance as opposed to when using their mean performance.[26] Moreover, the highest-performing lesson showcases the teacher's capacity and better sketches the potential effect of the intervention. This approach also mitigates variability caused by other external factors (e.g., content of the lesson, contextual factors that might impinge on teachers' attempt to teach ambitiously). For the qualitative analysis, all the lessons of the three selected cases (N=12) were used, as explained next.

3.7 Data Analysis

Kirkpatrick's model, as outlined in Section 2.7, served as the *overarching* framework for the data analysis, which was conducted at two levels: initially at the group level and then at the level of selected teacher cases. Table 9 summarizes all the analyses employed to address the research questions of the study, both at the group level and for selected cases. In what follows, the data analysis approaches followed per each of these two levels are presented, organized per each of the four Kirkpatrick's dimensions within each level.

[26] The same analysis was conducted for mean scores, revealing largely similar patterns in teachers' practice.

Table 9

Summary of Data Sources and Analyses for Addressing Each Research Question

Research question	Data Sources[1]	Data Analysis at the Group Level[2]	Data Sources	Data Analysis for Selected Cases[3]
Teachers' Final Results				
1. How does teachers' performance look at the final timepoint?				
1.1 What teaching practices do teachers experiment with in their concluding lessons?	▪ 29 VLs	Calculating the percentages of occurrences of the lesson phase[4] in which a certain practice was noted in *teachers' lessons* at the final timepoint over the total number of occurrences of the lesson phase at that timepoint. (e.g., There were 13 occurrences of TL at the *final* timepoint, and in five of those the focal practice was observed).	-	Not applicable.
1.2 What is the teachers' performance in their concluding lessons?	▪ 29 VLs	Aggregating the scores of teachers' lessons delivered in the final timepoint at the lesson level, considering both the *mean* and *maximum* performance metrics. The mean metric reflected the group's average performance for each teaching practice and the mean overall performance, in each lesson phase per axis. The maximum metric indicated the peak performance in each teaching practice and the best overall performance, in each lesson phase per axis.	▪ 12 VLs (one lesson each)	Aggregating the scores of *each case's* final lesson at the lesson level, considering the *average* performance metrics. The mean metric reflected each teacher's average performance for each lesson phase per axis.
Teachers' Behavior				
2. How does their teaching behavior evolve over time?				
2.1 What teaching practices do teachers experiment with most frequently?	▪ 29 VLs	Calculating the percentages of occurrences of the lesson phase in which a certain practice was noted in *teachers'* lessons during a particular timepoint over the total number of occurrences of the lesson phase at that timepoint (e.g., There were 13 occurrences of TL at a *certain* timepoint, and in five of those the focal practice was observed).	▪ **12 VLs**[5] ▪ 12 PRI ▪ 12 POIs ▪ 9 VCSs ▪ 3 EPIs	Using *thematic analysis* (cf. Braun & Clarke 2006) for a multi-pass review of the data. The initial step involved familiarization with the data noting emerging patterns and drafting *analytical memos for each teacher*, highlighting key lesson episodes and interactions. A second review focused on identifying initial codes based on data excerpts that showcased changes in teaching practices, challenges, or conceptual shifts within the PLD's axes. The memos were then organized into meaningful units and initial codes. In the third step, following a top-down and a bottom-up approach (Braun & Clarke, 2006), an iterative coding process of the sections of each analytic memo (see above) pertaining to the case's teacher practice was used. A coding scheme

				was developed categorizing teaching practices within lesson phases per axis and assessed their frequency (checkmarks were used to illustrate the experimentation with each practice).
2.2 What changes do teachers introduce in their practice, during their participation in the PLD program?	▪ 29 VLs	Using the one-sample Wilcoxon signed-rank test for pairwise comparisons across the five timepoints. The aggregated scores of all teachers' videotaped lessons per timepoint, for both teachers' average and maximum performance were used. The test aimed to identify significant differences in teaching quality across the five timepoints, conducting in two separate sets of pairwise comparisons for average and maximum performance to detect any statistical changes.	▪ **12 VLs** ▪ 12 PRIs ▪ 12 POIs ▪ 9 VCS ▪ 3 EPIs	▪ *Quantitative*: Using scores from four videotaped lessons per case aggregated at the lesson level to identify any patterns of change in the teaching practice of each case, across the timepoints. This process generated nine mean scores per lesson, by averaging the scores of the practices within each lesson phase per axis. ▪ *Qualitative*: Assessing the quality of *each case's* teaching practice in the videotaped lessons by using a color scale to indicate the depth and consistency of practice integration in the coding scheme developed from coding the memos in the third step of thematic analysis (see above). The color-coded analysis highlighted patterns in teaching quality, with themes and subthemes, providing a rich narrative of each teacher's pedagogical evolution. A subset of lessons was coded by an independent coder to ensure reliability, with over 85% agreement achieved.
2.3 What challenges do teachers encounter during this process?	-	Not applicable.	▪ **12 VLs** ▪ **12 PRIs** ▪ **12 POIs** ▪ **9 VCS** ▪ 27 TRCs ▪ **3 EPIs**	Coding the observed and perceived challenges identified in the pertinent memo sections for each case. Observed challenges in *each* teacher's VLs were evident when practices were inadequately applied, affecting the implementation of each axis. Perceived challenges emerged through teachers' expressions of difficulty, stress, or the need for additional resources in the VCSs and in *each* teacher's interviews. Challenges were categorized as either "addressed or mitigated" if they were overcome or reduced during the PLD, or "unresolved" if they persisted without resolution. This classification was based on changes in practice or teachers' reflections on overcoming specific difficulties.
		Teachers' Learning		
3. How does teachers' learning evolve over time?				

3.1 How do teachers (re)conceptualize cognitive activation, differentiation, and their interplay throughout their participation in a relevant PLD program?	-	Not applicable.	▪ **12 PRIs** ▪ **12 POIs** ▪ **9 VCS** ▪ 27 TRCs ▪ **3 EPIs**	Coding the memo sections related to each case's (re)conceptualizations to uncover changes in their understanding of each axis in the third step of thematic analysis. By comparing data across timepoints, themes and subthemes were identified, tracking the evolution of teachers' conceptualizations. These were categorized into initial (Timepoint 1), evolving (intermediate timepoints), and final (Timepoint 5) conceptualizations. Conceptual changes, both explicit and implicit, were noted. Explicit changes were marked by teachers' recognition of shifts in their understanding, while implicit changes were inferred from subtle cues in language, practice shifts, and evolving reflections.
3.2 How do teachers' (re)conceptualizations and practices compare?	-	Not applicable.	▪ **12 VLs** ▪ **12 PRIs** ▪ **12 POIs** ▪ **9 VCS** ▪ 27 TRCs ▪ **3 EPIs**	Examining the alignment between each case's conceptual shifts and their teaching practices over time, *mapping changes in both areas across timepoints*, during the fourth step of thematic analysis. Instances where teachers' evolving understandings matched or diverged from their teaching practice were identified, noting both congruent and incongruent examples. This comparison included *temporal analysis* to determine whether changes in conceptualization led to, followed, or coincided with changes in practice. Summaries of these alignments or misalignments concluded each case analysis.
		Teachers' Reaction		
4. What reactions do teachers have regarding the PLD program?	▪ 8 EPIs	Employing *thematic analysis* (cf. Braun & Clarke 2006) to selected excerpts of *all teachers'* end-of-program interviews that focused on their program experiences, especially its strengths, weaknesses, and suggested improvements. The process involved reading and coding the transcripts, developing and refining themes, leading to a comprehensive report *for the whole group* in Chapter 4.	▪ 3 EPIs	Employing *thematic analysis* (cf. Braun & Clarke 2006) to selected excerpts of *each teacher's* end-of-program interviews that focused on her program experiences, especially its strengths, weaknesses, and suggested improvements. The process involved reading and coding the transcripts, developing and refining the themes, leading to a comprehensive report *for each case* separately in Chapters 5, 6, & 7.

Notes.
1. VL: videotaped lessons; PRI: pre-lesson interview; POI: post-lesson interview; VCS: video-club session; TRC: teacher reflection card; EPI: end-of-program interviews.
2. Each analysis was conducted *once* for the whole group.
3. Each analysis was conducted *thrice*, for each case *separately*.
4. An occurrence denoted a specific instance or distinct time during which students are engaged in either TL, SAW, or WCI (see Section 3.6).
5. The main data sources for each data analysis are highlighted in bold, while the others served as supplementary data sources for that analysis.

3.7.1 Data Analysis at the Group Level

Results. The analysis initially focused on the (final) *Results* (Level 4 in Kirkpatrick's model) for the *teachers as a group*. Using descriptive statistics, teachers' collective teaching performance and the practices with which they most frequently experimented in their *final* lessons were quantitatively assessed. The analysis is detailed next.

Exploring teachers' collective experimentation in their final lessons. The analysis involved calculating the percentages of the use of teaching practices[27] used by the participating teachers in their final videotaped lessons. In particular, it entailed calculating the percentages of occurrences of the lesson phase in which a certain practice was noted in teachers' lessons during the last timepoint over the total number of occurrences of the lesson phase at that timepoint (e.g., there were 13 occurrences of student autonomous work at Timepoint 5 and in three of those the focal practice was observed). This analysis offered an overview of the *frequency* of the practice's occurrence but not the *quality* of the experimentation across the entire group (see Chapter 4).

Exploring teachers' collective teaching performance in their final lessons. To explore the quality of teachers' collective experimentation with cognitive activation and/or differentiation practices, the aggregated scores of their final videotaped lessons were used. The data were aggregated at the lesson level using both teachers' *mean* performance and *maximum* performance. The mean performance metric encapsulated the *average* level of (a) the teacher-group's performance in implementing the focal teaching practices in each lesson phase (i.e., task launching, student autonomous work, and whole-class interactions) for each axis (i.e., cognitive activation, differentiation, and their interplay) and (b) their overall performance across these phases. The maximum performance metric, on the other hand, signified the *highest* level of (a) the teacher-group's performance in employing each teaching practice in each lesson phase per axis and (b) their overall performance.

Behavior. To examine the changes in the *Behavior* (Level 3 in Kirkpatrick's model) of teachers as a *collective* entity, the study focused on the frequency and the quality of their experimentation in their videotaped lessons *across* the five timepoints using both descriptive and inferential statistics. Both analyses are detailed below.

[27] The teaching practices examined were included in the protocol used for coding the lessons.

Exploring collective changes in teachers' experimentation. The previously described method for calculating percentages was used for generating the *frequency* of the use of the teaching practices *for each of the five timepoints*. In particular, the percentages of occurrences of the lesson phase in which a certain practice was noted in teachers' lessons at *each* timepoint over the total occurrences of the lesson phase at that timepoint were calculated (e.g., there were 13 occurrences of student autonomous work at Timepoint 3 and in three of those the focal practice was observed). This approach enabled the examination of patterns or deviations in teachers' experimentation *across* these intervals.

Exploring collective changes introduced in teachers' practice. To investigate any possible changes in the teaching *quality* of the PLD group, the one-sample Wilcoxon signed-rank test was employed on the aggregated scores of all teachers' videotaped lessons per timepoint, for both teachers' *mean* and *maximum* performance. This non-parametric statistical approach was conducted as a paired difference test across the five repeated measurements (i.e., the timepoints) within a single sample. Using the test over the paired *t*-test was more appropriate due to the study's small sample size, which raised concerns about the normality of the data distribution (cf. Siegel & Castellan, 1988). Eventually, two sets of pairwise comparisons—i.e., one set for the aggregated mean scores and one set for the aggregated maximum scores, split by timepoint—were performed to thoroughly examine potential statistically significant differences in teaching quality across the five timepoints.

Learning. The assessment of teachers' *Learning* as a collective (Level 2 in Kirkpatrick's model) presented more complexity and required a qualitative approach. This *level* was examined only for the selected cases (not for the group-as-a-whole) because this investigation required a more targeted analysis, focusing on selected cases, to ensure a *meaningful* exploration of teacher learning (see more in Section 3.7.2).

Reaction. At Level 1 of Kirkpatrick's model, *Reaction*, a qualitative analysis of all teachers' end-of-program interviews was undertaken. A couple of interview questions were designed to prompt teachers to reflect on their experiences within the program, with a particular emphasis on identifying its strengths and weaknesses. Specifically, teachers were encouraged to articulate the benefits they perceived they accrued from their participation, providing concrete examples to support their insights, and suggest potential improvements to the program. The analysis of teachers' reactions to the PLD program was derived from isolating the interview excerpts that addressed the aforementioned topics. Thematic analysis, as outlined by Braun and Clarke (2006), was employed through a series of steps: carefully and repeatedly reading the interview transcripts; generating initial data-driven and theory-driven codes; grouping different codes that fit together into broader themes; iteratively refining these themes and their subthemes; assigning definitive names to the themes and subthemes; and finally compiling the findings into a report for Chapter 4 (see Table 10).

3.7.2 Data Analysis for Selected Cases

Following Kirkpatrick's organization, the data were approached through both quantitative and qualitative lenses to offer insights not only into the results and reactions to the program but also to delve into the learning trajectories and behavioral evolutions of selected teacher cases. This approach enabled a deeper investigation of the research questions, facilitating the detection of emergent patterns of change, where they existed.

Results. The scores of the final videotaped lesson of each of the selected cases were aggregated at the lesson level (N=1 lesson each). For each case, this involved averaging the scores of the practices within each lesson phase per axis. This analysis produced a set of nine mean scores per final lesson (three phases X three axes). This procedure was replicated three times, once for each of the three selected cases (see Chapters 5, 6, and 7).

Learning and Behavior. For the selected cases, the levels of *Learning and Behavior* were studied *concurrently* through mixed-method processes that operated complementarily; thereby these analyses are presented together. The quantitative component documented any quantifiable changes in the teaching practices in the videotaped lessons of the selected cases (*Behavior*). The teachers' (re)conceptualizations of cognitive activation, differentiation, and their interplay (*Learning*); any qualitative changes in their teaching quality across their videotaped lessons (*Behavior*); as well as the challenges encountered throughout their participation in the PLD (*both levels*), were qualitatively explored. Finally, the study explored the alignment of the evolution of the cases' conceptualizations and teaching practice (*both levels*). The mixed-methods approach adopted is further detailed below.

Exploring quantitative changes in the teaching practice of the selected cases. The scores of the videotaped lessons of each selected case were aggregated at lesson level

(N=4 lessons each). The mean scores for each lesson were obtained by averaging the scores of the practices within each lesson phase per axis. This analysis yielded a set of nine mean scores per lesson that reflected the selected case's performance during that particular lesson. This process was carried out separately for each of the three cases, a total of three times. These scores lent themselves to investigating changes in each teacher's teaching practices over time, underscoring any notable trends within the context of the PLD program.

Exploring qualitative changes in the conceptualizations and teaching practice, as well as challenges faced by the selected cases during the PLD program. For this investigation, all the available data of each of the three selected cases underwent *thematic analysis* involving multiple passes (cf. Riger & Sigurvinsdottir, 2016). This analysis was employed three times, once for the dataset of *each teacher*. The first pass involved watching the videotaped lessons of each teacher (N=4 each) and reading through the transcripts of their pre- and post-lesson interviews (N=8 each), end-of-program interviews (N=1 each), reflection cards (N=9 each), as well as the video-club sessions (N=9). The primary goal of this read-through phase was to get familiar with the data. During the data review, notes were taken on emerging ideas, meanings, and patterns, and a detailed analytical memo was developed for each teacher. These memos included rich descriptions of each teacher's videotaped lesson, along with characteristic excerpts and notable interactions.

Then, a second pass entailed a focused review of the data to identify an initial list of codes. The focus was on identifying data excerpts that reflected each case's (re)conceptualizations (from the video-club sessions and *each* teacher's interviews, changes in teaching practice (from *each* teacher's videotaped lesson), or perceived challenges faced during the PLD (from the video-club sessions and *each* teacher's interviews), with respect to the three axes of the program—cognitive activation, differentiation, and their interplay. Pertinent excerpts from the teachers' interviews and reflection cards, as well as from the videotaped video-clubs were incorporated in each analytic memo. The analysis was also deepened by a second review of the videotaped lessons in order to identify excerpts that related to each teacher's *observed* challenges (from each teacher's videotaped lessons). Gradually, each memo was organized into meaningful units, forming a set of initial codes. This organization was guided by existing literature, employing a top-down or theory-driven approach, while also remaining open to new, emerging ideas through a bottom-up or data-driven approach (cf. Braun & Clarke, 2006).

The third pass of the data adopted a three-layered approach, focusing on coding the memos and searching for themes to (a) identify changes in teaching practices; (b) highlight shifts in teachers' conceptualizations; and (c) identify and cluster the challenges faced. This coding is detailed below.

***Coding memos to identify changes in teachers' practices*.** The sections of the memos that pertained to (changes in) teachers' practice were coded following a top-down and bottom-up approach (Braun & Clarke, 2006). This approach integrated predetermined codes, derived from the coding protocol used for coding the videotaped lessons (top-down), while also accommodating new codes that emerged from the data from all three cases (bottom-up). This dynamic and iterative coding process involved multiple rounds of review to refine the codes, clarify their relationships, and identify overlaps among them, leading to the development of a coding scheme to ground the findings (see Tables 12, 13, and 14 in Chapters 5, 6, and 7, respectively).

This coding scheme was instrumental in systematically organizing the findings. It was structured around three key axes: cognitive activation, differentiation, and their interplay, reflecting the core emphases of the PLD program. For each axis, the teaching practices observed in each phase of the lesson—task launching, student autonomous work, and whole-class interactions—were incorporated (see the first column of Tables 12, 13, and 14 in Chapters 5, 6, and 7, respectively). The scheme incorporated both practices that *supported* or *impeded* each dimension.

The next step aimed at determining both the *frequency* and the *quality* of the implementation for each identified practice in the videotaped lessons (see the last four columns of Tables 12, 13, and 14, in Chapters 5, 6, and 7, respectively). A checkmark was employed to indicate that the teacher had experimented to some extent with a practice during their lesson, denoting its frequency. Then, a color scale[28] was utilized to illustrate aspects of the quality of practice implementation. Practices that were systematically integrated into the lesson, demonstrating depth rather than sporadic use, were highlighted in a bright green color, signifying a high level of performance. Practices that appeared fleetingly and the quality of their implementation was in pro-forma ways were depicted in a lighter shade of green. In contrast, the red shade

[28] The color shading signified a more qualitative or descriptive sense to indicate the richness, depth, and quality of the implementation of different teaching practices.

indicated that the cognitive level had decreased when implementing the practice. Practices that hindered the implementation of each axis were grey colored. To ensure interrater reliability, one third of the lessons (N=4 out of 12) underwent coding by an independent coder who was trained on using the coding scheme. The interrater reliability indicated a high level of agreement (over 85%) between the two trained coders (including the author). Any discrepancies in interpretation were thoroughly discussed and debated until new or refined codes and/or shading were established (which were taken into consideration in revising the coding for all three teachers, if necessary).

The color-coded representation of the use of teaching practices across the four lessons provided insightful *themes* about patterns in the teaching quality of each case. These themes showcased either progressive trends or consistent patterns, delineating the fundamental features of each teacher's teaching evolution. Within these overarching themes, subthemes that captured crucial elements of the coding scheme were identified and detailed. Each theme and subtheme are named and described in Chapters 5, 6, and 7 (for each Pina, Kate, and Michelle, respectively), enriched with narratives, detailed descriptions, and illustrative examples drawn from the videotaped lessons, offering a vivid portrayal of each teacher's teaching dynamics at play.

The coding scheme functioned as a heuristic tool, capturing the nuances of teaching practices in each teacher's lessons and enabling a detailed qualitative exploration of each teacher's teaching, revealing subtleties not necessarily evident in the quantitative analysis.

Coding memos to highlight shifts in teachers' conceptualizations. The second layer of coding focused on analyzing the memo sections containing data from teachers' pre- and post-lesson interviews, end-of-program interviews, and reflection cards, along with footage from video-club sessions. This analysis specifically targeted at highlighting patterns in teachers' conceptualizations of cognitive activation, differentiation, and their interplay. Constant comparisons were made across the data of consecutive timepoints, scrutinizing them for similarities, differences, and recurring ideas or concepts.

Through this process, themes and their corresponding subthemes were identified, delineating the patterns of change in the conceptualizations of each teacher within every axis. These themes were then systematically categorized into initial, evolving, and final conceptualizations for each axis, providing a trajectory of the teachers' understanding over time. Initial conceptualizations derived from data collected at Timepoint 1, whereas final conceptualizations were based on data from Timepoint 5. Any data collected between these timepoints pertained to the evolving conceptualizations of the teachers. For each case, the output of this analysis presents a detailed portrayal of their initial, evolving, and final conceptualizations, including the associated themes and supporting data excerpts that substantiate these themes.

To identify a conceptual change,[29] there needed to be a noticeable shift in teachers' understanding or perspective regarding a particular pedagogical concept or idea related to the axes of cognitive activation, differentiation, or their interplay. Recognizing that teachers' awareness of conceptual change could vary greatly, analysis was open to either *explicit* or *implicit* changes. Explicit conceptual changes were identified when teachers described 'aha moments' they experienced, discovering the limitations of their current understandings, and explaining their new conceptions. Implicit conceptual changes involved subtle shifts in understanding that the teacher might have not been fully aware of; these were recognized by paying close attention to the language used, the examples provided, the shift in focus, the incorporation of new practices in their teaching, and changes in the teacher's reflections over time.

Exploring how the changes in conceptualizations and teaching practice of the selected cases compare. The fourth pass of the data involved exploring the potential congruence or incongruence between teachers' (re)conceptualizations and the changes in their observed classroom practices over time. The quantitative and qualitative findings related to teachers' changes in teaching practice and conceptualization were systematically mapped against across the timepoints. This involved looking for instances in which shifts in learning (conceptualizations) were mirrored by corresponding changes in classroom behavior (teaching practices). For each case, instances in which changes in conceptualizations and teaching practices were aligned or misaligned were noted. For example, alignment could be indicated by an emphasis on challenging less advanced students in both the conceptualization and teaching practice of a teacher at the same timepoint. In contrast, misalignment could be observed if, for example, a teacher articulated a shift towards differentiated

[29] Conceptual change denotes how learners reorganize, revise, or replace their existing understanding of a concept or idea with new, more fruitful, intelligible, and plausible insights (cf. Gregoire, 2003). This change can be moderate or substantial. Revising a learner's conceptual schema often requires experiencing cognitive conflict or dissatisfaction with their existing beliefs or knowledge, prompting a re-evaluation of their conceptions (cf. Özdemir & Clark, 2007).

strategies but continued to employ predominantly uniform approaches. The comparison considered the timing of changes (temporal analysis), recognizing that shifts in conceptualization might precede, follow, or occur simultaneously with changes in teaching practice. These instances of alignment or misalignment between their conceptualizations and actual teaching practices were summarized at the end of each case.

Coding memos to identify and cluster teachers' challenges. The third layer of memo coding was dedicated to identifying and coding challenges[30] faced by the teachers. The videotaped lessons helped identify the observed challenges faced by teachers during teaching, while the remaining data sources revealed their perceived challenges. For instance, moments in the lessons in which the teacher *systematically* applied practices in a manner that needed improvement or hindered the implementation of one of the three axes suggested observed challenges. The coding scheme and its color shading proved to be supportive in further revealing observed challenges. In the contrary, a perceived challenge could manifest through various indications or expressions from teachers during their interviews or the video-club discussions, signalling difficulties, obstacles, or concerns in their teaching practice. For example, teachers could openly admit finding certain teaching aspects challenging, or express frustration or stress about specific facets of their work. Additionally, an increased tendency to seek out resources or support could indicate areas in which teachers felt less confident or faced challenges. In addition, recounting specific challenging incidents or moments from videotaped lessons and attention to non-verbal signals such as tone or hesitance could provide further insights into perceived challenges.

These challenges were then classified into "challenges addressed or mitigated during the PLD," which were, to some degree, alleviated during the program, and "unresolved challenges," which persisted throughout the program without being adequately addressed. For example, if a teacher described how adopting a specific practice or idea helped her navigate through a challenge she was facing, it would imply that the obstacle was successfully overcome or the challenge was reduced, categorizing it as an addressed challenge. On the other hand, there might be situations where a teacher consistently mentioned a persistent challenge that had not been mitigated yet,

[30] The term "challenges" is used to denote a range of observed or self-reported problems that teachers face while planning and enacting lessons (cf. Lampert, 1992, 2001).

even in the end-of program interview, placing it in the category of unresolved challenges. Similarly, if a teacher showed improvement in implementing specific practices or reduced practices that obstructed progress in any axis, this was seen as successfully overcoming the challenge. Conversely, if these issues persisted until the final lesson, it indicated that the challenge remained unresolved.

Reaction. The thematic analysis followed for the whole group was also employed to discern the individual cases' reactions to the program. This involved a focused examination of specific parts of the teachers' end-of-program interviews, particularly their reflections on the program's benefits, its strengths and limitations, and their recommendations for enhancing the PLD (see Section 3.7.1).

3.8 Researcher's Autobiographical Note

In the conduct of this study, the researcher brought forth a rich tapestry of personal interests, beliefs, and experiential knowledge, each layer contributing to the rich backdrop against which this research unfolded. These personal dimensions were invaluable in shaping her perspective and approach since they enabled the researcher to engage deeply with the data, bringing insights and empathetic understanding to the analysis of teachers' practices and challenges within the PLD program. At the same time, they necessitated vigilant awareness to mitigate their influence during key phases of data collection and analysis, as well as in presenting and interpreting the study findings. Below, factors that could potentially introduce researcher bias, followed by the measures undertaken to alleviate these influences are outlined.

The researcher had a background in advanced mathematics during her secondary education, combined with her specialization in mathematics education in her bachelor's degree of educational science. This blend of content knowledge and pedagogical expertise enriched her understanding of the subject matter and teaching methodologies, equipping her with a better understanding of the challenges and opportunities in teaching mathematics.

As a senior undergraduate student during her field practicum, the researcher participated in two extensive video-clubs. She had the opportunity to videotape her lessons, engage in video-club discussions with her peers, and reflect on her practice. This experience enabled her to better understand how teachers experience video-clubs. This firsthand experience fostered a strong advocacy for video-clubs within the researcher, as she witnessed their potential in offering a reflective and collaborative platform for teachers to critically analyze and enhance their teaching. However, she had her reservations about whether they yield positive results in the same way for all participating teachers since each teacher's participation is driven by unique motives and participants may have diverse initial knowledge and experience regarding the discussed topics.

As the Assistant Coordinator of the EDUCATE project, the researcher played a multifaceted role, deeply engaging in every phase of the program, including facilitating the video-club sessions, conducting the end-of-program interviews, preparing reports and intellectual outputs, and finalizing the modules and platform. Such extensive engagement inherently positioned the researcher in a unique stance, intertwining professional responsibilities with the analytical lens through which data was selected and examined.

The researcher's immersion into all the phases of the project afforded her a comprehensive understanding of the program's intricacies and objectives. However, this very immersion necessitated a conscientious reflection on the potential for personal biases to color the research process. As a facilitator, the researcher, being less experienced than the teachers in the video-club group, participated in discussions as needed without excessively challenging the teachers to avoid appearing disrespectful. The dual nature of her role—both as a facilitator and an analyst—could have subtly influenced the selection, framing, and interpretation of data, underscoring the importance of reflexivity in her methodological approach. Her participation in the program as an Assistant Coordinator positively influenced her stance towards the program, as she was inherently motivated to witness enhanced results and contribute to its success. This involvement also led her to view the PLD program as ambitious, recognizing that, despite its comprehensive goals, not all objectives might be fully realized within its scope.

Moreover, her previous research on cognitive activation in mathematics and her experience as a teacher for the Deaf made her an advocate for cognitive activation and differentiation. Recognizing the demands of ambitious teaching, she emphasizes the need for systematic support and PLD for teachers to realize the vision of ambitious teaching effectively. This predisposition could lead to affirming both axes in the collected data, possibly overlooking contrary evidence. Additionally, the focus might shift towards the complexities of ambitious teaching, potentially underplaying its benefits or strategies for addressing its challenges.

To mitigate potential biases, the researcher conscientiously employed a suite of rigorous methodological safeguards. These included engaging in constructive peer debriefing sessions and maintaining a reflexive journal (see Section 3.9). Peer debriefing allowed colleagues to critique and challenge assumptions, offering external viewpoints that could uncover overlooked biases and suggest alternative data interpretations, enhancing accountability and reducing the influence of preconceptions on the findings. Also, the reflexive journal served as a critical tool for introspection, allowing the researcher to carefully examine and question her own influence throughout the research process.

3.9 Strengthening the Trustworthiness of the Study

To enhance the study's rigor, several measures were taken considering Mertens's (2015) criteria for determining quality in mixed methods research. For the quantitative data, the following aspects were considered: validity, reliability, and objectivity. For the qualitative data, the following aspects of trustworthiness were taken into consideration: credibility, transferability, and confirmability.

With respect to quantitative criteria, to boost the study's validity this chapter meticulously documents all methodological decisions and transparently acknowledges limitations (see Section 3.10). This clear research path facilitates external audits and also reinforces the study's reliability. Moreover, teacher participants were videotaped multiple times during the program—crucial for associating any observed changes in teaching with the intervention, thereby strengthening internal validity and reliability. To further solidify the reliability and objectivity of the quantitative component, each videotaped lesson was meticulously coded by two independent coders, thoroughly trained in applying the coding protocol. Inter-rater agreement statistics reflect the reliability of the coding process. Finally, the use of reconciled scores in data analysis, derived from a consensus between the coders, enhanced the reliability of the research findings.

Regarding the qualitative criteria, credibility was enhanced through prolonged and persistent observation over six months, facilitating an in-depth exploration of teachers' conceptualizations, teaching practices, and challenges, and the interactive dynamics within video-club sessions. The analysis of teaching practices employed both top-down and bottom-up approaches (Braun & Clarke, 2006), ensuring a balanced integration of established concepts and emergent themes, thus bolstering the findings' credibility. Also, the coders' training in the qualitative coding scheme supported a credible analysis by promoting consistent and accurate application of codes. The inclusion of the developed coding schemes in Chapter 5 enhances confirmability. Closely related to this measure was the key methodological choice of selecting heterogeneous cases for analysis (see Section 3.5); it not only showcased diverse educational contexts but also aided in evaluating the transferability of the findings. Enhancing transferability, the study presents (often negative) data points contradicting emerging themes or suggesting alternative interpretations (see Chapters 5, 6, and 7). Acknowledging the inherently subjective nature of qualitative analysis and foregrounding the plurality of interpretations (Patton, 2015), the study sought to offer a transparent account of the various possible readings of the data, inviting a critical dialogue around the findings. Another measure taken by the researcher to reduce bias was to present thick descriptions and multiple interpretations of the findings, further enabling the transferability of results. Finally, the author's autobiographical note and reflexive journaling (see Section 3.8) were instrumental in increasing the study's trustworthiness, serving to introspectively assess and address potential biases, thereby reinforcing the credibility and confirmability of the research.

Finally, an effort was made to strengthen data collection and analysis triangulation. Firstly, multiple data sources were collected to ensure a holistic view of the research topic and minimize single-source bias. Secondly, the integration of mixed methods, combining quantitative techniques such as the Wilcoxon test, mean scores, and percentages with qualitative case studies, significantly enhanced the study's depth and validity. Quantitative analysis offered statistical insights, whereas qualitative inquiry provided essential context, enriching the study's overall credibility. This mixed-methods approach enabled cross-validation of findings from diverse data forms, mitigating method-specific biases and bolstering the study's confirmability. Third, weekly peer debriefings with fellow PhD candidates and the supervisor offered a triangulated review of the methodology and findings, uncovering blind spots and providing new insights, thereby enhancing the study's credibility. This also allowed to establish the fit between the participants' views and researcher's representation of them. These discussions promoted self-reflection on personal biases, improving the study's credibility, dependability, and confirmability.

3.10 Methodological Limitations

The findings of this study have to be seen in light of some methodological limitations. The first relates to the convenient sample recruited. Considering that all teachers were volunteers, it is important to note that the findings might reflect a higher level of motivation and engagement compared to a more heterogeneous group, at least in terms of motivation. Although the insights obtained are informative, the voluntary participation of teachers in the PLD program may skew the results toward more favorable outcomes.

Secondly, despite the attempts made to have teacher participants be videotaped at regular intervals, this was not possible due to various personal and contextual factors; equally important, none of the teachers had videotaped lessons before the commencement of the PLD program, and not all of them at the culmination of each video-club session, thus making it difficult to trace changes in teachers' practice from session to session. Although the intention was for teachers to deliver lessons both before and after the intervention, this requirement was reevaluated to mitigate the risk of increased dropout rates. The primary concern was to prevent any undue pressure because of early-stage videotaping. Adjusting the research design demonstrated ethical sensitivity, enhancing the study's integrity. Despite the absence of a baseline lesson for comparison, the study's longitudinal design allowed for collecting data at several intervals during and post-intervention. This approach, alongside temporal analysis, revealed trends and shifts over time, complemented by participant reflections on pre-intervention practices. Also, to increase comparability and meaningful interpretations, the videotaped lessons were grouped into five timepoints (see Figure 14) based on the focus of each video-club session that preceded the lesson.

Thirdly, a closely related limitation concerns the lack of consistency in accompanying each lesson with a lesson plan or lesson materials. We refrained from pressing teachers to prepare and submit lesson plans for each lesson to avoid imposing unnecessary demands on teachers that might have deterred their continued participation in the program.

Fourth, due to the small sample size participating in the EDUCATE program (fewer than 100 teachers across the four countries), it was not feasible to conduct confirmatory factor analysis to ascertain the construct validity of the observation protocol.

Fifth, the collected self-reported data may have contained several sources of participants' bias that the researcher should have been alerted to; even so, this kind of data allowed participants to uncover and express their own ideas. To mitigate this limitation, the self-reported data were complemented with observation data over a relatively extended time period, which allowed to reveal descriptions and recurring patterns of behaviors. The mixed research design employed herein helped compensate for the weaknesses of either a strictly qualitative or a strictly quantitative approach.

A sixth limitation is related to the potential researcher bias who also acted as a video-club facilitator. The researcher's autobiographical note was a way of remedying this limitation, reducing any possible biases the author could hold in analyzing and interpreting the data (see Section 3.8).

Finally, investigating the sustainability of the changes in teachers' conceptualizations and teaching practice could provide further insights into the long-term effect of the project after the intervention was over. Yet, doing so was beyond the scope of the current study, suggesting a topic for future research.

CHAPTER 4. GROUP-LEVEL FINDINGS

This chapter evaluates the effectiveness of a PLD program for a group of teachers through the lens of Kirkpatrick's model to focus on their *Reactions*, *Results*, and *Behavior*. Thematic analysis of end-of-program interviews revealed teachers' positive reception of the program, with a few negative aspects and suggestions for improvement. Results from their concluding lessons showed mixed outcomes. Particularly, percentages showed a strong commitment to practices of cognitive activation, albeit less experimentation with practices promoting differentiation and the interplay of both axes. Teachers' performance in their concluding lessons indicated moderate to high scores in enhancing cognitive activation and its interplay with differentiation, though variability in differentiation. Behavioral changes throughout the program were also mixed. Analysis of teachers' lessons across the program duration revealed that while cognitive activation practices were consistently implemented, the experimentation with practices that promote differentiation and their interplay revealed fluctuations. Wilcoxon signed-rank tests showed significant improvements in practices that enhance the interplay of both axes and less in each axis alone, observed predominantly around the program's midpoint, mostly during task launching and student autonomous work. However, not all improvements were maintained until the program's end. The findings from analyzing reactions, results, and behavior of the teachers as a group provided a holistic and complementary view of the program's effectiveness while identifying unresolved issues.

4.1 Teachers' Reaction to the PLD Program

Table 10 delineates teachers' feedback on the PLD program—corresponding to the Reaction level of Kirkpatrick's model. The findings are organized into themes and subthemes that emerged from the thematic analysis of their end-of-program interviews. Specifically, their reactions fall into three overarching categories: (A) positive aspects, (B) negative aspects, and (C) suggestions for program improvement. Within each category, the analysis revealed six themes reflecting positive perceptions of the program, one theme identifying negative sides, and one theme dedicated to proposing refinements. The multitude of positive aspects illustrates that teachers overwhelmingly perceived the PLD program positively. As described next, they highlighted several key aspects of the video-clubs' operation and their impacts on their teaching behavior and learning.

Table 10

Teachers' Evaluation of the PLD Program

Theme	Subthemes
A. Positive Aspects of the PLD Program	
1. Developing a strong sense of learning community in which teachers support each other through interaction and collaboration	**a.** Enhancing the community feeling and fostering closer relationships and personalized interactions because of the small group size
	b. Sharing common concerns and challenges openly, and joint and reflective problem-solving
	c. Collaborating and discussing lesson content, common concerns, and challenges with teachers teaching at the same grade level
	d. Working collectively on planning a lesson that promotes both cognitive activation and differentiation
	e. Viewing videotaped lessons allowing them to observe and learn from each other
2. Becoming more aware of their teaching and getting ideas from others' teaching due to the video-club component of the program	**a.** Having multiple opportunities to experiment with various practices in their videotaped lessons
	b. Watching and discussing their teaching to become more conscious and more capable of reflecting upon and analyzing their practice and mistakes
	c. Watching other teachers' teaching during the sharing part of the video-clubs to get ideas and help them relate their own work to that of other teachers
	d. Observing teaching methods and ideas from colleagues, through video clips or discussions to help them see and develop diverse perspectives
3. Becoming (more) familiar with the ideas of cognitive activation and differentiation, putting a name on certain teaching aspects used before rather unconsciously, and legitimizing their use in teaching	**a.** Getting to know (better) the ideas of cognitive activation and/or differentiation
	b. Seeing the practical manifestations of the ideas they were learning more theoretically in other PLD seminars
	c. Naming certain teaching moves discussed in the program they had already been using unconsciously
	d. Exploring and discussing ideas and practices that were previously overlooked or undervalued in their teaching
	e. Realizing the legitimacy of their work
4. Combining theory and practice and enriching participants' teaching toolkit with specific ideas and teaching practices	**a.** Designing and using enablers and extenders
	b. Learning how to activate the more advanced students who are often neglected during teaching
	c. Learning how to pose appropriate and diverse questions in different phases of a lesson
	d. Learning to analyze the tasks more carefully during planning and rank them according to their cognitive challenge

5. Feeling ownership, not being overwhelmed with ideas, and feeling that knowledge was co-constructed step by step and in ways that responded to their needs	**a.** Feeling that the experience they were bringing to the sessions was well-respected
	b. Co-constructing ideas with the session facilitators, giving them a sense of ownership
	c. Identifying their strengths and weaknesses, and customizing the program components flexibly to respond to their ideas, experiences, and specific needs rather than using preset presentations and pre-imposed ideas
	d. Introducing new ideas incrementally and implementing the discussed practices in small, manageable steps, to maintain a pace that enables participants to gradually build on shared ideas and develop cumulative knowledge
6. Identifying general organizational aspects of the PLD program, both structural and procedural, underpinning its effectiveness	**a.** Establishing clear learning objectives and structure for each session to guide learning
	b. Ensuring that the program follows a cyclical process of planning, acting, observing, and reflecting
	c. Providing valuable materials and resources and summarizing the key ideas of each session to support learning and reflection
	d. Engaging in thought-provoking peer discussions
	e. Designing the program with an extended duration and spacing out sessions to facilitate the absorption and application of new knowledge
	f. Providing (diverse) feedback from (multiple) teacher educators on their lesson plans and videotaped lessons, offering a range of perspectives and supporting them overcome challenges faced
B. Negative Aspects of the PLD Program	
1. Facing emotional and physical well-being challenges	**a.** Addressing or mitigating the initial stress and nervousness of videotaping their lessons
	b. Counteracting physical fatigue from continuous after-school PLD sessions
C. Suggestions for Improvement	
1. Optimizing PLD program organization and scheduling	**a.** Having more time during the video-club component of the session to watch more or longer lesson videoclips
	b. Developing a more defined schedule for videotaping lessons to enhance program structure and teacher learning
	c. Providing fewer PLD sessions could have been sufficient, without of course compromising content quality
	d. Accommodating the need for individualized feedback on their videotaped lessons more systematically within the constraints of participants' time
	e. Providing more examples of both effective and ineffective practices.
	f. Focusing on and discussing content that would apply to lower elementary grades
	g. Reinforcing the community of practice by providing teachers with other forums to communicate beyond their regular sessions during the program

A cornerstone of the program was its ability to cultivate a *learning community*, in which participants engaged in mutual support and shared reflections (Theme A.1):

> The warm atmosphere [of the program] was very important, allowing us to freely express and exchange ideas, discuss challenges and concerns, take initiative, and receive the necessary support whenever we felt the need for it. [We developed] a learning community where mistakes were permissible, and one could easily admit misunderstandings, freely ask for clarifications, and readily accept and learn from failures. This group was more like a family than just a team. (EPI-T1, lines 430-443)

Georgia described the group's dynamic as familial, emphasizing the importance of a safe, open, and supportive space for discussing mistakes, seeking clarification, and fostering a strong sense of belonging. This sense of community was further enriched by its small group size that facilitated intimate discussions (EPI-T6, line 419), allowing teachers to observe and discuss each other's work (EPI-T7, lines 423-430), delve into common challenges (EPI-T1, line 76-78; 451-458), collaboratively search for ideas and solutions (EPI-T4, lines 395-396), and collectively plan lessons that promoted both cognitive activation and differentiation (EPI-T4, line 434-435; EPI-T5, lines 413-416).

The second theme suggested that the *video-club component* of the PLD sessions emerged as a vital tool in elevating teachers' self-awareness and reflection upon their teaching (Theme A.2):

> Watching others' teaching was key because you realize that this ideal classroom and this perfect teaching do not exist; you also get to realize some things that you do that are not appropriate and which you wouldn't have been able to notice if it hadn't been for the video. (EPI-T5, lines 440-444)

Through observing and discussing their own and their peers' video-clips, teachers became more adept at analyzing their teaching methods and mistakes. This process facilitated a valuable exchange of ideas and practices: "[I enjoyed] going the next day or week [after videotaping my lesson] to discuss what worked, what didn't go well, what challenged me, or what I could do differently [in my lesson] while interacting with other teachers who see it from their perspective, experience, and knowledge." (EPI-T9, lines 1008-1014). Such interactions not only helped teachers to relate their work to that of their colleagues but also exposed them to a variety of teaching practices and perspectives (EPI-T1, lines 323-329; EPI-T7, lines 432-435). This exposure was instrumental in broadening their teaching viewpoints and *consciously* implementing certain practices (EPI-T1, lines 48-51; EPI-T9, lines 682-687). Stella's viewpoint encompasses all these ideas:

> Although in the beginning, it was very inconvenient to watch yourself teaching, you gain a lot [by doing so]; it highlights strengths and areas for improvement. Viewing your teaching from an external perspective allows for a deeper understanding and reflection on your practice. It also allows you to adopt the perspective of students, to understand how they perceive teaching. (EPI-T7, lines 48-56)

Stella clearly articulated the profound benefits that emerged from the video-club component, despite the initial struggles she also faced, which are discussed later. Her point about adopting a student's perspective was particularly compelling, helping teachers better understand the challenges and understanding of their students.

Another positive aspect of the program was enhanced *familiarity with the concepts of cognitive activation and differentiation* (Theme A.3):

> I realized that differentiation can be done in an easier and less cumbersome way than what I had originally thought. My experience was limited to theories from my participation in one-shot seminars. I had never actually worked on or learned how it can be applied in a real classroom context. Same thing with work on challenging tasks. (EPI-T1, lines 33-40)

Teachers were introduced to and developed a deeper understanding of cognitive activation and differentiation (EPI-T6, line 293; EPI-T7 line 34), attaching formal terminology to practices some were already employing intuitively ("It helped me realize what I am doing and why I am doing it. Many things that I had been doing unconsciously, based on my experience, are now named, and supported by the literature." EPI-T1, lines 48-51). Discussing the effectiveness of multiple teaching practices was pivotal in legitimizing teaching moves ("I wanted to know if what I'm already doing is indeed correct, which, in the end, it was." EPI-T9, lines 942-943) and exploring ideas and practices that were previously overlooked, thus expanding the teachers' pedagogical toolkit ("We touched upon and discussed matters that we previously did not give the necessary attention to." EPI-T8, lines 450-451).

Fourth, the *integration of theory and practice* was another positive theme echoed by participants (Theme A.4): "What I really liked about the program is that it combined theory and practice; during our meetings, there was something like a theoretical part and a practical part—and then, applying the ideas in our classrooms." (EPI-T8, lines

29-32). This approach helped teachers learn to utilize different tools discussed during the sessions (e.g., design and utilize enablers and extenders: "The use of enablers and extenders was one of the most appealing and important tools for making a lesson successful and ensuring that all my students benefited from it." EPI-T2, lines 502-506) and engage more advanced students, who might otherwise remain under-challenged in the classroom ("It was really important that I learned how to activate these charismatic[31] students." EPI-T8, lines 445-446). Furthermore, teachers developed the skill to formulate appropriate and varied questions at different phases of a lesson ("The questions I pose to less advanced students will be different from those I ask students who have already solved the main task." EPI-T6, lines 168-170). They also acquired the ability to meticulously analyze tasks during the planning phase, categorizing them by their cognitive demands:

> Now, [when I plan my lessons], I think of every single task; I analyze the task in depth, even consider the questions to pose, or the different parts of the task if they need to be reorganized. I am trying to consider the different levels of challenge of different parts of the task. In other words, I am much more conscious and deliberate in how I plan and organize my lesson activities. (EPI-T9, lines 263-273)

This blending of theory with practice provided teachers with tangible ideas to engage all students, which is often not the case in traditional PLD models:

> What was really important to me is that, unlike other seminars in which you just get the theory, and you think 'These are nice ideas, but how can I implement them in practice?', in this program, every time I was leaving a session, I had a hundred of practical ideas to experiment with! Plus, I didn't have hesitations that these ideas are time demanding or difficult to implement. (EPI-T1, lines 335-343)

Fifth, the program fostered a profound *sense of ownership* among participants, attempting to carefully balance the introduction of new ideas without overwhelming them (Theme A.5). This was achieved by valuing the experiences teachers brought to the sessions, creating an environment in which knowledge was co-constructed with facilitators ("You [as the session facilitators] did not say, 'These are the five things that you need to be doing for differentiation'. Instead, there was a discussion, and ideas were developed collaboratively." EPI-T5, lines 271-274). The program was tailored to

[31] The term 'charismatic' ('χαρισματικός' in Greek) was used by Kate to refer to exceptionally able students. The term was retained as used by Kate to precisely convey the teacher's intended meaning.

meet the specific needs of its participants, thus allowing for the customization of its components ("The program was really responsive to our needs. I did not feel that there was a predetermined agenda from the beginning; the program really responded to our actual needs and offered ideas for practice." EPI-T4, lines 398-401). This approach facilitated the identification and reinforcement of teacher strengths and weaknesses, as Ariana suggested:

> We were learning something new, we could then implement it, and then discuss what we learned from this experimentation. [...] So, we were building on it [this new aspect], by identifying difficulties that we encountered and further working on it. This enhanced learning. (EPI-T9, lines 31-41)

By incrementally introducing concepts and practices, the program ensured that participants could digest and apply these ideas at a manageable pace, fostering an accumulative development of knowledge and skills:

> I felt that it was nice that we were getting ideas gradually, not everything at the same time, and that we were experimenting with them piece by piece: first in launching a task, then in autonomous work, and then during the discussion with the whole class. It was really nice that it was a piece each time because we could then implement them in our work, without saying, 'Lord, how can I do all these?'" (EPI-T5, lines 186-192)

In addition, the structural and procedural organization of the PLD program was positively perceived by the teachers (Theme A.6). By establishing clear learning objectives and structures for each session, the program ensured that participants had a guided learning experience (EPI-T4, lines 432-434). It adhered to a cyclical process of planning, teaching, observing, and reflecting, which fostered continuous improvement and deeper understanding ("All program aspects were important because they formed a continuous cycle; one thing led to another, and what we gained from one, we utilized it in the next." EPI-T1, lines 66-67). The provision of valuable materials and resources, along with summaries of key session ideas, supported ongoing learning and reflection, as Souzana described:

> The materials were extremely helpful because I would refer to them whenever I was preparing a lesson. [...] I really liked the notes that we were taking, because [in there] we were summarizing key ideas, such as practices, explanations, examples of questions, enablers and extenders, and lesson plans. I still refer to them frequently to see what I can utilize. (EPI-T6, lines 369-379)

These notes, as well as throught-provoking peer discussions, were integral, encouraging collaborative learning and the exchange of diverse perspectives ("Despite all my teaching experience, the discussions really problematized me and made me reflect a lot." EPI-T2, lines 71-72). The program's design, featuring an extended duration and spaced sessions, allowed participants ample time to absorb and apply new knowledge (EPI-T9, lines 1093-1094). Furthermore, receiving diverse feedback from multiple teacher educators on lesson plans and videotaped lessons enriched teachers' learning experience ("Each of you [as teacher educators] offered us feedback on our lessons, allowing us to gain a variety of perspectives." EPI-T1, lines 73-74).

Despite the positive feedback on the PLD program, certain negative aspects were identified by the participants, primarily concerning their *emotional and physical well-being* (Theme B.1). The process of videotaping lessons, a central component of the program, induced anxiety and discomfort among *all* teachers during the first stages of the program ("Videotaping presents a challenge we, as teachers, unanimously seek to minimize. The presence of an observer induces anxiety and can sometimes alter classroom dynamics." EPI-T9, lines 53-59). Additionally, the scheduling of continuous after-school sessions contributed to physical fatigue, suggesting a need for reconsidering the program's temporal structure to better accommodate teachers' well-being ("A challenge I faced was frequently feeling fatigued, as we would head to the sessions directly after school. Upon arriving I was feeling weary; however, I consistently left [the sessions] with a sense of enthusiasm." EPI-T1, lines 488-490). This challenge, however, was superficially encountered by only two teachers, suggesting that while the issue was present, its impact was perhaps not as widespread or deeply felt among the participants. Addressing these concerns could further improve the effectiveness and receptivity of the program, ensuring it not only fosters pedagogical growth but also maintains the overall well-being of its participants.

To augment the effectiveness of the PLD program, several constructive suggestions for improvement have been proposed by the teachers (Theme C.1). *Optimizing the organization and scheduling of the program* stood out as a crucial area for enhancement. Expanding the video-club component to allow more time for the viewing and discussion of longer lesson videoclips was also thought to facilitate deeper analysis and reflection ("It might have been more helpful to have more time to watch more examples, more excerpts from our teaching or others' teaching and to discuss and evaluate them." EPI-T9, lines 1242-1246). Additionally, a more defined schedule for videotaping lessons would likely improve the program's structure and teacher learning:

> It would be good to have a clearer, structured timeline for videotaping our lessons. This means learning a new practice, videotaping our experimentation with it, then moving on to learn a new practice and videotaping our trials with that immediately after its introduction. (EPI-T1, lines 544-548)

Furthermore, and contrary to what most teachers believed about the duration of the PLD program (see Theme A.6.d), a teacher suggested reducing the number of PLD sessions or arranging the sessions to be more closely spaced, without compromising the quality of content, to alleviate the burden and fatigue on teachers (EPI-T7, lines 345-357). Feedback mechanisms also emerged as an area for improvement; participants expressed a desire for more individualized feedback that could be tailored to their time constraints (EPI-T5, lines 308-311). Some teachers also suggested providing more examples of both effective and ineffective practices in their PLD discussions to enrich learning and understanding (EPI-T9, lines 1279-1281), while also tailoring PLD content/ideas to specifically address the needs and challenges of teaching lower elementary grades they may teach in the future (EPI-T9, lines 1284-1294). Lastly, extending the community of practice by providing additional I for communication would encourage continuous interaction and support among teachers outside of scheduled sessions:

> "It would have been great to have something like a platform which we could use to discuss ideas among ourselves from the one meeting to the other, during these, say, three weeks or a month [that elapsed] till the next meeting." (EPI-T9, lines 1303-1316)

Implementing these suggestions could significantly refine the program's design and delivery, ensuring it remains responsive to the needs of teachers.

In sum, the reactions of teachers to the program paint a relatively positive picture, albeit with certain suggestions for improvement and some points that were deemed as less favorable. The next section explores what the results level of Kirkpatrick's model reveals about the PLD's success.

4.2 Teachers' Performance at the Final Timepoint of the PLD Program

This section delves into the *Results* level of Kirkpatrick's model, outlining the *frequencies* of teachers' experimentation with teaching practices that promote cognitive activation, differentiation, or their interplay in their concluding lessons. It also evaluates the *quality* of these efforts.

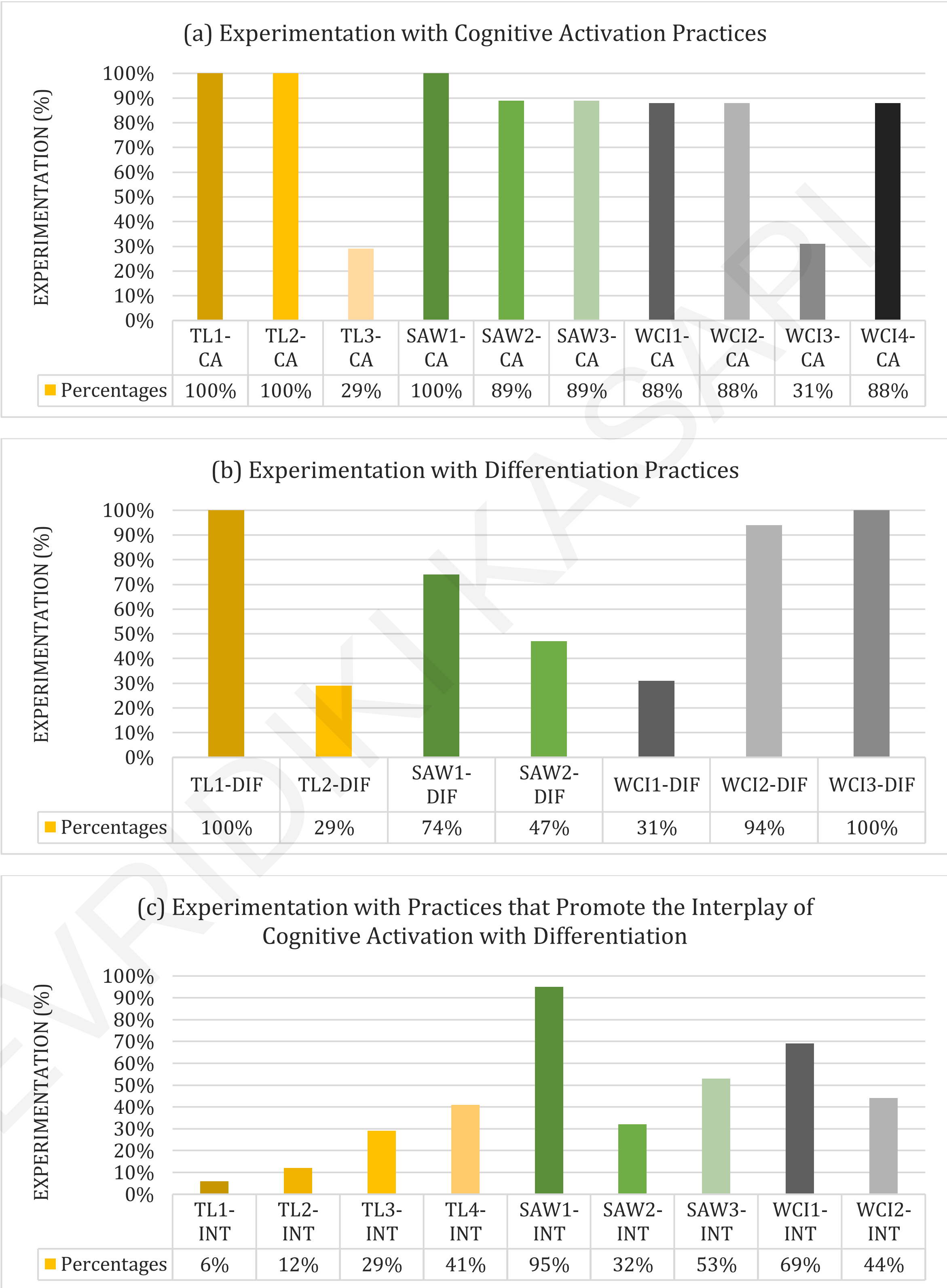
4.2.1 Frequency of Teachers' Use of Teaching Practices in their Concluding Lesson

This section presents the *frequency* of teachers' experimentation with various teaching practices, as captured in the coded videotaped lessons at the final timepoint. The figures presented in the next sections should not be interpreted as indicators of the *quality* of teaching. Rather, they reflect the practices with which teachers experimented in their videotaped lessons. Additionally, it should be acknowledged that the reported percentages might appear fairly high. This could be attributed to the methodology used for calculating these percentages, which involved comparing the number of occurrences of the lesson phase featuring a particular practice at the final timepoint over the total number of occurrences of the lesson phase at that timepoint (e.g., there were 13 occurrences of SAW at Timepoint 5 and in three of those the focal practice was observed).

Figure 15 captures the frequency of teachers' experimentation with (a) cognitive activation practices, (b) differentiation practices, and (c) practices that promote the interplay of both axes, at the final timepoint of the PLD program. Overall, the analysis revealed that the group experimented with *all* practices, though the extent of their experimentation varied. Notably, teachers were most actively experimenting with cognitive activation practices, followed by differentiation practices, and finally, the least engagement was seen in practices that combine both axes.

Figure 15

Teachers' Experimentation with Practices that Promote (a) Cognitive Activation, (b) Differentiation, and (c) the Interplay Between the Two Axes, at the Final Timepoint



Note.
The names of the practices referenced in this graph are listed in the Appendix 1.

As Figure 15a depicts, a strong commitment to selecting mathematically challenging tasks (TL1-CA) and maintaining their cognitive demands (TL2-CA) was observed in teachers' final lessons, with both practices showing 100% experimentation. This was reasonable since all the lessons were structured around such tasks. In addition, teachers were highly engaged in providing prompts (SAW1-CA) and in facilitating student engagement in mathematical reasoning without trivializing the challenge (SAW2-CA and SAW3-CA), as evidenced by high percentages close to or at 100%. While most whole-class interaction practices (WCI1-CA, WCI2-CA, and WCI4-CA) maintained high levels of experimentation, there was a significant drop for WCI3-CA (31%), in which teachers experimented less with asking students to compare or evaluate different solution approaches, highlighting a potential challenge and an area for further development. Less experimentation was also observed in the practice of discussing mathematical ideas at the outset of tasks (TL3-CA) with only 29% experimentation, something that teachers might have postponed for student autonomous work.

Figure 15b suggests that teachers experimented less with differentiation over cognitive activation practices in their final lesson. The data revealed more experimentation with selecting universally accessible tasks (TL1-DIF at 100%) and ensuring that student expressions of mathematical ideas were accessible to everyone (WCI3-DIF also at 100%). Additionally, in most occurrences of WCI, teachers emphasized making mathematical ideas prominent during lessons (WCI2-DIF at 94%), which underscores a commitment to shared understanding in the classroom. Experimentation with differentiation practices was less frequent in the phase of student autonomous work compared to the other phases (see Figure 15b); this is surprising given that this phase is particularly suitable for exploring and implementing varied differentiation ideas.

In each of the other phases, there was also one differentiation practice with which experimentation was limited. Specifically, only a minority of teachers focused on clarifying organizational decisions or management procedures for students working autonomously (TL2-DIF at 29%). This could indicate a potential oversight in equipping students to manage their work without guidance, or it might suggest that classroom norms and procedures had already been well-established in earlier lessons and did not require reiteration. Similarly, the practice of logically sequencing student solutions

(WCI1-DIF) appeared less frequently, hinting at challenges in arranging student contributions in a way that enhances group understanding.

Figure 15c reflects a *diverse landscape* of teacher experimentation with practices that foster the interplay of both axes. The engagement with practices that foster the synergy between cognitive activation and differentiation was less pronounced than the experimentation with each axis separately. This could be attributed to the complexity of merging both axes, requiring a solid understanding of each individually and how they can effectively complement one another—a conceptualization that teachers may not have fully mastered. Notably, the least experimentation was observed with practices that promote the interplay of both axes during task launching, especially with clarifying task aspects, both non-mathematical (TL1-INT at 6%) and mathematical (TL2-INT at 12%).

However, a notable observation was the significant use of targeted questioning (SAW1-INT) at 95%, underscoring a strategic emphasis on fostering deep understanding and customized learning. This practice was extensively discussed during the sessions, thus offering teachers concrete examples of appropriate questions that support both axes. Additionally, while questions are a fundamental tool for teachers, there is no data in Figure 15c illustrating if the quality of these questions followed the pattern of their usage frequency.

Using enablers (SAW3-INT) and extenders (SAW4-INT) were two of the practices that were less observed in student autonomous work (32% and 53%, respectively, see Figure 15c). The difference in experimentation with extenders over enablers suggested a greater focus or ease on challenging advanced learners or early finishers than on facilitating access to the task for less advanced students. This might be because advanced learners' needs were more visible, prompting a more frequent response or that less advanced students did not require enabling prompts to work on the tasks. However, enablers required a better understanding and anticipation of diverse student needs, which may account for their lower implementation rate. This implies that enhancing support for less advanced students remained an area for development.

While these figures offer a snapshot of experimentation frequencies at the final timepoint, there are several aspects they cannot convey. In particular, they cannot illustrate any *changes in the frequency* of experimentation across different timepoints, which will be discussed later (see Section 4.2.1). Moreover, they cannot show how effectively they were executed (i.e., the *quality aspect*). In the next section, the quality of teachers' performance at the final timepoint is presented, complementing the current findings.

4.2.2 Quality of Teachers' Performance in their Concluding Lesson

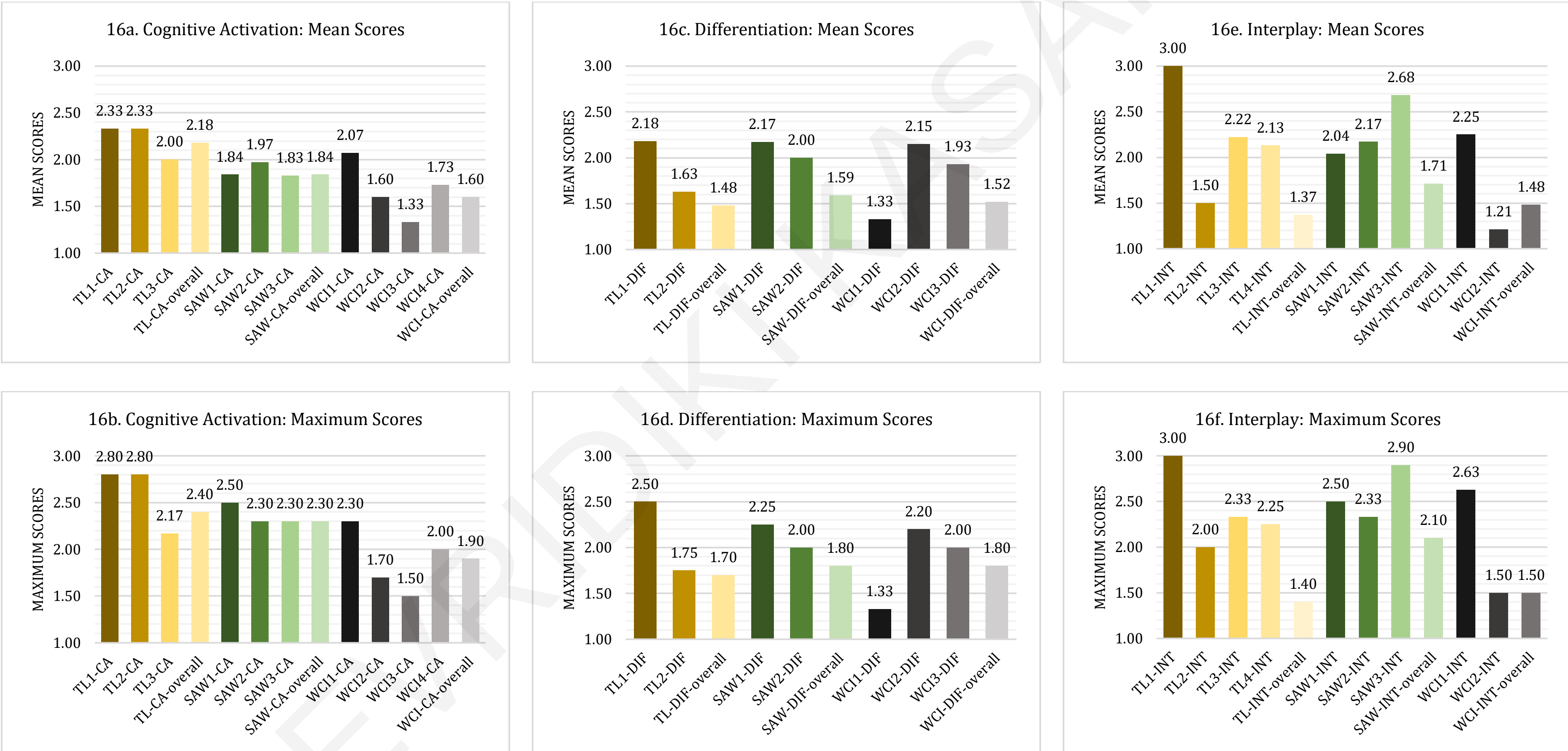
This section presents the *quality* of teachers' experimentation with the teaching practices at the final timepoint. To evaluate teachers' performance at the final timepoint, the scores from lessons videotaped at this timepoint were aggregated at the lesson level. This evaluation considered both the *mean* and *maximum* performance indicators (see Section 3.7.1 in Chapter 3).

It should be noted that the maximum scores are consistently higher than the mean scores across all practices. This is typical as maximum scores capture exceptional cases or outliers, whereas mean scores provide an average that is affected by lower-scoring instances. Also, the *overall* scores of each phase are lower than the individual practice scores within each phase. This may suggest that when coders considered all aspects of a phase collectively, they might have applied more stringent criteria or noticed more shortcomings than when they evaluated individual practices. It may also be indicative of the challenge of maintaining high quality consistently across all practices.

Figure 16 displays the mean and maximum performance of teachers with practices that enhance cognitive activation (16a and 16b), differentiation (16c and 16d), and their interplay (16e and 16f) as observed at the final timepoint of the PLD program.

Figure 16

Mean and Maximum Quality of Teachers' Experimentation with Practices Promoting Cognitive Activation, Differentiation, and their Interplay at the Final Timepoint of the PLD Program



Note.
The names of the practices referenced in this graph are listed in the Appendix 1.

To start, a pattern is clearly observable for cognitive activation: teachers performed better at launching the tasks, did less well with student autonomous work, and had the lowest performance in whole-class interactions. Also, the data reveal a generally uniform performance level within each lesson phase of this axis, except in whole-class interactions in which this consistency diminishes (see Figure 16a & 16b). In contrast, when considering differentiation and its interplay with cognitive activation, a significant variation in the quality of teachers' work regarding various practices becomes evident—especially in the latter, ranging from 1.21-3.00 (see Figure 16c, 16d, 16e, & 16f). This indicates that while practices for cognitive activation were applied uniformly, the incorporation of practices from the other axes showed greater variability in their implementation.

Interestingly, although teachers experimented *less* with practices promoting the interplay of both axes compared to practices that promote each axis separately (see Figure 16c), their mean and maximum performances in *most* practices were *higher*, ranging between 2 and 3 on the 3-point scale (see Figure 16e & 16f). This indicates that their modest experimentation yielded notably successful outcomes.

Among all the practices with which teachers experimented, two stood out for their high performance, both related to the *interplay* of the two axes: clarifying non-mathematical aspects of the task (TL1-INT) and using extenders (SAW3-INT). Following those are the practices of selecting and maintaining the demands for challenging tasks (see TL1-CA & TL2-CA, Figure 16a & 16b), which share identical mean and maximum scores close to the upper end of the scale (2.33 and 2.80). There are additional practices, mainly promoting the interplay of both axes, in which teachers have recorded mean scores above 2, reflecting a relatively moderate to high performance during their final lessons as a group.

The lowest performance was observed in practices from whole-class interactions across all axes, whose mean scores were well below the median of the scale (1.50): asking students to compare or evaluate different solutions (WCI3-CA); sequencing student solutions in a way that builds understanding (WCI1-DIF); and using incorrect or incomplete student solutions as resources for all student learning (WCI2-INT). These trends might suggest challenges in this phase—which is quite demanding— and a need for targeted PLD in orchestrating whole-class interactions.

A notable inconsistency between the extent of teachers' experimentation with certain practices and the quality of their performance was observed, highlighting a complex dynamic between their intentions and the effectiveness of their practices. For instance, despite limited attempts (6%, see TL1-INT, Figure 15c) at clarifying non-mathematical aspects of task wording during task launching, these efforts were notably successful (achieving a top score of 3, see Figure 16e).[32] Conversely, while there was a greater focus on directing differentiated questions (95%, see SAW1-INT, Figure 15c), the quality of these efforts did not translate into such high scores (2.04, see SAW1-INT, Figure 16e). In other practices such as SAW2-CA and SAW3-CA, teachers engaged in experimentation at similar levels (89%, see Figure 15a). However, on average, teachers were more effective at *prompting* students to engage in reasoning and/or meaning-making activities than in *actually* engaging them in these processes (see Figure 16a). Additionally, while SAW3-INT has high mean and maximum scores (2.68 and 2.90, respectively), indicating relatively successful engagement with advanced learners or early finishers, SAW2-INT is lower (2.17 and 2.33), pointing to potential challenges in providing appropriate enablers to less advanced students. The recurring pattern in both the quality and frequency of experimentation suggests a consistent area of strength in addressing the needs of advanced students, but also a consistent area for improvement in optimizing support for less advanced students without compromising the rigor of the tasks.

Finally, in all graphs, the mean and average scores generally align in their relative rankings; yet a distinct variation emerges during the phase of student autonomous work, particularly in practices aimed at cognitive activation and its interplay with differentiation. For instance, some teachers were particularly successful in providing mathematical prompts that challenged students without over-simplifying their thinking processes (SAW1-CA) and in directing diverse questions to different students to stimulate engagement in meaning-making, conceptual connections, or mathematical reasoning (SAW1-INT), as evidenced by the maximum score patterns (both at 2.50, see Figure 16b & 16f). As a group, they seem to have had a relatively moderate performance (SAW1-CA at 1.84 and SAW-INT at 2.04, see Figure 16a & 16e). This discrepancy between mean and maximum scores indicates a range in teacher performance levels, not captured in this analysis of the teachers as a group, as discussed later (see Chapters 5, 6, and 7).

---

[32] The larger gap between the mean scores of TL1-INT (3.00) and TL2-INT (1.50) suggests some inconsistency in how teachers clarify tasks. Perhaps, in their attempt to clarify *mathematical* aspects, they may have been doing the thinking for their students, whereas, with non-mathematical aspects, they did not run this risk.

4.2.3 Summarizing Key Findings on the Frequency of Teachers' Experimentation and the Quality of Teachers' Performance at the Final Timepoint

The exploration of teachers' experimentation with different teaching practices and the quality of their work thereof at the final timepoint of the PLD program unveils key insights. First, teachers displayed higher engagement with cognitive activation practices, underscoring their commitment to fostering students' thinking and problem-solving skills. However, while teachers also engaged in differentiation practices, the extent of their experimentation was less consistent compared to cognitive activation. Experimentation with practices that integrate cognitive activation and differentiation was even less frequent, suggesting challenges in merging these two axes. This may indicate the complexity of effectively combining these approaches or a need for more focused PLD support in this area.

Second, the quality of teachers' experimentation, as reflected in their performance scores, varied across different practices. While there was high engagement in certain practices, the quality of these efforts (as measured by mean and maximum performance scores) varied. Notably, practices that merged cognitive activation with differentiation often resulted in higher performance, despite being less frequently experimented with, suggesting that when such integrated practices were employed, they were generally of higher quality.

Third, the analysis also revealed a mismatch between the frequency of experimenting with certain practices and their quality. For instance, limited attempts at clarifying non-mathematical task aspects led to remarkably high performance, highlighting the potency of small amounts of targeted experimentation.

Strengths were consistently observed in targeted questioning and the use of extenders. Conversely, areas needing improvement were identified, especially in supporting less advanced students and optimizing the management of student autonomous work. Additionally, challenges in maintaining high-quality whole-class interactions, such as facilitating discussions that compare different solution approaches, point to potential development areas in teacher PLD.

Fourth, the range between mean and maximum scores, especially during the phase of student autonomous work, suggests variability in individual teacher performance. This variability underscores the diversity in teacher effectiveness and the potential need for personalized PLD to address specific areas of need.

Finally, the PLD program's evaluation at the final timepoint showcases its success in specific axes, notably in practices that enhance cognitive activation and the interplay of cognitive activation with differentiation (albeit with less experimentation with the latter). The program's strengths were evident in the complete engagement of teachers in selecting and maintaining mathematically challenging tasks, a crucial aspect of cognitive activation, and the high-quality outcomes in task clarification and the use of extenders. However, the overall success of the PLD program was moderated by the variability in the extent and quality of experimentation across certain practices, especially in whole-class interactions and in providing support for less advanced students. Also, the quality of teachers' experimentation with differentiation was less consistent. The notable discrepancy between mean and maximum scores, particularly evident in student autonomous work, underscores the diversity in teacher performance and indicates a potential need for more personalized PLD approaches to meet individual teacher needs. The Results level of the PLD program presents a mixed picture of teacher experimentation and performance, with some areas of strength and others needing improvement.

4.3 The Evolution of Teachers' Behavioral Processes Across the Timepoints

Transitioning to the Behavior level of Kirkpatrick's model, the upcoming sections detail the manifestation and evolution of the teaching practices used by the video-club group. These sections cover the variance in the frequency of experimentation, as well as the teaching quality of each practice over the five timepoints.

4.3.1 Teachers' Experimentation Across the Timepoints

This section presents the findings related to teachers' experimentation with various teaching practices across five timepoints. As outlined in Chapter 3, the analysis utilized percentages to represent the use of teaching practices that promote cognitive activation, differentiation, or their interplay, observed in the videotaped lessons within *each* of the five timepoints (see Section 3.7.1).

The analysis uncovered several prominent patterns in teachers' experimentation across the three axes—cognitive activation, differentiation, and their interplay—throughout the PLD program. A significant insight was that the teacher participants experimented with *all* cognitive activation, differentiation, and interplay practices that surfaced and discussed during the PLD program in their videotaped lessons, albeit in *varying degrees* across timepoints.

Teachers consistently engaged in cognitive activation practices with high levels of experimentation, indicating a strong commitment to enhancing students' mathematical thinking. While they also used differentiation practices, their levels of experimentation varied, particularly during phases of student autonomous work, revealing areas of inconsistency. The integration of cognitive activation and differentiation practices was observed less frequently, likely due to the challenges inherent in merging these complex pedagogical axes. This indicates a need for more focused support within the PLD program to navigate the subtleties of applying practices that promote the interplay of cognitive activation and differentiation effectively. However, notably, there was a clear shift towards prioritizing some practices promoting the interplay of both axes that foster independent mathematical reasoning among students, highlighting a progressive move towards student-centered learning strategies.

A detailed account of teachers' experimentation during different phases of their lessons—namely, task launching, student autonomous work, and whole-class interactions—is presented in the following sections. These sections provide detailed insights into specific practices employed. As noted in Section 4.2, the figures in subsequent sections do not reflect teaching *quality,* but rather show the practices with which teachers opted to experiment in their videotaped lessons. The relatively high reported percentages result from using the quotient of the number of occurrences of the lesson phase featuring a particular practice at *each* timepoint over the total number of occurrences of the lesson phase at that timepoint (e.g., there were 13 occurrences of SAW at Timepoint 3 and in three of those the focal practice was observed). The presentation of the findings begins with the practices associated with cognitive activation, proceeds to those related to differentiation, and finally addresses practices that lie at the interplay of both axes.

4.3.1.1 Experimentation with Cognitive Activation Practices Across the Timepoints

Figure 17 showcases the frequency of teachers' experimentation with cognitive activation practices during the three phases of their lessons, across the five timepoints. Teachers' experimentation with nearly all identified cognitive activation practices was comparable across timepoints, with only two practices being less frequently explored. Across most practices and timepoints, the experimentation rates were either sustained or showed an increase, indicating a growing emphasis on these cognitive activation strategies as the program progressed. The slight fluctuations observed in the consistently high practices might reflect the natural ebb and flow of classroom dynamics or could be indicative of the evolving nature of teachers' comfort and skill in implementing these practices.

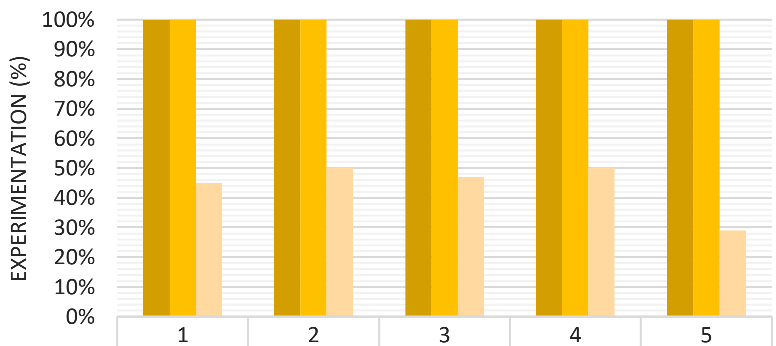
Figure 17a shows a high and consistent experimentation with challenging tasks. The unwavering 100% experimentation in selecting mathematically challenging tasks (TL1-CA) and in maintaining the cognitive demands during task launching (TL2-CA) was expected since all videotaped lessons involved challenging tasks. It could also stem from the repeated emphasis on task analysis and challenge in the early video-club sessions. This, along with the revamped student textbooks offering more complex tasks, could have supported the increased use of such tasks in lessons.

However, the experimentation of teachers with discussing mathematical ideas with students (TL3-CA) during task launching was notably lower than TL1-CA and TL2-CA. This practice also saw a significant drop at the fifth timepoint, which could suggest challenges or shifts in focus as the program progressed. The reasons for this could be multifaceted, ranging from time constraints within lessons to possible uncertainty in guiding conversations that delved deeply into mathematical concepts. Teachers likely refrained from too much guidance to prevent giving away hints or solutions before students' autonomous work.

In student autonomous work phases, experimentation with providing mathematical prompts (SAW1-CA) and soliciting student engagement in reasoning and meaning-making activities (SAW2-CA and SAW3-CA) was overall high, with a notable increase to 100% by the fifth timepoint for SAW1-CA (see Figure 17b). This suggests an increasing emphasis on promoting independent mathematical thinking over time.

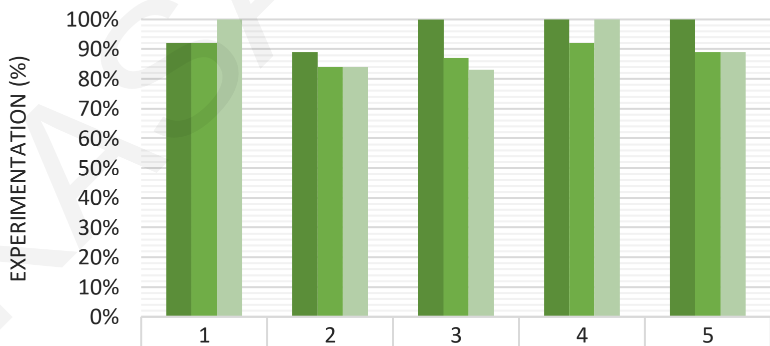
Figure 17

Teachers' Experimentation with Cognitive Activation Practices During (a) Task Launching; (b) Student Autonomous Work; and (c) Whole-class Interactions, Across the Five Timepoints.



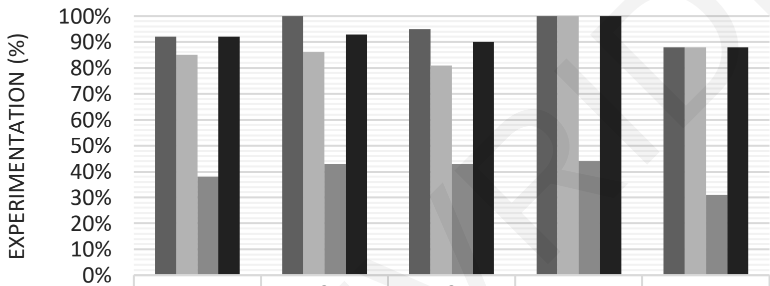
	1	2	3	4	5
TL1-CA	100%	100%	100%	100%	100%
TL2-CA	100%	100%	100%	100%	100%
TL3-CA	45%	50%	47%	50%	29%

(a)



	1	2	3	4	5
SAW1-CA	92%	89%	100%	100%	100%
SAW2-CA	92%	84%	87%	92%	89%
SAW3-CA	100%	84%	83%	100%	89%

(b)



	1	2	3	4	5
WCI1-CA	92%	100%	95%	100%	88%
WCI2-CA	85%	86%	81%	100%	88%
WCI3-CA	38%	43%	43%	44%	31%
WCI4-CA	92%	93%	90%	100%	88%

(c)

TL1-CA-Selecting mathematically challenging tasks; **TL2-CA**-Maintaining the cognitive demands of the task as presented to students during task launching; **TL3-CA**-Discussing mathematical ideas as presented to students; **SAW1**-Providing mathematical prompts to students without trivializing their thinking; **SAW2-CA**-Asking students to engage in mathematical reasoning and/or meaning-making activities; **SAW3-CA**-Engaging students in mathematical reasoning and/or meaning-making activities; **WCI1-CA**-Eliciting instances of student mathematical reasoning and/or meaning/making; **WCI2-CA**-Synthesizing and extending important mathematical ideas; **WCI3-CA**-Asking students to compare or evaluate different solution approaches; **WCI4-CA**-Engaging students in mathematical reasoning and meaning-making activities.

Figure 17c also reveals that during whole-class interactions, teachers' experimentation with eliciting instances of student reasoning (WCI1-CA) and synthesizing and extending important mathematical ideas (WCI2-CA) was robust and consistent. These high rates of experimentation suggest a dynamic classroom environment. However, the lower and less consistent experimentation rates for asking students to compare or evaluate different solution approaches (WCI3-CA), with a downward trend to 31% by the fifth timepoint, could signal a missed opportunity for peer learning and critical evaluation.

Interestingly, the engagement in mathematical reasoning and meaning-making activities (WCI4-CA) was high, paralleling the emphasis on eliciting and extending mathematical reasoning (WCI1-CA). This indicates that, while comparing solutions was less common, the overall goal of engaging students in higher-order thinking during whole-class sessions was a focal point of the lessons observed.

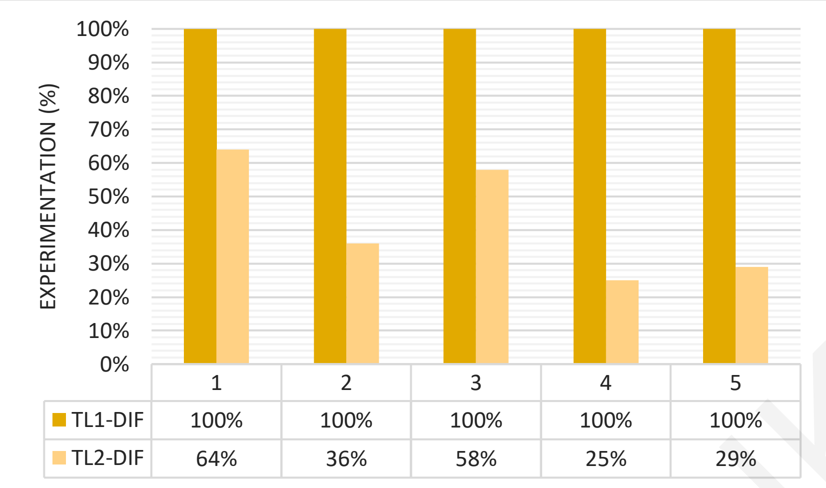
Weaving these threads together, it appears that the teachers involved in the PLD program were deeply invested in cognitive activation, demonstrated by their consistent choices in tasks and sustained efforts to engage students in autonomous and collective mathematical thinking. While the implementation was robust, areas such as facilitating in-depth discussions and comparing different solution approaches highlight opportunities for further professional growth.

4.3.1.2 Experimentation with Differentiation Practices Across the Timepoints

Figure 18 sheds light on the frequency of teachers' experimentation with differentiation practices across the five timepoints. Overall, teachers experimented to a lesser extent with differentiation practices—especially during student autonomous work—than with cognitive activation practices. Furthermore, the graphs indicated that while there was a commitment to or increment in experimentation with certain differentiation practices, teachers exhibited varying or low levels of experimentation with others, which could have benefited from additional support and development within the PLD program.

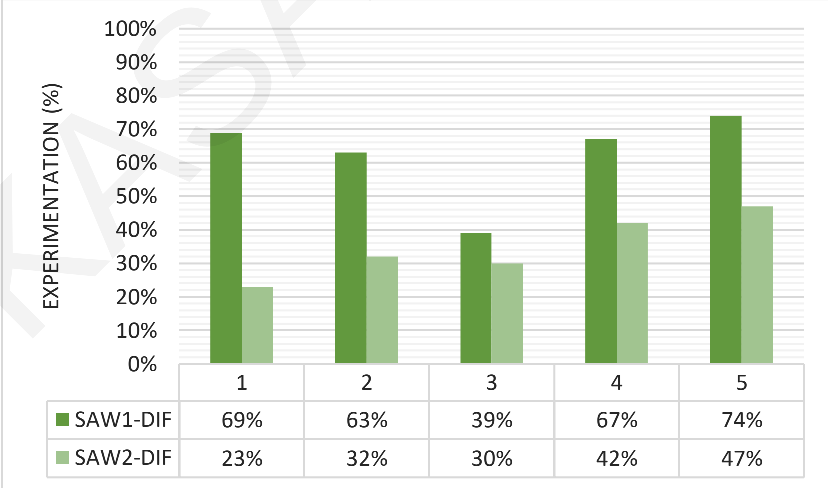
Figure 18

Teachers' Experimentation with Differentiation Practices During (a) Task Launching; (b) Student Autonomous Work; and (c) Whole-class Interactions, Across the Five Timepoints.



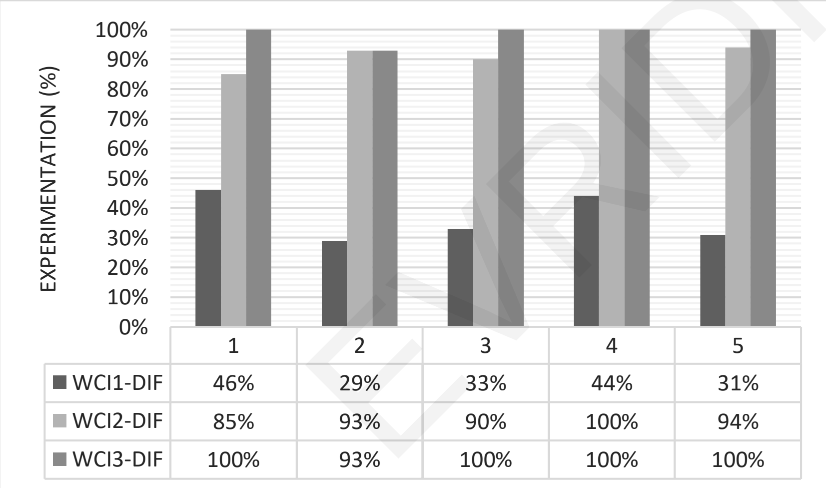
	1	2	3	4	5
TL1-DIF	100%	100%	100%	100%	100%
TL2-DIF	64%	36%	58%	25%	29%

(a)



	1	2	3	4	5
SAW1-DIF	69%	63%	39%	67%	74%
SAW2-DIF	23%	32%	30%	42%	47%

(b)



	1	2	3	4	5
WCI1-DIF	46%	29%	33%	44%	31%
WCI2-DIF	85%	93%	90%	100%	94%
WCI3-DIF	100%	93%	100%	100%	100%

(c)

TL1-DIF-Selecting tasks which are potentially accessible to all students; **TL2-DIF**-Making clear the organizational decisions or management procedures for working autonomously on the task; **SAW1-DIF**-Using student asynchronous work to accommodate different learning readiness levels and needs; **SAW2-DIF**-Encouraging multiple expressions of content, process, and/or product; **WCI1-DIF**-Sequencing student solutions in a reasonable progression to support student access to the ideas shared; **WCI2-DIF**-Highlighting important mathematical ideas during the sharing to ensure that these are made clear to as many students as possible; **WCI3-DIF**-Students express mathematical ideas that are visible and/or audible to all students (as well as the teacher).

For some practices, such as selecting tasks accessible to all students (TL1-DIF, see Figure 18a) and ensuring that students' mathematical ideas were visible/audible to all (WCI3-DIF, see Figure 18c), there was a high and consistent rate of experimentation across all timepoints, often reaching 100%. The tasks might have had an inherently relatable context and content, which naturally invited multiple routes of accessibility. The 169elevated levels of experimentation with practice WCI3-DIF suggest that student ideas and solutions were presented in some way.

In student autonomous work (see Figure 18b), there was a general upward trend in the experimentation with practices SAW1-DIF and SAW2-DIF over the five timepoints, suggesting that teachers had been progressively engaging more with these practices as the program continued. The uptick in SAW1-DIF, which generally required less preparatory effort from teachers, gained traction likely because of the introduction and exploration of asynchronous work modalities during VCS2. Despite this upward trend, the rate of experimentation for SAW2-DIF was lower than for SAW1-DIF. This could suggest that promoting multiple expressions of content, process, and product remained less frequently implemented, indicating potential complexities in its application that may not have been fully addressed in the video-club sessions.

The frequency of teachers' experimentation with highlighting key mathematical concepts during discussions to ensure student comprehension (WCI2-DIF) was consistently high and increasing (see Figure 18c). It started at a high rate of 85% and showed a growing trend, reaching 100% by the fourth timepoint and slightly decreasing to 94% at the fifth timepoint. This pattern suggests a generally increasing emphasis on this practice over time, with a slight dip towards the end that did not significantly deviate from the overall high level of experimentation. This increment may reflect the use of practices explored during the sessions, such as revoicing; explaining; and using representations.

Other practices showed variability in experimentation across different timepoints. For example, making organizational decisions or procedures clear for autonomous work (TL2-DIF, see Figure 18a) fluctuated significantly, suggesting inconsistencies in teachers' approaches or possibly varying levels of necessity for such clarity over time. A slight increase in TL2-DIF during the third timepoint might reflect the impact of discussions in VCS3.

Lower and fluctuated rates of experimentation with complex practices such as sequencing student solutions to facilitate understanding (WCI1-DIF, see Figure 18c), which never reached 50% across any timepoint, suggest that teachers may have found consistent implementation challenging. Despite discussions in VCS3 and VCS4, the application of WCI1 was sporadic.

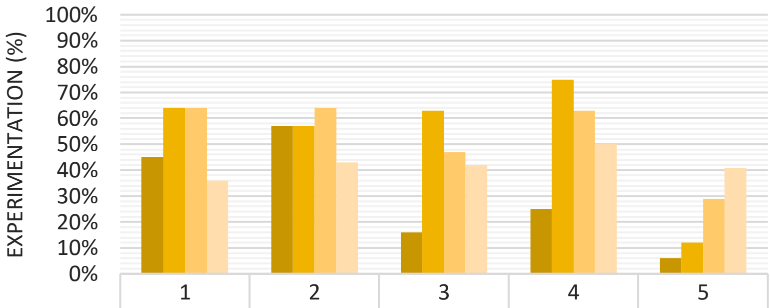
In sum, throughout the PLD program, teachers consistently engaged in practices to ensure accessibility and comprehension of mathematical tasks and concepts, with certain practices like task selection and highlighting key ideas showing high and increasing rates of experimentation. While there was a positive trend in engaging students in asynchronous work, some complex practices like sequencing solutions during whole-class interactions were less consistently applied, pointing to a need for ongoing support in these areas.

4.3.1.3 Experimentation with Practices that Promote the Interplay of Both Axes Across the Timepoints

Figure 19 illustrates the frequency of teachers' experimentation with practices that potentially foster the interplay of cognitive activation with differentiation across the five timepoints. Overall, teachers were less involved with practices combining cognitive activation and differentiation compared to individual axis-focused practices, likely due to the complexity of integrating both. This trend may stem from introducing this interplay later in the PLD program (from Timepoint 3). Moreover, applying all practices at once could be challenging, and for many teachers, it was their first experience simultaneously considering both pedagogical axes, accounting for the low to moderate experimentation with such integrated practices.

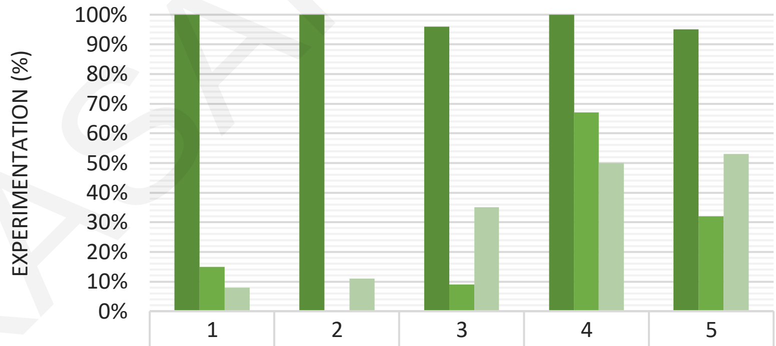
Figure 19

Teachers' Experimentation with Practices that Promote the Interplay of Both Axes During (a) Task Launching; (b) Student Autonomous Work; and (c) Whole-class Interactions, Across the Five Timepoints.



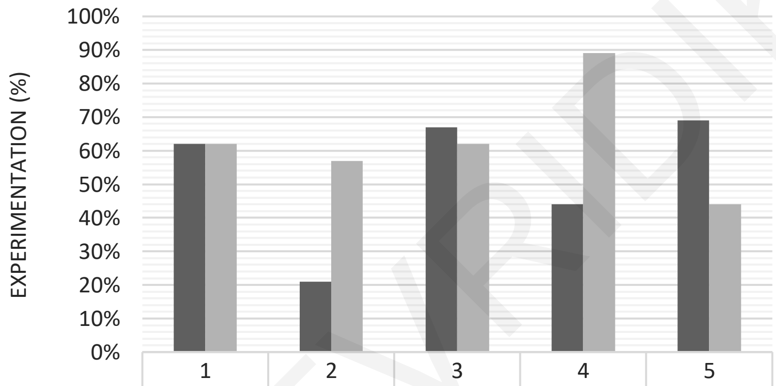
	1	2	3	4	5
TL1-INT	45%	57%	16%	25%	6%
TL2-INT	64%	57%	63%	75%	12%
TL3-INT	64%	64%	47%	63%	29%
TL4-INT	36%	43%	42%	50%	41%

(a)



	1	2	3	4	5
SAW1-INT	100%	100%	96%	100%	95%
SAW2-INT	15%	0%	9%	67%	32%
SAW3-INT	8%	11%	35%	50%	53%

(b)



	1	2	3	4	5
WCI1-INT	62%	21%	67%	44%	69%
WCI2-INT	62%	57%	62%	89%	44%

(c)

TL1-INT-Explaining potentially unfamiliar non-mathematical aspects of the wording of the task; **TL2-INT**-Clarifying mathematical aspects of the task; **TL3-INT**-Posing questions that indicate the level of support that students need in order to engage in the task; **TL4-INT**-Activating relevant existing mathematical knowledge and strategies; **SAW1-INT**-Directing different types of questions to different students for engaging them in meaning-making, conceptual connections, or mathematical reasoning; **SAW2-INT**-Providing enablers to facilitate access to the task at hand without reducing the challenge; **SAW3-INT**-Providing extenders to advanced learners or early finishers; **WCI1-INT**-Holding students accountable for attending to and understanding their classmates' sharing; **WCI2-INT**-Using incorrect or incomplete student solutions as resources for all student learning.

Student autonomous work practices showed a consistent use of varied questioning and a notable rise in support and challenge after focused PLD (see Figure 19b). In particular, the strategic deployment of diverse questions to individual students or groups during student autonomous work exhibited a high level of experimentation consistently (SAW1). In contrast, SAW2 and SAW3 started with very low rates of experimentation which then increased notably by the fourth timepoint. This indicated growing attention to providing support (enablers) and extending the challenge (extenders) to students as the program progressed, although they never reached the high levels of SAW1. This was not surprising since the tool of enablers and extenders was introduced in VCS5 (corresponding to Timepoint 4), and teachers endorsed and experimented with both tools in their lessons significantly. Notably, the use of extenders saw an increase earlier (at Timepoint 3), indicating an interest among teachers in extending learning for advanced learners or early finishers—a concern that was actively discussed during VCS2 and VCS3. This aspect is further discussed in the next chapters.

Variability was observed in experimentation with task-launching practices over the timepoints, revealing an intricate narrative (see Figure 19a). Initially, teachers engaged moderately well with explaining both the non-mathematical (TL1) and mathematical aspects of tasks (TL2), reflecting an understanding of the importance of clarity in task presentation. An early focus on TL2 may have indicated a particular emphasis on the precision of mathematical language and concepts, essential for students to begin their work on solid footing (as discussed during VCS5). TL3—asking probing questions to determine students' needs—remained relatively stable but decreased towards the end. This could reflect the complexity of accurately assessing and responding to students' varying needs. Fluctuations in TL4—activating students' existing knowledge to approach new tasks—might have been due to the challenges in seamlessly integrating new content with students' prior learning or varying the focus of different units within the curriculum that may not have lent themselves easily to such connections. The significant decline in all task-launching practices by the fifth timepoint suggested a potential shift in focus or a change in perceived necessity. This might indicate that teachers felt that students had become more adept at understanding tasks without extensive discussions, or a need for re-emphasizing the importance of these practices in PLD sessions.

A closer examination of whole-class interaction practices yielded several patterns (see Figure 19c). The rates for WCI1—ensuring that students were attentive and comprehended their classmates' contributions—showed variability. This suggests that teachers recognized the importance of active listening and peer learning but may have found it challenging to consistently hold students to this standard. The dip at certain timepoints could indicate periods in which less emphasis was placed on this practice or in which it was overshadowed by other priorities in the classroom. The experimentation with using incorrect or incomplete student solutions as learning resources (WCI2) showed a significant peak at the fourth timepoint, suggesting that there might have been a concerted effort or a specific focus during the PLD program on using mistakes constructively. In particular, these shifts can be linked to discussions in VCS 3-4, which highlighted strategies such as encouraging students to share, discuss, revoice, rephrase, or paraphrase their classmates' correct and incorrect ideas, strategies, or solutions. However, the subsequent decline indicates that sustaining such focus over time might be difficult. Overall, fluctuating experimentation may reflect teachers' ongoing learning curve in managing and facilitating effective whole-class discussions.

4.3.1.4 Summarizing Key Findings on the Frequency of Teachers' Experimentation Across the Timepoints

Throughout the PLD program, teachers embarked on a journey of pedagogical evolution, marked by their experimentation with a spectrum of teaching strategies across cognitive activation, differentiation, and their interplay. The data across the five timepoints paints a picture of a committed teaching cohort, keen on deepening their students' mathematical understanding while grappling with the inherent complexities of differentiating teaching and enhancing the interplay between the two axes. The findings not only highlight the areas of strength but also illuminate the pathways for professional growth.

In the realm of cognitive activation, teachers collectively demonstrated a steadfast dedication, particularly in choosing mathematically challenging tasks and trying to maintain their cognitive rigor when presented to students. However, the journey was not without its hurdles. The practice of engaging students in discussions about mathematical ideas and the comparison and evaluation of different student ideas/solutions, critical aspects of cognitive activation, saw a decline in experimentation, especially noticeable towards the program's conclusion. This trend might hint at potential obstacles—be it time constraints or the intricacies of facilitating meaningful mathematical discourse—that could benefit from targeted support in future PLD initiatives.

When it came to differentiation practices, the landscape was more varied. Teachers seemed to navigate this terrain with more caution, particularly in the phase of student autonomous work. While certain practices such as ensuring task accessibility and the visibility/audibility of mathematical ideas were embraced consistently, others saw a more fluctuating level of engagement. This variability underscores the nuanced challenges teachers face in differentiating learning experiences and suggests a need for ongoing PLD to increase confidence and knowledge in applying differentiation practices.

The interplay of cognitive activation and differentiation practices presented its own set of challenges, reflected in the less frequent experimentation with practices that embody this interplay. The experimentation with these practices seems to be a quite complex endeavor for teachers, as they navigate through a multitude of ideas and practices in their teaching. Yet, the gradual uptick in the use of enablers and extenders is a beacon of progress, indicating a growing teacher attentiveness to providing tailored support and extending challenges to meet diverse student needs.

Perhaps the variability in experimentation with some practices could be attributed to factors, such as the content of the lessons, the objectives or focal points of each video-club session, the challenges posed by implementing certain teaching practices, and teachers' priorities and goals (as will be discussed in Chapters 5, 6 and 7), some teachers had predetermined objectives that might not fully align with or only partially intersect with the aims and content of the PLD program). Therefore, it appears that teachers' experimentation, as a facet of the processes of learning, unlearning, or relearning, might not be a straightforward process of knowing and implementing the "right" or effective practices. Instead, it is a journey marked by its ebbs and flows, and one cannot expect a steady linear progression over time.

It is important to note that while we can speculate on the PLD's influence on experimentation, how its structure and content specifically shaped this experimentation cannot be determined. Furthermore, the reasons and the context behind the choice of certain practices over others can only be inferred, such as the specific challenges of a classroom or individual teacher's needs, motivations, and hesitations. Therefore, only assumptions about the sources and conditions that influence the frequency of teachers' experimentation with teaching practices can be made without the analysis of qualitative data.

Finally, it is important to note that high frequency does not necessarily equate to high quality; a teacher may frequently implement a practice but not necessarily in a way that maximizes student learning. Therefore, while the consistency and variation in experimentation frequency are informative, they are just one piece of the puzzle. A comprehensive assessment would require a qualitative analysis to determine the depth and quality of this experimentation. While the aforementioned findings highlight that teachers seized the opportunity to experiment with various practices across multiple videotaped lessons, they do not address the quality of these experimentations, which is the research point we now turn to.

4.3.2 Quality of Teachers' Performance Across the Timepoints

This section presents the findings related to the quality of teachers' performance in various teaching practices across the five timepoints. Notably, the figures presented in this section reflect the collective situation of the entire sample and do not account for individual changes within teachers, a topic that is explored in the following chapters. To identify any statistically significant changes in teachers' collective performance, the non-parametric test Wilcoxon signed-rank test was used to compare the mean and maximum aggregated scores of teacher performance for each practice within a timepoint between matched pairs of timepoints (see Section 3.7 in Chapter 3).

In all figures presented in the following sections, the crosses and asterisks serve as visual indicators of statistical significance, each symbol denoting a specific level of statistical difference between practices over two timepoints. Asterisks (*) were used to highlight differences that were statistically significant at the .05 level, while crosses (†) were employed to indicate differences that were statistically significant at the .10 level. The same color for crosses or asterisks over certain bars implies that the practices represented by those bars showed statistically significant differences at the specified levels (.05 for asterisks, .10 for crosses) when comparing outcomes at two different timepoints.

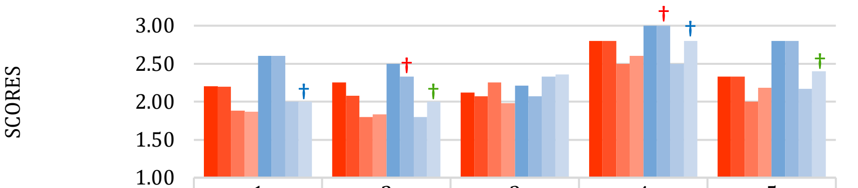
Across the timepoints, teachers' quality of performance exhibited fluctuations with moments of significant improvement and instances of decline. A notable peak in the task launching and student autonomous work phases was primarily observed around Timepoints 3 and 4. However, this high level of performance was not consistently sustained into Timepoint 5. In contrast, whole-class interactions did not show a clear trend of improvement, with some practices even experiencing a decline.

The reader is notified of some limitations mentioned earlier in Chapter 3. For instance, only four lessons were videotaped per teacher. Also, teachers had the discretion to videotape lessons at different junctures during the PLD program, potentially choosing to do so after gaining comfort with the program's concepts or with their selected lesson content. The absence of a baseline lesson for each teacher, which would have benchmarked their initial teaching quality, complicates the accurate tracking of performance trends in their experimentation. In what follows, the findings on teachers' performance are presented beginning with cognitive activation, followed by differentiation, and concluding with practices that foster the interplay of both cognitive activation and differentiation.

4.3.2.1 Quality of Teachers' Experimentation with Cognitive Activation Practices Across the Timepoints

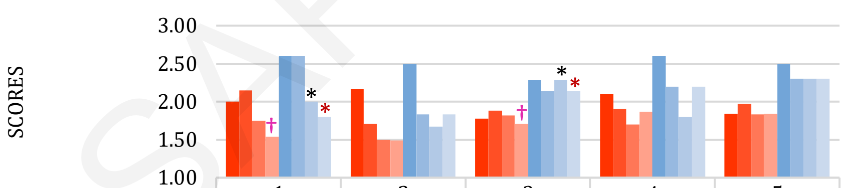
Figure 20 delineates the teaching quality of teachers' experimentation with practices of cognitive activation within each phase, across the five timepoints. Throughout the PLD program, most cognitive activation practices in task launching, student autonomous work, and whole-class interactions showed notable fluctuations and no dramatic changes, with a peak in performance at Timepoint 4. This peak, especially evident in maintaining task demands during task launching and overall task launching, suggests a moment of exemplary practice. However, this improvement was not sustained into Timepoint 5. While some practices related to student autonomous work, such as engaging students in mathematical reasoning, showed significant improvement by Timepoint 3, overall consistency remained elusive. Whole-class interactions varied, with no clear trend of improvement, despite some positive shifts observed up to Timepoint 4.

Figure 20. *Quality of Teachers' Experimentation with Cognitive Activation Practices During (a) Task Launching; (b) Student Autonomous Work; and (c) Whole-class Interactions, Across the Five Timepoints*



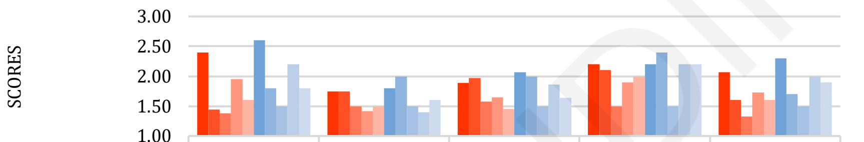
	1	2	3	4	5
TL1-mean	2.20	2.25	2.12	2.80	2.33
TL2-mean	2.20	2.08	2.07	2.80	2.33
TL3-mean	1.88	1.80	2.25	2.50	2.00
TL-overall-mean	1.87	1.83	1.98	2.60	2.18
TL1-max	2.60	2.50	2.21	3.00	2.80
TL2-max	2.60	2.33	2.07	3.00	2.80
TL3-max	2.00	1.80	2.33	2.50	2.17
TL-overall-max	2.00	2.00	2.36	2.80	2.40

(a)



	1	2	3	4	5
SAW1-mean	2.00	2.17	1.78	2.10	1.84
SAW2-mean	2.15	1.71	1.88	1.90	1.97
SAW3-mean	1.75	1.50	1.82	1.70	1.83
SAW-overall-mean	1.54	1.49	1.71	1.87	1.84
SAW1-max	2.60	2.50	2.29	2.60	2.50
SAW2-max	2.60	1.83	2.14	2.20	2.30
SAW3-max	2.00	1.67	2.29	1.80	2.30
SAW-overall-max	1.80	1.83	2.14	2.20	2.30

(b)



	1	2	3	4	5
WCI1-mean	2.40	1.75	1.89	2.20	2.07
WCI2-mean	1.44	1.75	1.97	2.10	1.60
WCI3-mean	1.38	1.50	1.58	1.50	1.33
WCI4-mean	1.95	1.42	1.65	1.90	1.73
WCI-mean-overall	1.60	1.50	1.45	2.00	1.60
WCI1-max	2.60	1.80	2.07	2.20	2.30
WCI2-max	1.80	2.00	2.00	2.40	1.70
WCI3-max	1.50	1.50	1.50	1.50	1.50
WCI4-max	2.20	1.40	1.86	2.20	2.00
WCI-overall-max	1.80	1.60	1.64	2.20	1.90

(c)

TL1-Selecting mathematically challenging tasks; **TL2**-Maintaining the cognitive demands of the task as presented to students during task launching; **TL3**-Discussing mathematical ideas as presented to students; **TL-overall**-Holistic estimate of the challenging work during this phase; **SAW1**-Providing mathematical prompts to students without trivializing their thinking; **SAW2**-Asking students to engage in mathematical reasoning and/or meaning-making activities; **SAW3**- Engaging students in mathematical reasoning and/or meaning-making activities; **SAW-overall**-Holistic estimate of the challenging work during this phase; **WCI1**-Eliciting instances of student mathematical reasoning and/or meaning/making; **WCI2**-Synthesizing and extending important mathematical ideas; **WCI3**-Asking students to compare or evaluate different solution approaches; **WCI4**-Engaging students in mathematical reasoning and meaning-making activities; **WCI-overall**-Holistic estimate of the challenging work during this phase.

Same color * signifies statistical significance at .05 level, and † at .1 level, between two timepoints.

In the task launching phase (see Figure 20a), teachers' mean and maximum performance for almost all practices under consideration was higher than 2.00. Notably, the performance of these practices illustrated a *mixed trend* across the initial three timepoints: while there were fluctuations in discussing mathematical ideas (TL3), the practices of selecting challenging tasks (TL1) and maintaining the cognitive demands of task presented to students (TL2) displayed a declining trend.

These fluctuations were punctuated by an increase in both mean and maximum scores at the fourth timepoint—approaching (nearly) 3.00—before experiencing a slight downturn at the fifth. This progression was more pronounced in the maximum performance of teachers at Timepoint 4, particularly noticeable in the practice of maintaining the cognitive demands of tasks (TL2-max) and in the maximum overall quality of task launching practices (TL-overall-max), which reached their zenith. The Wilcoxon Signed-ranks test revealed a significant increase in the teachers' maximum performance for maintaining task demands during the task launching at Timepoint 4 when compared to Timepoint 2 ($W_{(3)}=0$, $z=-1.73$, $p<.10$).[33] Similarly, the overall maximum performance during task launching at Timepoint 4 was significantly higher than at Timepoint 1 ($W_{(2)}=0$, $z=-1.73$, $p<.10$) and Timepoint 2 ($W_{(3)}=0$, $z=-1.34$, $p<.10$).

This peak signifies a moment of exemplary practice in the task launching phase during Timepoint 4, warranting further investigation into the factors contributing to this success to inform sustained teacher development. Some factors could be the accumulated experience with these practices, the impact of PLD, or a reflection of a particular focus in the curriculum or teaching cycle at that point. However, without sustained improvement at the fifth timepoint, it raises questions about the consistency and durability of these practices.

In the phase of student autonomous work (see Figure 20b), teachers' mean and maximum performance was higher than 1.50 and often slightly higher than 2.00. Notably, teachers' *mean performance* was rather *inconsistent*. The mean scores for providing mathematical prompts (SAW1-mean) fluctuate across the five timepoints, with a modest rise at the fourth, indicating a brief period in which teachers effectively challenged students without oversimplifying their mathematical thinking. The mean scores of asking students to engage in mathematical reasoning and/or meaning-making activities (SAW2-mean) dipped initially (Timepoints 1-3) but improved by the

[33] Given the small sample size and the constraints outlined in Section 4.2.2, it was determined to include a discussion on differences deemed significant at an alpha level of .10. (Kim & Choi, 2021; Zimmerman, 2000).

fourth timepoint. However, SAW3-mean mean scores remained erratic, lacking the notable peak of SAW1-mean, implying inconsistent translation of teachers' encouragement for reasoning (SAW2-mean) into active student engagement (SAW3-mean).

Teachers' maximum performance in facilitating student engagement in mathematical reasoning and meaning-making activities (SAW3-max) showed a significant increase at Timepoint 3 compared to Timepoint 1 ($W_{(4)}=0$, $z=-2.00$, $p<.05$). Furthermore, in this phase, teachers demonstrated a significant improvement in their maximum overall performance at Timepoint 3 than at Timepoint 1 (SAW-overall-max, $W_{(4)}=0$, $z=-2.00$, $p<.05$; SAW-overall-mean, $W_{(4)}=0$, $z=-1.79$, $p<.10$). These performance differences in SAW3-max could be attributed to the discussions held around questioning as a technique to draw out student thinking in VCS3 (Timepoint 3).

The significant rise only in *maximum* scores between certain timepoints in both phases (i.e., TL and SAW) signified instances of enhanced teacher performance, suggesting both improvement and a potential high level of quality in key teaching aspects. Interestingly, most practices in TL and SAW reached or even exceeded a score of 2.50 on the 1-to-3 point scale at some timepoint, indicating that teachers demonstrated high performance at least once, yet, leaving some room for improvement. Furthermore, statistical differences could highlight *performance variability among teachers* over time, indicating that while some teachers were consistent, others achieved breakthroughs at certain moments—an issue we revisit in the next chapters.

The trends in whole-class interactions (see Figure 20c) reveal that plenary discussions and interactions experienced varying levels of success (notably in WCI2 and WCI4), without a definitive trend of steady improvement or decline. However, the practice of eliciting student reasoning (WCI1) saw a decline over timepoints. Meanwhile, the trend in teachers' ability to prompt students to compare or evaluate different solutions (WCI3) and the overall use of WCI practices showed some positive shifts up to Timepoint 4, without a clear, confirmable pattern of continuous improvement. This variability highlights the challenges in consistently engaging an entire class in high-level mathematical thinking, emphasizing the need to improve consistency. The fluctuations might have been influenced by factors such as the topic being taught, the class dynamics, or the specific practices employed by teachers at different timepoints.

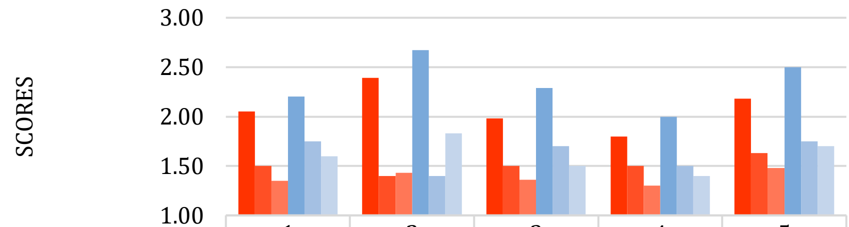
From the above, it can be concluded that the mean and maximum performance for several cognitive activation practices peaked at Timepoints 3 and 4. Initially, during Timepoint 1, teachers focused on integrating various practices introduced in video-club sessions. By Timepoint 4, as they delved deeper into these ideas through continuous discussion and application in the video-club sessions, there might have been a noticeable refinement in their teaching methods. This led to an enhancement in the quality of their experimentations. This improvement could be linked to the systematic organization of practices into a list of steps at Timepoint 4 (see Table 7, in Chapter 3), offering teachers a chance to explore and debate the concepts as a cohesive set of practices.

Finally, a slight drop in teacher performance from the first to the second timepoint was noted, albeit not significant and not surprising. Specifically, at the first timepoint, the focus was on identifying factors that could either facilitate or impede cognitive activation. By the second timepoint, the discussion shifted towards issues related to differentiation. Therefore, after VCS2 at Timepoint 2, it is plausible that teachers temporarily put aside issues of cognitive activation to explore and apply the newly discussed practices in differentiation.

4.3.2.2 Quality of Teachers' Experimentation with Differentiation Practices Across the Timepoints

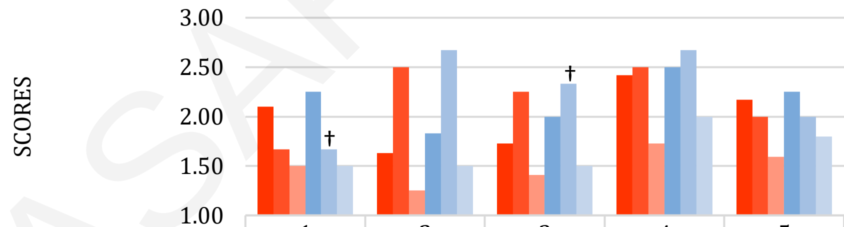
Figure 21 presents the mean and maximum performance of teachers in experimenting with differentiation practices. No clear patterns of progression were observed since teachers' performance was largely fluctuating across the timepoints. While the progression was not consistently upward at every timepoint, by the fifth timepoint, there was a slight increase in all practices related to task launching and student autonomous work, as well as in one aspect of whole-class interaction (WCI2) when compared to the initial timepoint. Overall, teachers scored lower on their experimentation with practices representing differentiation as opposed to the practices of cognitive activation, suggesting that more scaffolds seem to be warranted to improve the quality of their work on this front.

Figure 21. *Quality of Teachers' Experimentation with Differentiation Practices During (a) Task Launching; (b) Student Autonomous Work; and (c) Whole-class Interactions, Across the Five Timepoints.*



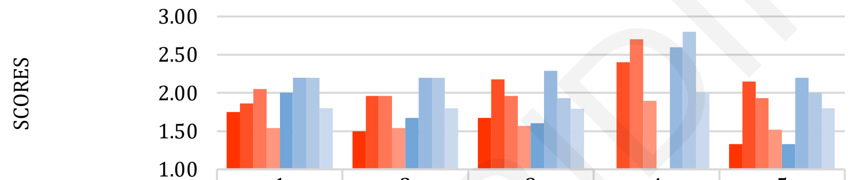
	1	2	3	4	5
TL1-mean	2.05	2.39	1.98	1.80	2.18
TL2-mean	1.50	1.40	1.50	1.50	1.63
TL-mean-overall	1.35	1.43	1.36	1.30	1.48
TL1-max	2.20	2.67	2.29	2.00	2.50
TL2-max	1.75	1.40	1.70	1.50	1.75
TL-overall-max	1.60	1.83	1.50	1.40	1.70

(a)



	1	2	3	4	5
SAW1-mean	2.10	1.63	1.73	2.42	2.17
SAW2-mean	1.67	2.50	2.25	2.50	2.00
SAW-mean-overall	1.50	1.25	1.41	1.73	1.59
SAW1-max	2.25	1.83	2.00	2.50	2.25
SAW2-max	1.67	2.67	2.33	2.67	2.00
SAW-overall-max	1.50	1.50	1.50	2.00	1.80

(b)



	1	2	3	4	5
WCI1-mean	1.75	1.50	1.67	1.00	1.33
WCI2-mean	1.86	1.96	2.18	2.40	2.15
WCI3-mean	2.05	1.96	1.96	2.70	1.93
WCI-overall-mean	1.54	1.54	1.57	1.90	1.52
WCI1-max	2.00	1.67	1.60	1.00	1.33
WCI2-max	2.20	2.20	2.29	2.60	2.20
WCI3-max	2.20	2.20	1.93	2.80	2.00
WCI-overall-max	1.80	1.80	1.79	2.00	1.80

(c)

TL1-Selecting tasks which are potentially accessible to all students; **TL2**-Making clear the organizational decisions or management procedures for working autonomously on the task; **TL-overall**-Holistic estimate of the differentiation during this phase; **SAW1**-Using student asynchronous work to accommodate different learning readiness levels and needs; **SAW2**-Encouraging multiple expressions of content, process, and/or product; **SAW-overall**-Holistic estimate of the differentiation during this phase; **WCI1**-Sequencing student solutions in a reasonable progression to support student access to the ideas shared; **WCI2**-Highlighting important mathematical ideas during the sharing to ensure that these are made clear to as many students as possible; **WCI3**-Students express mathematical ideas that are visible and/or audible to all students (as well as the teacher); **WCI-overall**-Holistic estimate of the differentiation during this phase.

Same color * signifies statistical significance at .05 level, and † at .1 level, between two timepoints.

The presentation of the results commences with the student autonomous work phase, which stood out as the sole phase in which statistically significant differences between timepoints were discerned. At first glance, Figure 21b suggests that while there were instances of high performance across all practices in student autonomous work, peaking notably at Timepoint 4, there was considerable variability across timepoints. Despite the indications for some progressive trends, only teachers' maximum performance on encouraging multiple expressions of content, process, and product (SAW2-max) demonstrated a significant improvement at Timepoint 3 compared to Timepoint 1 ($W_{(2)}=0$, $z=-1.73$, $p<.10$). However, the scores suggest that the consistency of differentiation practices during student autonomous work was an area for improvement.

Figure 21a reflects a non-linear fluctuating progression in teachers' differentiation practices during task launching, with moments of higher performance interspersed with periods of decline, before an eventual rebound in quality towards the end of the observed period (i.e., Timepoint 5). *Timepoint 2* appears to be a strong moment for the accessibility of tasks (TL1) and the overall performance in task launching (TL-overall), with both mean and maximum scores being at their peak. Timepoint 4 represents the lowest overall maximum scores, suggesting that this timepoint was particularly challenging for achieving high-quality in these practices. In VCS2, issues of differentiation were prominently addressed, whereas in VCS5 and VCS6, the focus shifted towards exploring the interplay between cognitive activation and differentiation, potentially leading teachers to sideline differentiation concerns and revealing that their learning journey was not straightforward.

The whole-class interactions phase shows variability across the different practices and timepoints (see Figure 21c). A notable trend is the peak in both mean and maximum scores for highlighting important mathematical ideas (WCI2) and having student expressions visible/audible to all (WCI3) at Timepoint 4, suggesting a period of heightened focus and possibly effective practice in these areas. However, the dips at Timepoint 4 for WCI1 indicate a challenge in sequencing student solutions effectively during this time. The overall data suggest that while there are points of high performance in whole-class interactions, maintaining consistent high-quality differentiation practices remained a challenge.

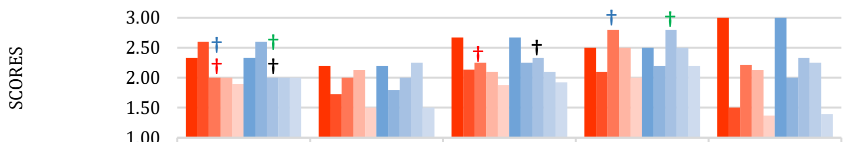
In all, the data suggests a peak in effective differentiation practices during the middle timepoints, which might indicate a learning curve that improved with experience before potentially facing new challenges or a decline in some aspects. It is important to note that the highest teacher scores (maximum scores) were scattered across different timepoints, reflecting the individual journeys and peaks in teacher performance.

4.3.2.3 Quality of Teachers' Experimentation with Practices that Promote the Interplay of Cognitive Activation and Differentiation Across the Timepoints

Figure 22 gives an overview of the teaching quality of teachers' experimentation with practices that represent the interplay of cognitive activation and differentiation across the five timepoints. The analysis demonstrated more statistically significant shifts in the implementation of practices that enhance the interplay between cognitive activation and differentiation at different timepoints of the program, compared to those observed in the implementation of practices related to each axis independently. Significant differences were most evident during the student autonomous work phase (which is presented first), with fewer notable changes occurring in task launching, and the least in whole-class interactions.

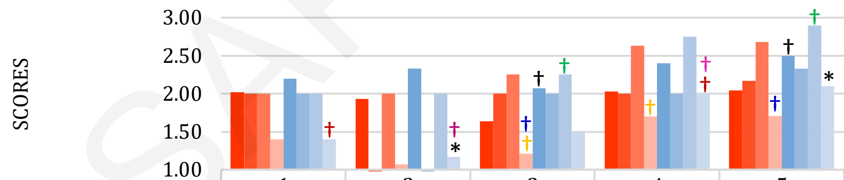
The findings suggest that teachers were supported more in their work with the students during the autonomous work phase, as demonstrated by the notable increase in statistically significant results observed in this phase (see Figure 22b). The increase in the overall maximum performance (SAW-interplay-max) highlighted this trend, with significantly higher scores observed at Timepoint 5 ($W_{(3)}=0$, $z=-2.00$, $p<.05$) relative to Timepoint 2. Moreover, Timepoint 4 also saw a significant improvement in performance compared to both Timepoint 1 ($W_{(2)}=0$, $z=-1.73$, $p<.10$) and Timepoint 2 ($W_{(3)}=0$, $z=-1.73$, $p<.10$), though with a lower level of statistical significance. Similar trends were evident in the holistic mean score of this phase (SAW-overall-mean) in which teachers showed significant improvement at the last two timepoints ($W_{(3)}=0$, $z=-1.83$, $p<.10$ for Timepoint 4 and $W_{(3)}=0$, $z=-1.84$, $p<.10$ for Timepoint 5) compared to Timepoint 3.

Figure 22. *Quality of Teachers' Experimentation with Practices that Promote the Interplay of Both Cognitive Activation and Differentiation During (a) Task Launching; (b) Student Autonomous Work; and (c) Whole-class Interactions.*



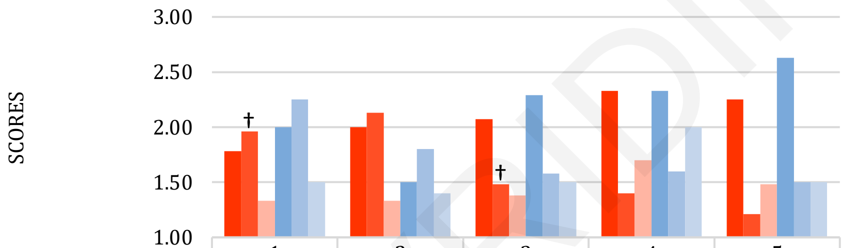
	1	2	3	4	5
TL1-mean	2.33	2.20	2.67	2.50	3.00
TL2-mean	2.60	1.73	2.14	2.10	1.50
TL3-mean	2.00	2.00	2.25	2.80	2.22
TL4-mean	2.00	2.13	2.10	2.50	2.13
TL-overall-mean	1.90	1.50	1.88	2.00	1.37
TL1-max	2.33	2.20	2.67	2.50	3.00
TL2-max	2.60	1.80	2.25	2.20	2.00
TL3-max	2.00	2.00	2.33	2.80	2.33
TL4-max	2.00	2.25	2.10	2.50	2.25
TL-overall-max	2.00	1.50	1.92	2.20	1.40

(a)



	1	2	3	4	5
SAW1-mean	2.02	1.93	1.64	2.03	2.04
SAW2-mean	2.00	0.00	2.00	2.00	2.17
SAW3-mean	2.00	2.00	2.25	2.63	2.68
SAW-overall-mean	1.40	1.07	1.21	1.70	1.71
SAW1-max	2.20	2.33	2.07	2.40	2.50
SAW2-max	2.00	0.00	2.00	2.00	2.33
SAW3-max	2.00	2.00	2.25	2.75	2.90
SAW-overall-max	1.40	1.17	1.50	2.00	2.10

(b)



	1	2	3	4	5
WCI1-mean	1.78	2.00	2.07	2.33	2.25
WCI2-mean	1.96	2.13	1.48	1.40	1.21
WCI-overall-mean	1.33	1.33	1.38	1.70	1.48
WCI1-max	2.00	1.50	2.29	2.33	2.63
WCI2-max	2.25	1.80	1.58	1.60	1.50
WCI-overall-max	1.50	1.40	1.50	2.00	1.50

(c)

TL1-Explaining potentially unfamiliar non-mathematical aspects of the wording of the task; **TL2**-Clarifying mathematical aspects of the task; **TL3**-Posing questions that indicate the level of support that students need in order to engage in the task; **TL4**-Activating relevant existing mathematical knowledge and strategies; **TL-overall**-Holistic estimate of the differentiation during this phase; **SAW1**-Directing different types of questions to different students for engaging them in meaning-making, conceptual connections, or mathematical reasoning; **SAW2**-Providing enablers to facilitate access to the task at hand without reducing the challenge; **SAW3**-Providing extenders to advanced learners or early finishers; **SAW-overall**-Holistic estimate of the differentiation during this phase; **WCI1**-Holding students accountable for attending to and understanding their classmates' sharing; **WCI2**-Using incorrect or incomplete student solutions as resources for all student learning; **WCI-overall**-Holistic estimate of the differentiation during this phase.

Same color * signifies statistical significance at .05 level, and † at .1 level, between two timepoints.

The trend towards increased maximum performance in employing differentiated questions (SAW1-max) is encouraging, as it suggests teachers were improving their ability to engage all students in meaningful mathematical discussions. Notably, the quality of their questioning significantly improved by Timepoint 5 in comparison to Timepoint 3 ($W_{(3)}=0$, $z=-1.73$, $p<.10$), a development that is particularly promising considering the emphasis placed on this aspect of teaching during the video-club sessions (i.e., VCS3 and VCS4).

Similarly, the increasing trend of using extenders (SAW3-max) demonstrated that teachers were effectively meeting the needs of advanced learners by offering them more opportunities to deepen their understanding and further explore concepts. This was evidenced by a significant rise in their maximum performance on this aspect by Timepoint 5 compared to Timepoint 3 ($W_{(2)}=0$, $z=-1.73$, $p<.10$). This progression aligned with the observed trend in teachers' expanding use of extenders as illustrated in Figure 22b, a pattern not mirrored in their use of enablers, which remained consistently around the value of 2.00. A potential reason for the gap between the heightened experimentation with enablers and the diminished quality observed might be the increased challenge teachers faced in crafting appropriate enablers for their students. This situation was echoed in the qualitative findings from teacher reports (see Chapters 5, 6, and 7). Teachers highlighted that the concept of using both enablers and extenders was a completely new idea introduced by the project, which they integrated into their teaching. This indicates a learning curve and an adaptation process as teachers familiarize themselves with and implement these innovative practices in their classrooms.

In the task launching phase (Figure 22a), an upward trend was generally observed across most practices over the timepoints, with the exception of clarifying mathematical aspects of the task (TL2), which demonstrated a decline throughout the PLD program. Teachers showed stronger performance in explaining potentially unfamiliar non-mathematical aspects (TL1), as indicated by the higher mean scores. Both posing questions to gauge the level of support students needed for engagement (TL3) and activating relevant existing mathematical knowledge and strategies (TL4) reached their highest point at the (third and more at the) fourth timepoint. The Wilcoxon signed-ranks test confirmed that teachers significantly improved in eliciting the knowledge necessary for students to work on tasks at Timepoint 3 (TL3-mean, $W_{(3)}=0$, $z=-1.73$, $p<.10$; TL3-max, $W_{(3)}=0$, $z=-1.73$, $p<.10$) and Timepoint 4 (TL3-mean,

$W_{(2)}=0$, $z=-1.73$, $p<.10$; TL3-max, $W_{(2)}=0$, $z=-1.73$, $p<.10$), as evidenced by both their mean and maximum performance metrics, in comparison to Timepoint 1. This result is promising, particularly because the nature of the questions—tailored to appropriately challenge each student and to concentrate on key mathematical concepts—was addressed during VCS3 and VCS4. Questions designed to collect insights into students' knowledge, solutions, strategies, and challenges served as exemplars during the sessions. Other fluctuations noted, such as those at the fifth timepoint, were considered minor and could be attributed to factors such as the nature of the tasks assigned, or the specific content covered in the lessons.

Throughout the whole-class interactions (see Figure 22c), a consistent yet non-significant trend of steady improvement was noted in teachers' efforts to ensure students were attentive to and comprehended their peers' contributions (WCI1). Conversely, the approach to utilizing incorrect or incomplete solutions as educational opportunities (WCI2) demonstrated a declining trend, indicating a potential gap in transforming these moments into constructive learning experiences for all students. This decline in teachers' performance, as reflected in both mean and maximum scores (WCI2-mean and -max), was particularly notable. Notably, the decrease in the mean performance for this practice from Timepoint 1 to Timepoint 3 was statistically significant ($W_{(3)}=0$, $z=-1.83$, $p<.10$).

4.3.2.4 Summarizing Key Findings on the Quality of Teachers' Performance Across the Timepoints

The findings illustrated that teachers experimented with the concepts discussed in the video-club sessions and captured by the coding protocol with varying degrees of frequency in use and in quality across different practices. The Wilcoxon signed-rank tests showed statistically significant improvements in about one-third of the phase-level practices, particularly from the third timepoint onwards. This improvement was most notable in practices that combine cognitive activation with differentiation, rather than in practices focusing on each axis separately. Looking across phases, significant growth was observed mostly in practices pertaining to the interplay of cognitive activation and differentiation (five practices)[34], followed by those that emphasize cognitive activation alone (four practices), and to a lesser extent, differentiation

[34] Practices showing statistically significant changes in both mean and maximum scores are counted only once.

practices (one practice). Of these practices, six pertained to practices related to the phase of student autonomous work. Notably, teachers' performance in practices related to student autonomous work exhibited the most significant changes, especially in practices that promote the interplay of both axes (three practices).

Certain practices were distinguished by statistically significant improvements in both mean and maximum performance throughout the PLD program. These include the *use of extenders* and *differentiated questioning* during student autonomous work, as well as *maintaining cognitive challenge* and *identifying student support needs* during task launching. These practices were highlighted in the PLD program, particularly in the sessions associated with Timepoints 3 and 4, and were positively received by teachers, as indicated in their end-of-program interviews (see Chapters 5, 6, and 7). This suggests a collective improvement in areas emphasized during the program, despite the complexity of the concepts and practices introduced.

However, significant improvements were limited to one differentiation practice and one whole-class interaction practice, indicating areas that require further qualitative investigation to understand the underlying challenges. The most significant improvements were observed between the final two timepoints and earlier ones, suggesting that organizing the practices into a dynamic list of steps for lesson planning in the fourth timepoint may have aided teachers in enhancing their implementation.

Additionally, the analysis of the quality of teachers' experimentation, alongside the frequency, revealed that more frequent practice implementation was *not* necessarily tantamount to higher quality. Factors beyond mere experimentation possibly contributed to the observed improvements, warranting further exploration. The data analysis, encompassing both mean and maximum scores, provided a multifaceted view of the teachers' development, showing more significant growth in maximum performance (13 differences) compared to mean performance (six differences). This indicates some progress, although there is still room for enhancement.

The improvements observed in the quality of teachers' experimentation are noteworthy, especially considering the limitations of the study—only four lessons per teacher across eight teachers, the six-month duration of the PLD program totaling 22.5 hours, and the absence of baseline or follow-up lesson recordings. To delve deeper, a more detailed qualitative examination of the teachers' practices was essential. This deeper exploration is the focus of the analysis presented in the forthcoming chapters.

To conclude, based on the trends observed in teachers' collective performance across the timepoints, they did benefit from the PLD program. Key indicators of benefit include peak performances, especially notable at Timepoint 4; statistically significant improvements; and the introduction of new practices, such as enablers and extenders. However, the benefits were not uniformly sustained or consistent across all practices or timepoints, indicating areas where further support and development might be needed. The observed fluctuations and the decline in certain practices by Timepoint 5 suggest that achieving lasting change in teaching practices may require ongoing support and engagement beyond the scope of the program. Also, the variability in performance may suggest that while the program was beneficial on a group level, individual experiences and outcomes may vary, pointing to the need for exploring individual teachers within the collective framework.

4.4 Shifting from Collective Insights to Individual Stories

The different delineations of *Reactions*, *Results*, and *Behavioral* evolution of the teachers as a group provided a comprehensive and complementary overview of the PLD program's effectiveness. Specifically, immediate teachers' perceptions captured through reactions underlined the program's strengths and areas needing refinement. Together, the levels of Results and Behavior showed mixed outcomes rather than a clear-cut success or failure, suggesting that while the PLD program has led to significant pedagogical advancements, it also highlighted areas in which further development was crucial. Results highlighted the necessity for *continued* support, particularly in differentiation. Behavioral analysis over time emphasized the importance of *responsive* PLD to meet evolving needs, illustrating that increased frequency in practice implementation does not always equate to higher quality.

Having outlined the *collective* patterns and outcomes from the examination of the teacher group participating in the PLD program, the subsequent chapters focus on the individual journeys of three selected cases: Pina, Kate, and Michelle. While the collective results offer a panoramic view of the program's impact on teachers' experimentation and teaching performance, the detailed stories of these specific cases can deepen our understanding of the complexity and diversity of diverse teacher development trajectories.

As explained in Chapter 3 (see Section 3.3), the selected cases were heterogeneous in various characteristics, while also differing in the quantitative delineation of their lessons across the timepoints, as is presented in their results and behavioral patterns in the coming chapters. Table 11 outlines the content of the four videotaped lessons taught by each of the selected teachers over the course of the PLD program.

Table 11

Content Overview of the Videotaped Lessons for Each Case Study

Pina	Kate	Michelle
L1 (occurring after VCS1): Formulating and applying the divisibility criteria of 2, 5, and 10, and investigating their inter-relationship	**L1** (occurring after VCS2): Identifying, estimating, and constructing complementary and supplementary angles	**L1** (occurring after VCS1): Sequencing negative numbers and understanding their value and use in daily life
L2 (occurring after VCS2): Developing generalization of an algebraic pattern and finding the value of any given term	**L2** (occurring after VCS6): Investigating and identifying a general rule for calculating the sum of the interior angles of a polygon	**L2** (occurring after VCS2): Investigating the associative and commutative properties of addition and using them to do mental calculations
L3 (occurring after VCS4): Classifying different types of triangles according to the measure of their angles	**L3** (occurring after VCS7): Multiplying a fraction by an integer, and vice versa	**L3** (occurring after VCS3): Performing long divisions with two-digit divisors based on the distributive property
L4 (occurring after VCS6): Identifying the relationship between the multiplication patterns of 2, 4, and 8	**L4** (occurring after VCS8): Estimating and investigating ways of multiplying integers with mixed numbers	**L4** (occurring after VCS9): Using proportional reasoning to compare and represent different ratios

Notes.
1. All teacher names are pseudonyms.
2. L1: Lesson 1, L2: Lesson 2, L3: Lesson 3, L4: Lesson 4; VCS: video-club session (see their content in Table 7).

The content of teachers' lessons showcases a diverse range of mathematical concepts from various areas of mathematics, including numbers, operations, geometry, algebraic ideas, ratios, and proportional reasoning. As discussed in Chapter 3, teachers had the autonomy to select and teach any mathematical content they wanted for videotaping, based on their curriculum and personal preferences, without the pressure to cover specific or the same topics/areas. This approach aimed to prevent overwhelming teachers, which could lead to disengagement and additional stress.

The findings of the selected cases are presented in the coming chapters. For each case study, a corresponding separate chapter has been created, containing the findings of each teacher's data analyses. Specifically, Chapter 5 explores how Pina represents a relatively successful case of continuous improvement and a teacher who viewed and practiced cognitive activation and differentiation synergistically. Chapter 6 delves into Kate's experiences, highlighting the challenges of maintaining consistency and quality when juggling multiple new practices. Chapter 7 focuses on Michelle's story, shedding light on the complexity of teacher-student interactions and illustrating how well-intentioned support on the teacher's part can inadvertently hinder student thinking. Each case illustrates distinct pathways in their teaching and learning within the same program. As such, the next three chapters complement the evaluation of the program as a collective undertaken in this chapter.

CHAPTER 5. THE CASE OF PINA: A SYMPHONY OF GROWTH[35]

Pina's journey through the PLD program illustrates a steady progression in her teaching across all axes, with notable growth in cognitive activation and its interplay with differentiation. She experimented with a wide array of practices, aligned with the video-club session emphases, while gradually polishing their implementation. Contrary to the second case (i.e., Kate, see Chapter 6) Pina adeptly layered new practices atop existing ones, without abandoning one practice for another. At the same time, her conceptual evolution was characterized by a deliberate shift from treating the axes separately to adopting a more synergistic approach. Although identified as a success within the program, Pina encountered challenges with specific concepts and practices, some of which she overcame, while others remained even at the end of the program. Her case highlights the importance of ongoing experimentation in teaching and the potential of PLD programs to support such development.

5.1 Pina's Background

Pina was an experienced teacher with fifteen years of teaching experience, including a long career in lower grades (12 out of 15 years). During the PLD program, she was a second-year third-grade teacher serving in a school, which was attended by many children with immigrant backgrounds. Her educational background included a bachelor's degree in elementary education and a master's degree in mathematics education.

The focus of the EDUCATE PLD program as it was advertised sounded very interesting to Pina (EPI, line 24). Moreover, her earlier positive experience with videotaping herself while teaching mathematics in another research program conducted by one of the teacher educators (focusing on issues of cognitive activation and challenging tasks) made her more open to joining the current program and being videotaped again:

Pina: We worked with [teacher educator's name] in a previous research program during which I was videotaped multiple times while teaching… Having experience with being videotaped, I thought "Why not? This [PLD program] will be [as] good [as my prior experience]".

Interviewer: Were you afraid of being videotaped again?

Pina: No, I was not. Of course, you always feel weird when you first see yourself teaching but you learn a lot of things by observing yourself. (EPI, lines 24-33)

[35] The title's metaphorical parallelism with a musical symphony underscores the harmonious integration of various teaching practices in Pina's teaching practice, akin to the coordinated expression found in orchestral music.

As it turned out, because of her previous experience, Pina considered the use of videos from her teaching as a valuable source for her learning because it could offer her a window into her actual teaching practice (lines 6-7). Despite already being an experienced teacher, she further explained that by analyzing and reflecting on her teaching and that of her colleagues during the current program, she got involved in a process of experimenting; learning new ideas; identifying and correcting teaching "errors"; trying alternative approaches and practices; and even noticing things she had never thought of, such as the volume of her voice (EPI, lines 53-58; 71-75).

Pina entered the program believing that not all ideas would work perfectly in her lesson, without this discouraging her from experimenting with them (EPI, lines 63-66). This indicates that she had adopted a more interpretive rather than an evaluative stance towards her teaching, an attitude that could be supportive when focusing on and experimenting with certain practices. This skill proved to be helpful along the way in the PLD program, as will be explained later.

5.2 Reaction Level: Pina's Evaluation of the EDUCATE PLD Program

Pina had an especially favorable attitude towards the PLD program, identifying several positive aspects of it that contributed to her learning. Firstly, she appreciated its interactive nature, since participants were actively engaged in all program activities and discussions (EPI, lines 38-40). Satisfied with this aspect, she stated that she "always left the video-club sessions quite concerned—not in the negative sense—about what I can do better or how I can do something differently" (EPI, lines 76-78).

Secondly, she valued the opportunity to put into practice various ideas she had learned (lines 71-72). Specifically, experimenting, watching, and discussing her teaching supported her in noticing, interpreting, and improving her work of teaching: "What matters is experimenting with new ideas and noticing a difference in your teaching… to see if and how a practice works… And if it doesn't work, consider what and how your practice can change to see improvements" (EPI, lines 66-69, emphasis added).

Thirdly, despite the meticulous organization behind the PLD program, she enjoyed that there was also flexibility (EPI, lines 675-679). Specifically, at the end of each session, participants were asked to reflect on what they had learned and what needs they still had related to the topics under discussion, allowing subsequent sessions to focus on these (EPI, lines 679-682; 684-688). She mentioned:

> The program wasn't just about the teacher educators presenting ideas. […] They respected us and took into consideration our opinions and needs, as well as what we wanted to pursue or what topics we wanted to delve deeper into. Not being static was very important for me. There was flexibility. (EPI, lines 682-683; 687-691)

Furthermore, she claimed that the small group size facilitated personal interactions among teacher participants and teacher educators, helped streamline the planning of the videotaped lessons, and even allowed for receiving individual feedback by the teacher educators via phone calls (EPI, lines 691-693). "It was very important to have someone to tell me some things about my lesson, either positive or negative, especially what didn't work so well," she stated, once again highlighting her growth mindset (EPI, lines 694-697). The small group size was also beneficial in creating a "learning community" where participants "observed and discussed each other's lessons" (EPI, lines 871-876; 697-699). She felt that everyone could speak comfortably and share their concerns or aspects of their work that did not work well in their lessons, without feeling embarrassed or hesitant (EPI, lines 721-728) or being judged (EPI, lines 697-699). The following phrase epitomizes her impression of the program: "The atmosphere of the sessions was very friendly; it felt like a family. You don't encounter it often." (EPI, lines 933-935)

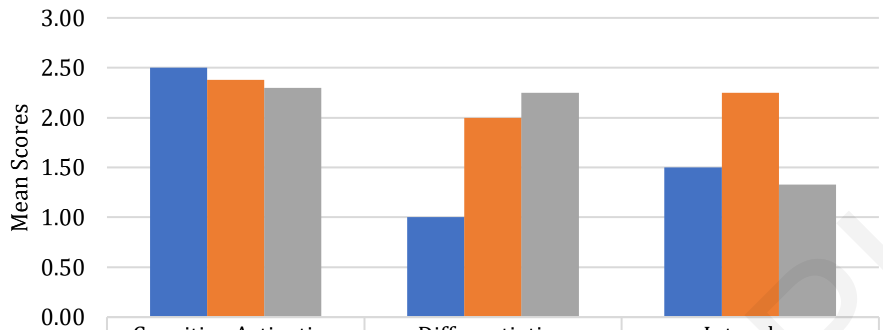
Lastly, the extended duration of the program was essential because "there was a time span from one meeting to the next for videotaping the lessons and reflecting on them. It wasn't five meetings, one right after the other" (lines 703-709). This point emphasizes not only the need for multiple PLD meetings but also the significance of spreading them over time.

5.3 Results Level: Pina's Teaching Performance in Her Concluding Lesson

To examine the teaching performance in Pina's concluding lesson, the aggregated mean scores of each phase per axis of that lesson were used (see Section 3.7, in Chapter 3). Figure 23 illustrates Pina's teaching performance in cognitive activation, differentiation, and the interplay of the two axes in that particular lesson, across the three lesson phases.

Figure 23

Pina's Performance in Cognitive Activation, Differentiation, and Their Interplay in Each Phase of Her Concluding Lesson



	Cognitive Activation	Differentiation	Interplay
TL	2.50	1.00	1.50
SAW	2.38	2.00	2.25
WCI	2.30	2.25	1.33

Note.
TL: Task Launching; SAW: Student Autonomous Work; WCI: Whole-Class Interactions.

The graph depicts a case of a teacher whose performance was overall moderate to high. Pina's scores in cognitive activation ranged from 2.30 to 2.50 across all phases, which were relatively close to the top of the scale (i.e., 3.00). Although she was slightly more effective during the task-launching phase, the differences between phases were minor. These performances were indicative of a teacher who prioritized and was successful in maintaining a relatively high level of mathematical challenge during the lesson.

Considering differentiation alone, Pina's performance could be characterized as varying from low to moderate depending on the lesson phase, with her scores ranging from 1.00 to 2.25. A score of 1.00 during task launching, which was the lowest across all axes and phases, indicates a low performance in addressing student differences during this lesson phase. However, the scores doubled to 2.00 and 2.25 during student autonomous work and whole-class interactions, respectively. This suggests she was better at engaging with students individually, in groups, or plenary after they had been working autonomously.

Her performance in the interplay between the two axes fluctuated across different task phases. Pina showed a better ability to integrate these two components during student autonomous work (scoring 2.25). Still, there was more room for improvement during the other phases (with scores ranging from 1.33 to 1.50). This implies that Pina was more effectively balancing the cognitive demands of the tasks with students' individual needs during their autonomous work.

In all, Pina ended up being a relatively successful case of the PLD program who was particularly adept at engaging students in challenging work across all three lesson phases and could adjust her teaching to meet diverse needs, especially during student autonomous work and whole-class interactions. Her ability to integrate both axes appeared to be more effective during student autonomous work. Her scores in the concluding lesson points to the need to develop her skills in the interplay of both axes during task launching and whole-class interactions, and her differentiation skills during task launching. However, without knowing her baseline scores or any other scores from other timepoints of the PLD program, it is challenging to measure her exact benefit. Consequently, an in-depth examination of Pina's teaching practice *progression* was conducted across her four videotaped lessons, which is presented next.

5.4 Behavior Level: Evolution of Pina's Teaching Performance

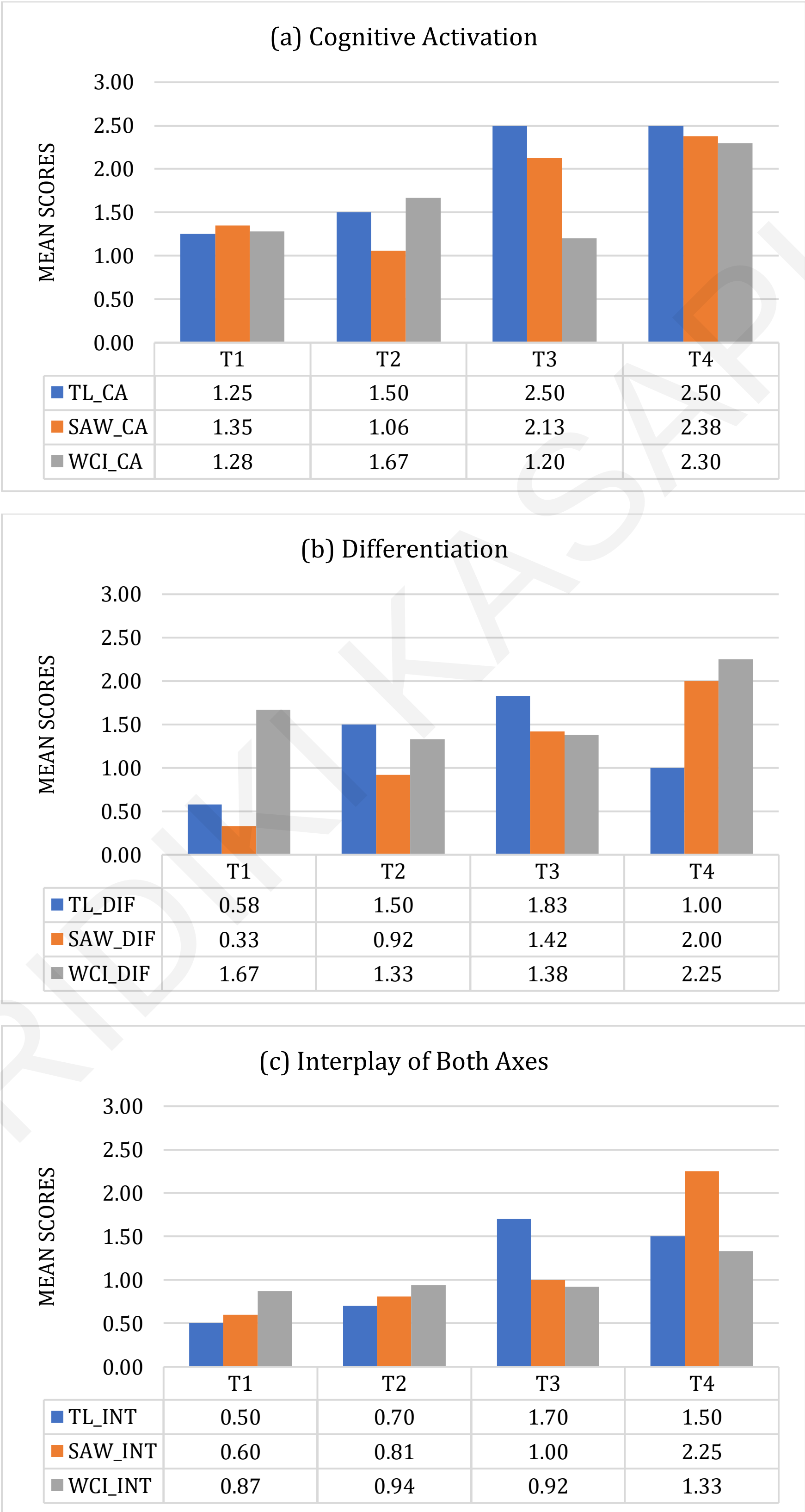
This section presents the evolution of Pina's teaching performance, as illuminated by both quantitative and qualitative analyses. The aggregated mean scores of each lesson in each phase per axis showcased a promising upward trajectory in her teaching, particularly in cognitive activation and its interplay with differentiation. The qualitative analysis suggested consistency and quality of Pina's experimentation with respect to these pedagogical aspects, highlighting her continual growth and the challenges she encountered.

5.4.1 Quantitative Analysis

Figure 24 represents these mean scores of all Pina's videotaped lessons across the teaching phases within each axis.

Figure 24

Pina's Performance in (A) Cognitive Activation, (B) Differentiation, and (C) Their Interplay, Per Phase Across Her Lessons



Note.
TL: Task Launching; SAW: Student Autonomous Work; WCI: Whole-Class Interactions.

What stands out in this figure is the upward trajectory in her mean performance in most respects, which reached much above the scale median (1.50) after her last lesson. The greatest growth was observed in her mean performance in cognitive activation which increased dramatically from below the median of the scale in her first lesson (almost 1.50, see Figure 24a) to much higher than the median in her final lesson (around 2.50, see Figure 24a). Also, her mean performance in the interplay of cognitive activation and differentiation rose steadily across the four lessons, shifting from much below the median of the scale (around 1.00, see Figure 24c) to near or above the median (between 1.00 and 2.50, see Figure 24c)—yet, not reaching the scores of the mean performance of cognitive activation in the phases of task launching and whole-class interactions (see Figure 24a). Additionally, a noticeable increase was observed in differentiation, with remarkable growth in her mean performance in the phase of student autonomous work when comparing her first and last lessons (from around 0.50 to over 2.00, see Figure 24b). Some drops were observed (e.g., in cognitive activation during whole-class interactions from Timepoint 2 to Timepoint 3, or in differentiation during task launching from Timepoint 3 to Timepoint 4) but they did not deviate dramatically from the overall trend of growth noticed across the four lessons.

Overall, the quantitative analysis suggests that despite still having some room for improvement in all three axes, Pina was a case of a teacher who steadily solidified and improved practices that promoted the dual goal of cognitive activation and differentiation with which she experimented across her four lessons. The upward trend in almost all phases within each axis across the lessons indicates that Pina had benefitted from the PLD program. Nevertheless, this analysis falls short of elucidating the specific methodologies and practices in which enhancements were observed, nor does it shed light on areas in which potentially divergent patterns may have emerged. The following section transitions from this surface-level evaluation of Pina's performance towards a deeper, more nuanced qualitative exploration of her teaching progression.

5.4.2 Qualitative Analysis

The results obtained from the qualitative analysis of Pina's lessons are presented in Table 12. As the checkmarks in Table 12 imply, Pina experimented with multiple ideas discussed in the PLD program, especially in the phases of task launching and whole class discussion. Specifically, they illustrate that the number of practices with which she experimented slightly increased from the first to the second or third lesson. Then, she was persistently experimenting with most practices until her concluding lesson.

Congruent with the quantitative results (see Figure 24), a clear trend of constant improvement regarding the quality of her experimentation across the four lessons is observed. In particular, the shading of the cells in Table 12, shifting from light green or red to bright green for most practices from the second or third lesson and onwards, shows a general pattern of consistency in the frequency and quality of Pina's experimentation. Remarkably, the greatest improvement is identified in her third and fourth lessons, since most cells have a bright green color. Hence, judging by the continual growth in her teaching practice, Pina seems to be a relatively successful case of the PLD program, improved in various respects, especially in cognitive activation and its interplay with differentiation.

Despite the steady progress in her teaching practice, the experimentation with some practices slightly differs from the pattern described above. For instance, some practices were sporadically implemented across the four lessons principally in medium or high quality (including, handling unexpected student solutions; handling unexpected student interference which could probably steal the thinking; handling alternative conceptions around mathematical ideas; and using flexible grouping, see Table 12). These practices were related to the teacher's handling of *unexpected* classroom events and student responses/reactions, which did not always arise in every lesson, and hence, reasonably, they did not follow the steady increase pattern.

All in all, Table 12 suggests that Pina experienced a steady increase and great gains from the PLD program, while also facing some challenges (some of them can be observed from the check marks in the grey-colored practices in Table 12). The steady increase pattern was more discernible for cognitive activation and its interplay with differentiation. The qualitative analysis of Pina's lessons and the individual themes that emerged are detailed next.

Table 12

Pina's Experimentation with Practices That Promote (A) Cognitive Activation, (B) Differentiation, and (C) Their Interplay, Across Her Lessons

A. COGNITIVE ACTIVATION				
TASK LAUNCHING	**L1**	**L2**	**L3**	**L4**
• Selecting mathematically challenging tasks	✓	✓	✓	✓
• Maintaining the cognitive demands of the task as presented to students during task launching	✓	✓	✓	✓
• Asking students to read the task instructions and implement the think-pair-share strategy	✓			
• Discussing (key) mathematical ideas (to the task and/or the goal of the lesson) without reducing the demands		✓	✓	✓
• Asking students to explain/make sense of mathematical symbols, sentences, or representations		✓	✓	✓
• Handling unexpected student interference which could probably steal the thinking		✓	✓	✓
• Relaunching the task ending up initiating a whole-class discussion				✓
STUDENT AUTONOMOUS WORK	**L1**	**L2**	**L3**	**L4**
• Providing mathematical prompts to students to help them make some progress on the task and take up the challenge	✓	✓	✓	✓
• Pressing students for explanation/meaning, for making conceptual connections, or engaging in mathematical reasoning	✓	✓	✓	✓
• Posing mostly close-ended questions	✓			
• Monitoring students' work and being more directive than needed	✓	✓		
• Allowing another student to steal his pair's thinking				
• Asking for explanations that focus on describing the procedure used				
• Telling the students precisely *how* to work (step-by-step) on the task				
• Pointing out errors in students' work and remediating with procedures				
WHOLE-CLASS INTERACTIONS	**L1**	**L2**	**L3**	**L4**
• Eliciting or providing opportunities for student mathematical reasoning and meaning-making without reducing the challenge	✓	✓	✓	✓
• Rephrasing student ideas to address key mathematical ideas related to the task at hand (in interaction with students)	✓	✓	✓	✓
• Synthesizing and extending student contributions to address key mathematical ideas related to the task at hand (in interaction with students)	✓	✓	✓	✓
• Presenting and discussing multiple solutions	✓	✓	✓	✓
• Comparing and evaluating multiple solutions	✓	✓	✓	✓
• Bringing the class to the plenary at appropriate checkpoints		✓	✓	✓
• Having a clear direction during the discussion	✓	✓	✓	✓
• Extending the discussion by posing a question that is more challenging than the task-at-hand	✓	✓		
• Handling unexpected student solutions	✓		✓	
• Handling unexpected student interference which could probably steal the thinking		✓		
• Handling alternative conceptions around mathematical ideas			✓	
• Providing directive hints or ready-made answers	✓	✓		
• Enacting IRE interactions when checking in plenary	✓			
• Introducing important mathematical ideas very early				
B. DIFFERENTIATION				
TASK LAUNCHING	**L1**	**L2**	**L3**	**L4**
• Having the resources or materials available to be used by the students		✓	✓	✓
• Providing a clear way of working on the task	✓	✓	✓	✓

	L1	L2	L3	L4
• Providing materials to students and not ensuring that they have understood how they can be used while working on the task				
STUDENT AUTONOMOUS WORK	**L1**	**L2**	**L3**	**L4**
• Implementing asynchronous work	✓	✓	✓	✓
• Using learning aids				
• Using an entry card				
• Monitoring students' work and formatively assessing their needs		✓	✓	✓
• Using a tic-tac-toe board as a final assessment activity				
• Using an exit card	✓			
• Using anchoring activities	✓	✓		
• Encouraging multiple expressions of content, process, and/or product		✓	✓	✓
• Grouping students according to their proficiency levels				
WHOLE-CLASS INTERACTIONS	**L1**	**L2**	**L3**	**L4**
• Highlighting important mathematical ideas	✓	✓	✓	✓
• Sequencing student solutions in a reasonable progression		✓	✓	✓
• Allowing students to start explaining any method/solution they want		✓	✓	✓
C. THE INTERPLAY OF COGNITIVE ACTIVATION AND DIFFERENTIATION				
TASK LAUNCHING	**L1**	**L2**	**L3**	**L4**
• Explaining potentially unfamiliar non-mathematical aspects of the wording of the task or difficult words (context- or scenario-wise)	✓	✓	✓	✓
• Clarifying mathematical aspects of the task		✓	✓	✓
• Activating relevant existing mathematical knowledge and strategies			✓	✓
• Posing questions that indicate the level of support that students need in order to engage in the task without reducing the level of challenge			✓	✓
• Spending no time clarifying the task instructions during task launching				
STUDENT AUTONOMOUS WORK	**L1**	**L2**	**L3**	**L4**
• Directing different types of questions to different (groups) of students		✓	✓	✓
• Circulating within all the groups attempting to attend to all students			✓	
• Using enabler(s)				✓
• Using extender(s)			✓	✓
• Sharing a strategy devised by a student group with the rest of the class, to support them make progress on the task and taking up the challenge				
• Using flexible grouping			✓	
• Maintaining the demand for more advanced students and trivializing the thinking of less advanced students	✓	✓		
• Devoting more time to scaffolding students who are facing difficulties				✓
• Not establishing a routine for what the early finishers could do once they complete the main task				
• Facing difficulties in supporting less-advanced students to make progress on the task		✓		✓
• Directing the exact same questions to all (groups of) students				
WHOLE-CLASS INTERACTIONS	**L1**	**L2**	**L3**	**L4**
• Using incorrect or incomplete student solutions as resources for all students' learning	✓	✓	✓	✓
• Holding students accountable to attend to their classmates' thinking		✓	✓	✓

Notes.
1. L#: Lesson ordinal number
2. The check mark illustrates that the teacher experimented to some extent with the practice in that lesson.
3. The color scale demonstrates the quality and frequency of the implementation (see Section 3.7 in Chapter 3).
4. Bright green: the practice was implemented in high quality and the demand was not reduced; Light green: it was implemented in medium quality and the demand was somehow reduced; Red: it was implemented in low quality and the demand was reduced; Grey: this practice hinders the particular axis.

The in-depth qualitative analysis of the four lessons revealed two main themes, regarding the changes or lack thereof in Pina's teaching: (a) *steadily improving in promoting the dual goal of cognitive activation and differentiation and (b) struggling with certain ideas inherent in concurrently attending to both axes*. In what follows, the findings are discussed considering these two themes, accompanied by indicative examples.

Steadily improving in promoting the dual goal of cognitive activation and differentiation. The first theme pertains to the teacher's constant improvement in employing different practices that promote cognitive activation and its interplay with differentiation (and less differentiation alone) throughout the PLD program. Pina was observed to (a) *experiment with various ideas, following the video-club sessions' emphases,* while (b) *building on program ideas and gradually polishing certain practices.* The first subtheme illustrates that Pina was increasing the number of ideas with which she experimented over time (see Figure 14 in Chapter 3). The second subtheme concerns mostly the changes observed in the quality of Pina's experimentation over time. The two subthemes are further analyzed below.

Experimenting with various ideas following the video-club sessions' emphases. The first subtheme shows that in each lesson Pina deliberately experimented with practices that were the focus of the discussion in each previous video-club session, rather than randomly experimenting with multiple ideas or cherry-picking ideas. Specifically, her lessons were taught and videotaped after VCS1, VCS2, VCS4, and VCS6 (see Table 11 in Chapter 3). Unlike other teacher-participants who taught and videotaped their lessons at a timepoint that was more convenient for them or their school unit (e.g., as in the second case, Kate), Pina taught and videotaped a lesson focusing on what the aim of the prior meeting was, in between the sessions, remaining consistent to the PLD program structure and video-club emphases. In every subsequent session, videoclips of her lesson were shared, initiating a cycle of discussing, experimenting, and reflecting upon her practice combined with rich activities developed within the framework of the EDUCATE project.

Following the content and the unfolding of the video-club sessions (see Table 7 in Chapter 3), in L1, she mainly experimented with practices that promoted cognitive activation as was the focus of the discussions and the activities of VCS1 (see the number of checkmarks in cognitive activation practices compared to the practices from the other axes in L1 in Table 12). Then, in L2, she increased her experimentation with differentiation practices, following the content of VCS2 (see the increased number of checkmarks in differentiation practices in L2 compared to L1 in Table 12). In L3, which was videotaped after VCS4 (focusing on questioning as a means to achieve both cognitive activation and differentiation), growth in the experimentation with practices that promote the interplay of both axes was observed (see the increased number of checkmarks in practices that promote the interplay in L3 compared to earlier lessons, in Table 12). Finally, her fourth lesson included various practices from all three axes (see the number of checkmarks in L4 across all axes compared to her previous lessons, in Table 12). Her concluding lesson was videotaped after VCS6 in which the video-club group focused on generating a list of steps to be used during lesson planning aiming to help teachers deliberately design and teach lessons incorporating teaching practices that promote both axes. Perhaps the incremental introduction of certain Ideas during the PLD program allowed a logical progression from specific to more complicated aspects of the ideas/practices for Pina.

It is worth noting that although she increased her experimentation with practices from a certain axis in each lesson depending on the focus of the sessions, she did not terminate her experimentation with the practices she used in previous lessons, as evident in Table 12. For example, Pina started experimenting with presenting and discussing multiple solutions in L1 (axis of cognitive activation). In L2, she introduced the practice of sequencing student solutions in a reasonable progression (axis of differentiation), while continuing to experiment with the former practice, Then in L3, she activated relevant existing mathematical knowledge and strategies (axis of interplay), without ceasing her experimentation with previous practices. The next section further explores how Pina's experimentation with certain teaching practices had improved in quality over time.

Building on program ideas and gradually polishing certain practices. As can be seen from Table 12, Pina's experimentation could be described as spiraling[36] since she was revisiting the ideas/practices in each lesson, each time building on her prior learning (i.e., the discussions of the previous video-club sessions and experimentation) thus, further polishing her practice. Instead of experimenting with a practice and then abandoning it, Pina incrementally added them into her teaching and re-experimented with them in her next lessons, whereby the practices were built on to improve. This

[36] The idea of "spiral" is adapted from Bruner's (1960) cognitive theory to convey the idea of building upon previously discussed/learned ideas with deepening layers of complexity. In the case of Pina, she was coming back to the same practices, building on her previous experimentation.

subtheme was more pronounced in cognitive activation and its interplay with differentiation and not so much in differentiation in isolation. Two distinctive examples that highlight this specific pattern are presented and elaborated below.[37]

Polishing her skills in engaging students in mathematically challenging discussions. Throughout her lessons, Pina honed her techniques in questioning, scaffolding, and orchestrating whole-class discussions. In L1, which focused on the articulation and application of the divisibility rules for 2, 5, and 10, and their interconnectedness, Pina provided some opportunities for students to share their observations and explanations. Nevertheless, at various lesson junctures, she *was more prescriptive than perhaps needed or spoon-fed students some answers.* For instance, after the students had worked in pairs and made observations to deduce the conditions under which each bulb would light up (see questions a and b, Figure 25), the whole class convened for a plenary session to reflect on their findings regarding the green bulb:

Pina: When does the green bulb light up?
Student 1: [In numbers:] 2, 4, 6, 8, 10, 12, 14.
Pina: *What observations do you have about these numbers?*
Student 2: They are even numbers.
Student 3: They [i.e., even numbers] can be partitioned into two numbers, which are the same.
Pina: So, I can divide them by which number?
Student 4: Two.
Pina: *So, these are the multiples of which number?*
Student 4: 2 (1x2, 2x2, 3x2, 4x2…)
Pina: Which number would come after 14?
Student 5: 16, 18, 20, 22, 24.
Pina: *What do <u>even</u> numbers have? How do I distinguish them? <u>Think of their digits' value</u>.*
Student 6: They could have 2, 4, 6, 8.
Pina: You forgot one.
Student 6: Zero.
Pina: [repeats:] 0, 2, 4, 6, 8. *Any other observations from this table? Do you notice anything for some numbers in this table?*
Student 7: The odd numbers do not have a tick.
Pina: All of them? For which number no bulb is lit up?
Student 8: 3, 1, 7, 11, 13, 9.
Pina: Which are these numbers? How do I call all these numbers?
Student 9: Odd numbers.
Pina: Which is missing?
Student 9: [Number] 5
Pina: *No bulb is lighting up in odd numbers, <u>except those finishing in</u>…?*
Student 10: [Number] 5 (L1, min 12:10, emphasis added)

[37] The examples come from Pina's experimentation with practices that enhance cognitive activation and its interplay with differentiation since the pattern of refinement was more prominent in these axes rather than solely in differentiation.

Within this excerpt, Pina used open-ended inquiries to elicit students' insights at two distinct junctures (lines 3 and 18-19). However, when endeavoring to accentuate the difference between odd and even numbers and to direct the students' focus towards the characteristics of common multiples of two and five (lines 21-26), she adopted a more prescriptive approach, by explicitly indicating the aspects to be noted (lines 13-17) and prompting students to complete her statements (lines 27-28). A similar pattern of interaction was observed in the whole-class discussion around the final question of L1's task (Figure 25). In a discussion around this episode, Pina acknowledged the decrease in task demands and explained that she intended to provide more clarifications to avoid student confusion—probably being pressed for time since the lesson was nearing its end (VCS2, lines 193-194).

Figure 25

Pina's Main Task in Lesson 1

Demetris works on the computer. At the top of the screen, there is a green and a blue bulb.
He consecutively typed the numbers 1 through 15 and noted when each bulb turned on.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
GREEN		✔		✔		✔		✔		✔		✔		✔	
BLUE					✔					✔					✔

(a) What do you observe?

(b) Complete the table to show in which other numbers each bulb will light up, according to your observations.

	16	17	18	19	20	21	22	23	24	25	38	43	75	100
GREEN														
BLUE														

(c)When will the green bulb light up?

____

(d) When will the blue bulb light up?

____

(e)How many times will both bulbs light up concurrently, if we type numbers 1 to 100 consecutively?

____

(Mathematics Curriculum, 2016, Grade 3, Unit 3, p. 8, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

In light of this discussion in VCS2, notable advancements were made in L2 (which aimed at formulating a general principle for the number of dots in any specific term of a V-formation in a bird flight pattern, see Figure 26), regarding the facilitation of opportunities for mathematical reasoning. For instance, in her attempt to link odd and even numbers to the overarching pattern rule, Pina posed questions such as: "What do you notice? […] Why do the dots always amount to an odd number? How helpful is the number of the term in identifying the number of dots of any given term of the pattern? (L2, min 37:48). Also, she broadened the discussion, challenging students to consider if a formation could comprise 60 dots, employing reverse thinking. The incremental complexity of the main task supported Pina towards this more investigative approach.

Despite these improvements, Pina's approach in L2 was occasionally directive—yet less than in L1. For example, following the students' autonomous efforts to extend the pattern and fill in the accompanying table (Figure 26), Pina steered the group discussion with leading questions like, "Which dot should I draw *first*? How many additional *pairs* are needed for this term? […] What should we consider when forming *pairs*? How are they arranged?" (L2, min 8:26-12:36). Reflecting on the discussion, Pina acknowledged that whilst more advanced students had quickly grasped the pattern, it was unclear whether less advanced learners had the chance to *independently* make discoveries (POI2, lines 3-10). She admitted to repeatedly emphasizing "pairs", the "bird-leader" and the "arrangement of birds" to *prevent* their errors (POI2, lines 10-13).

Figure 26

Pina's Main Task in Lesson 2

Sometimes flocks of birds fly in impressive formations, like the following:

Helen used dots to show the pattern created by the bird formation.

(a) Draw the next two formations of the pattern.
(b) Fill in the table below.

V-formation number	Number of dots
1	3
2	5
3	7
4	
5	
6	

(c) What pattern can you observe in the table?

(d) How many dots will the 7th and 10th formations of the pattern consist of?

(e) Draw a formation of the pattern, consisting of 19 dots.

(f) Is it possible for a formation of the pattern to consist of 40 dots? Justify your answer.

(Mathematics Curriculum, 2016, Grade 3, Unit 3, pp. 53-54, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

Drawing on the list of criteria for selecting which students' solutions to share, developed in VCS3, Pina effectively sustained the intellectual rigor for whole-class interactions across her final lessons, facilitating substantive mathematical discourse. Departing from a pattern in earlier lessons in which *all* student responses were virtually presented, during L3 and L4), Pina adeptly presented and sequenced *selected* student solutions. Her selection process was informed by her observations and interactions during students' autonomous work, with a focus on how each solution aligned with the lesson objectives.

For instance, in L3, which centered on the categorization of triangles based on their angles, Pina employed a strategic approach to sharing student responses. Initiating the conversation, she chose to begin with student solutions that sorted triangles into three or five categories, some based on non-conventional criteria (e.g., "curved triangles", "narrow triangles", "open triangles", "close triangles", non-triangles, L3 min 19:49). She prompted students to articulate the reasoning behind their classifications and to identify commonalities within and across different groups (e.g., "Into which group from those noted on the board could these cards fit? [...] If you wanted to put the shapes into only two groups, which groups would you make?"). This encouraged a deeper exploration beyond their initial groupings, eventually consolidating them into two main categories: triangles and non-triangles. Through this process, the distinctive features of triangles were unveiled, seamlessly transitioning the discussion towards classifying triangles by their angles, aligning with the core objective of the lesson.

A significant shift was also observed in orchestrating a whole-class discussion that sustained and expanded students' mathematical reasoning in L4, focusing on discerning the interconnected multiplication patterns of 2, 4, and 8. The insights gained from the strategies discussed for doing so during VCS4 proved to be beneficial for Pina. An illustrative example was seen in the synthesis phase of the lesson. The scene unfolds as the class reconvenes into plenary after autonomous work on question A (Figure 27). Students were encouraged to present the solutions they had devised:

Student 1: [shares his solution:] 3X8.
Pina: Please, *explain what your mathematical sentence means*.
Student 1: 3 cupcakes and 8 boxes…
Pina: Are you sure?
Student 1: [revises his solution:] 3 boxes and 8 cupcakes.
Pina: *Can someone say this more accurately?* Are we going to add boxes and cupcakes? *Think about it*.
Student 2: 3 boxes each having 8 cupcakes.
Pina: Good. Having 8 cakes each. Does their sum equal 24?
Students: Yes.
Pina: Bravo. *Any other solutions? [to all students:] Listen to your classmates*.
Student 3: 6X3.
Pina: *Why 3?*
Student 3: [revises:] (6X2) + (3X4)
Pina: *What does 6X2 mean? What does 3X4 mean?*
Student 3: 6 boxes each having 2 cakes and 3 boxes each having 4 cakes.
Pina: Bravo. [addressing the whole class:] Does this get us 24 cakes?
Yes: Yes.
Pina: OK. *Let's hear other ways.* (L4, min 29:08, emphasis added)

In this episode, Pina continually prompted students to articulate their reasoning behind their solutions (lines 2 and 13) and to relate their mathematical ideas to the specifics of the task at hand (line 15). She adeptly rephrased students' contributions to clarify and validate their thinking (line 9), while also creating opportunities for students to evaluate (lines 6-7) and attend to (line 11) their peers' approaches. Moreover, Pina encouraged the presentation of diverse solutions (lines 11 and 20) and urged students to verify whether their proposed solution aligned with the task's stipulation of a total of 24 cupcakes (lines 9 and 18).

As the discussion progressed, additional solutions were brought forward. Pina strategically chose to focus the class's attention on a select few (3X8, 6X4, and 12X2). She asked students to identify the types of boxes utilized in each instance, before zeroing in on two specific calculations (6X4 and 12X2):

Pina: Look at these two mathematical sentences. What do you observe?
Student 9: [multiple students raise their hands]: 12 is *twice as much* as 6, and 2 is *half of* 4.
Pina: Nice. Why does this happen? Why did we need *twice as many* boxes?
Student 7: Because 4 is *double* as 2.
Pina: Yes …? How were we placing the cakes each time?
Student 7: In the second case, they were using half the cakes compared to the first case.
Pina: [rephrases]: So, because the box accommodates half of the cakes, we need twice the number of boxes. Nice. Do you notice anything

concerning 3X8 and 6X4? How many boxes of 8 did we use? How many boxes of 4 did we use? What happens to the boxes?

Student 10: The number of boxes has *doubled*.

Pina: Why?

Student 10: Because the number stays the same [apparently referring to the product, but then continues by mixing addition and multiplication] I add to 3 and then subtract from 6…

Pina: [interrupts and helps]: The number of boxes has doubled. What happens to the cakes in each box? [Not many students participate, and the teacher scolds them, saying "You seem to be still asleep. Come on!"]

Student 3 The cakes are *half*.

Pina: [rephrasing]: Yes, because the box accommodates half of the cakes, I now need twice as many boxes. [Finally, the teacher goes to the last pair of mathematical sentences—3X8 and 12X2—and again asks students to make observations if they notice any relationships]: What happens to the boxes, and what happens to the number of cakes in the boxes?

Student 11: It is 3 times.

Pina: [challenges her]: 3 times?

Student 12: 4 times.

Pina: [rephrases]: The number of boxes increased 4 times. What happened with the cakes?

Student 13: They are again *four times*.

Pina: The boxes quadrupled, what about the cakes in the boxes? Let's remember fractions.

Student 14: *They are ¼!!!*

Pina: Nice! (L4, min 39:34)

Pina steered the discussion towards uncovering the interconnections among the multiplication patterns through the use of open-ended questions (lines 1, 10-11, and 27-28), consistently probing the rationale behind student thinking (lines 4 and 14), and revoicing their contributions (lines 9, 23-24, and 32). Initially open to a variety of solutions, she later narrowed the focus to three of them, guiding the students to scrutinize the changes in the boxes and cupcakes with each solution (VCS7, line 512). She strategically encouraged students to draw parallels between the mathematical sentences, starting from simpler relationships, involving the concepts of "doubling" and "halving". She then progressed to more complex relationships, incorporating the notion of fractions. Reflecting on the lesson, Pina attributed her effectiveness in facilitating this discussion to her analytical lesson preparation and task sequencing, inspired by the insights gained during VCS3 (VCS7, lines 542-543).

Figure 27

Pina's Main Task in Lesson 4

Danae bakes cupcakes. She puts them in small boxes of 2, in medium boxes of 4, and in family boxes of 8.

a) If she baked 24 cupcakes, how many boxes would she need?

b) If she baked 32 cupcakes and used only boxes of the same size:

- How many small boxes did she use?
- How many medium boxes did she use?
- How many family boxes did she use?

c) What do you observe?

____

____

d) On Tuesday morning, Danae had 4 family boxes with cupcakes in the confectionery. She thought that they would not be sold. What can she do?

(Mathematics Curriculum, 2016, Grade 3, Unit 5, pp. 49-40, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

Gradually improving in directing different types of questions to different (groups) of students. This subtheme highlighted Pina's evolving approach to the interplay of both axes through questioning during autonomous work. Initially, in her first two lessons, Pina's strategy involved posing a *uniform* set of questions to all students. In L1, she circulated among pairs, asking each the same series of close-ended questions while they worked on identifying which numbers would light up each bulb (question a, Figure 25).

1	**Pina:**	*When does the green bulb light up?*
2	**Student 1:**	To those numbers [indicating the ticks].
3	**Pina:**	Read the numbers.
4	**Student 1:**	2, 4, 6, 8, 10, 12, 14

Pina: *When does the blue bulb light up? Tell me the numbers.*
Student 2: 5, 10, 15, 20.
Student 1: They are odd numbers.
Pina: Are they odd numbers? Did it light up on number 1? Think about it [heading to and addressing a second student pair:] *See the ticks. When does the blue bulb light up?*
Student 3: 5, 10, 15, 20.
Pina: This is correct. *And when does the green bulb light up? To which numbers?*
Student 4 To the odd numbers: 2, 4, 6…
Pina: Write your observations. Do you notice anything weird about some numbers? Think about it. [She heads to a third group:]
Student 5: [the student shares an inaudible observation]
Pina: Yes, this is correct… But *when does the green bulb lights up?*
Student 4: On multiples of 2.
Student 3: Both bulbs light up here [indicating the numbers].
Pina: Hmm... good. Do you notice anything about these numbers? Think about it. (L1, min 4:55, emphasis added).

In this excerpt, Pina interacted with three pairs of students, consistently using a similar set of closed-ended questions during these interactions (lines 1, 10, and 18). While she effectively refrained from diminishing the students' thinking and concluded each interaction with an open-ended question (lines 8, 15-16, and 21), her approach lacked customization necessary to engage students at their cognitive levels. In a reflective post-lesson interview, Pina acknowledged her need for further development in crafting learning opportunities that would allow more advanced students to deepen their understanding while simultaneously providing the right level of support for less advanced students, ensuring the task remained challenging for all (POI1, lines 28-38). Similar interactions occurred in L2.

VCS4 emphasized the critical role of diverse questioning in supporting students at different readiness levels during autonomous work. Pina actively contributed, offering various questioning strategies. This seemed to be associated with a noticeable evolution in her teaching by L3 and L4, in which she demonstrated heightened attentiveness to student responses. Drawing from their input, she crafted tailored questions that effectively stimulated their thinking, varying these questions to suit the individual or group she was addressing. For example, in L3, as students embarked on the task of categorizing triangles based on angle measures, Pina approached a group and posed an open-ended question:

Pina: What do you think?
Student 1: We put the triangles in one group and those which look like triangles in another group.

Pina: Good thinking. Which shapes will be put in each group?
Student 1: [S1 and S2 sort the cards] This… and this one…
Pina: [addressing S2] Can you repeat the criterion by which the cards in this group were classified?
Student 2: [repeats S1's production] We made two groups: the triangles and the non-triangles which look like triangles.
Pina: Why did you put that in the group with "the shapes that look like a triangle"? [pointing to a specific shape]
Student 2: Because they are not closed shapes.
Pina: So, what characteristics should triangles have?
Student 1: They should be closed shapes and not curved.
Pina: Which shapes are curved and [thus] were excluded from your group?
Student 2: [points to the shapes]
Pina: [repeats and asks:] Why did you exclude this one? [pointing to a specific shape] It is similar to another shape you have already kept out, to which one? [the students are thinking]
Student 2: [shows the similar shapes and says:] They both have empty spaces.
Pina: Your classification is very good. Think about the common features of the triangles and then find another way or criterion to sort the shapes. (L3, min 7:24)

In this episode, Pina asked a more advanced group to share their classification system, by asking for explanations, even if their solution was correct, to ensure their understanding (lines 2-3); having a student interpret another's thinking (lines 6-7); or challenging them to contemplate alternative classifications (lines 21-23). Conversely, Pina's approach with a less advanced group, which had an alternative idea about typical and non-typical triangles, differed largely:

Pina: Let's see what you have done here.
Student 3: We created groups.
Pina: What groups did you create?
Student 3: [In this group] we have the triangles [and in this group] the non-triangles.
Pina: [showing a scalene triangle] Why is this [card] not a triangle?
Student 4: Because it has one line [i.e., side] which is longer.
Pina: Isn't it a triangle?
Student 4: No. The triangles have equal lines [i.e., sides].
Pina: Which "lines" are equal?
Student 4: These two [pointing to the two sides of an isosceles triangle]
Student 3: Yes, they are equal.
Pina: [realizing that apparently, the students were able to identify only the typical representation of a triangle] OK… Well, do these shapes [i.e., those that were classified as non-triangles] have similarities and were put in the same group?
Student 3: No. Those two are similar and that is why they were grouped together [pointing to two shapes].
Student 4: But they are not triangles.
Pina: So, should the triangles always have equal "lines"? Did you measure the lines of these triangles, or did you assume that they were equal?

Student 4: [seems to have changed her mind] These shapes can be triangles.
Pina: Do you have a different opinion? Please share your ideas.
Student 3: [changes his mind too] They are triangles!
Pina: [addressing Student 4] What characteristics do they have to be triangles? Explain your thinking to [Student 3].
Student 4: They have three angles.
Pina: And what else?
Student 3: Three lines.
Pina: Can you think of a more suitable word for "lines"?
Student 3: Three sides!
Pina: Bravo! Do those shapes have three angles and three sides? [pointing to the cards which were classified as non-triangles]
Student 4: Yes.
Pina: Nice! Try to reconsider your classification and I will be back in a couple of minutes. (L3, min 10:09)

Pina initiated the interaction by asking students to explain their progress (lines 1 and 4). Upon realizing the misalignment with the criteria, she used focused questions to reveal their thinking process, without correcting them outright but attempting to help them identify their errors on their own (lines 7, 11, 14-17, 21-22, and 26-27). After the discussion, she allowed more time for reevaluation, highlighting reflection and self-correction in learning (lines 36-37).

These examples showcased Pina's attentive monitoring of her students, to tailor her questioning. In both cases, she initiated interactions with open-ended questions to gain valuable insights, then tailored her questions to the students' task progress. This ensured that upon her departure from a student group, the students could refine their work. Pina strategically navigated through the groups, carefully timing her interventions to align with each group's progress in the task. For instance, during an occurrence of student autonomous work in L3, she sequentially engaged with groups 1 through 3, revisited group 2, proceeded to groups 4 to 10, and made return visits to groups 5 and 7 (L3, min 7:09). This approach was similarly evident in L4, underscoring her commitment to responsive questioning.

Struggling with certain ideas inherent in concurrently attending to both fronts. The second theme pertained to Pina's observed or perceived challenges when concurrently pursuing cognitive activation and differentiation. The challenges could be organized into two subcategories, based on their occurrence or resolution timing. The first group pertained to *challenges that were mitigated or addressed* during the PLD program, such as doing the thinking for the students and extending the mathematical challenge for students without adequately taking into consideration their capabilities and performance on the main task. Clustered under the second group were different *unresolved challenges*, including contextual challenges that impeded the dual work of cognitive activation and differentiation (such as facing time constraints and overcoming the motivational barriers faced by students from immigrant backgrounds). Additionally, this category included challenges related to certain practices needing further polishing to serve both axes, such as becoming more judicious as to when to interrupt autonomous work at critical lesson junctures; working more closely with less advanced students at the expense of supporting more advanced students; and designing enablers that warrant some improvement. Each of these challenges is further elaborated upon below.

Challenges Addressed or Mitigated During the PLD

Doing the thinking for the students. Pina was gradually improving at preserving cognitive activation during student autonomous work. For example, in L1 (focusing on the multiples of 2, 5, and 10, see Figure 25), she seized some opportunities for students to provide explanations, by asking them to make observations as to when each bulb was going on or think of patterns (e.g., "When is each bulb going on? Is there anything weird in this table? […] Pay attention to the ticks. For which numbers is each bulb going on?" L1, min 5:00). However, in multiple instances, when monitoring students' work Pina was more directive than needed. For example, when the students were working on figuring out how the computer would behave when Demetris hit a particular number on the computer (see Figure 25), the following interaction occurred:

Pina: Which number do you have?
Student 1: [Number] 81.
Pina: Good. Which bulb will light up? *Which numbers end in 1?*
Student 2: The blue [bulb].
Pina: The blue bulb lights up on multiples of 5. So, will it light up? *Does your number end up in 5 or 0?*
Student 1: No...
Pina: When does the green bulb light up?
Student 2: In even numbers.
Pina: Is this number even? *No, it is not*. When do both bulbs light up? *Look at the numbers which are multiples of 10: 10, 20, 30…* Is your number a multiple of 10? Where should your number be included? [the teacher heads to another student pair and then brings the classroom back to plenary] (L2, min 32:54, emphasis added)

In this interaction, students were working with a number that did not fall into any category (line 2). Pina posed close-ended questions, providing some answers to students, and largely doing the thinking for them (lines 3; 5-6; 10-12). Perhaps, the fact that she rushed to proceed to a whole-class discussion suggests that she was pressed for time. This was more or less the pattern in L2, as well.

In her last two lessons, Pina consistently maintained the intellectual rigor of the task during student autonomous work. For instance, in L3 (on the classification of triangles), she navigated through all student groups, asking open-ended questions that spurred mathematical reasoning, such as "Describe the criterion used for classifying the cards in these groups. Why are these cards grouped together? Why did you put this card in the group named 'the shapes that look like triangles'? Why did you exclude this card from this group? What differences are there between the triangles and the non-triangles? What are the key features of a shape to be included in this group? (L3, min 7:09-19:49). For students who were unclear in their reasoning or who had set dubious classification criteria, she employed clarifying questions and suggestive prompts to encourage deeper reflection, to reassess their criteria (e.g., "You've categorized triangles as 'wide' and 'narrow'. What issues might arise with this criterion? How do we determine if a shape qualifies as 'wide' enough for this category? [...] Identify a criterion that is indisputable." L3, min 36:07-50:13). For students with well-defined classification systems, Pina challenged them to delve deeper. For instance, when a student pair categorized shapes as "triangles" and "non-triangles" based on the presence of "straight lines," she urged them to explore further commonalities within the groups. This led to a moment of reflection and slight disagreement among the students until they considered the angles of the triangles (L3, min 7:09). A similar approach was adopted in L4, demonstrating Pina's commitment to deepening students' understanding of the mathematical concepts at hand.

Extending the mathematical challenge for students without taking their capabilities and performance on the main task too much into account. In her initial lessons, Pina applied a *uniform* approach to elevating the task's complexity for all students, without fully considering their performance on the main task or their ability to cope with the heightened challenge. For instance, in L1, dealing with the multiples of 2, 5, and 10 (Figure 25), she broadened the inquiry to the *entire class*, prompting them to contemplate the outcomes for the bulbs upon reaching the number 81, briefly touching on its exclusion from any multiple groups (L1, min 36:51). In L2, focused on the V-formation pattern (Figure 26), she expanded the discussion by encouraging *all* students to apply reverse thinking and consider if a formation could comprise 60 dots (L2, min 37:46). Pina's guidance was overly directive for those students struggling with the extending questions.

Contrastingly, in her subsequent lessons, Pina adopted a different approach to offering extenders. These were provided based on students' completion and comprehension of the main task. Notably, in L3, besides carefully attending to students' work, Pina gave an extender to *early finishers*, challenging them to devise alternative classifications of the triangles, without having been officially and explicitly exposed to extenders during the video-club sessions. This additional task was given to three pairs after they had lucidly articulated their triangle sorting criteria for the main task.

The first formal experimentation with extenders, following their introduction in VCS5, occurred in L4. Pina planned *multiple extenders*, such as adding constraints or seeking alternative solutions (POI4, line 97). In this lesson, nine students received some type of extender. Due to the absence of some more advanced students expected to use these extenders on that day (POI4, lines 96-97; 103-105), adjustments were made to who would receive the planned extenders: imposing constraints on the number of boxes for one student (L4, min 26:01), suggesting different solution methods for six students (L4, min 23:15), and mixing box sizes for two students (L4, min 26:19).

Noticeably, in that lesson, Pina was also systematically checking early finishers' work around the main task, before providing them with an extender, by observing their written work and praising them (e.g., "Very well! Could you find an alternative solution using only five boxes of any size? Give it a try!", L4 min 26:01). In other cases, when students encountered some difficulties with the main task, she asked them to identify and rectify their error, before proceeding to extenders:

Pina: [The teacher noticed that the mathematical sentence did not match the wording of the problem: 2X12, 8X3, and 4X6] Do we have boxes of 12?
Student: No…
Pina: Do we have boxes of 3?
Student: No.
Pina: Please have a look at your mathematical sentences and I will be back in a minute [the teacher visited another student and returned to this student, who had corrected his mathematical sentences.] Bravo! Now they are correct. Can you mix the boxes?
Student: So, [write something like] 1X8 and…
Pina: Yes, do not use only boxes of the same size. Mix the boxes. (L4, min 25:31)

In this episode, Pina posed two questions to help this more advanced student identify his error (lines 2-3; 5). She gave him some time to work and when she came back (lines 7-10), she asked him to find a second way to solve the task by setting a constraint (i.e., mixing the boxes, line 12).

Developing extenders was an area in which Pina did very well, as evidenced in Table 12. Nevertheless, she acknowledged that thinking of appropriate extenders for a task without deviating from the original task's goal was time-consuming (EPI, lines 150-158). This does not undermine her success; instead, it indicates the sustained effort needed to refine her skills, making the process of creating extenders more intuitive and less time-consuming (EPI, lines 372-379).

Unresolved Challenges

Time constraints. During the PLD program, Pina repeatedly expressed her struggle to maintain the challenge for all students within a limited time frame. From L1, Pina was concerned about the lesson taking much longer than anticipated when trying to reach *all* students ("The lesson took longer than I expected" POI1, line 4). Time constraints were consistently mentioned as her main concern in the next lessons. The tension Pina faced pertained to balancing between refraining from sharing an answer or over-guiding the students and completing the lesson at the predetermined time: "I had to scaffold the whole class… The lesson took *so long* because I tried as much as possible to hold back from telling" (POI3, lines 88-89). She further explained that teaching a mixed-ability class was demanding as it tended to be more time-consuming:

Pina: The situation is difficult. In a mixed-ability class, like mine, *each student needs their own time*. Normally, there's a teaching assistant in my classroom, but she was absent from this lesson to support some students.

Interviewer Do you think that the situation would have been better if the teaching assistant had been in the classroom during today's lesson?

Pina: *Of course!* She could work with two or three students so that I could support the rest of the class. (POI4, lines 73-78, emphasis in the original)

In this excerpt, Pina was concerned with the different paces at which students work and learn (lines 1-2). Moreover, she brought into the equation the need for a teaching assistant to provide support to individual students or small groups in a mixed-ability class to increase the time of support each student receives (lines 2-3; 7-8).

This difficulty was not eliminated until the end of the PLD program. In her EPI, she explicated that she ran out of time when attempting to balance both axes:

> Time management was an issue; my lessons always took longer than planned. A forty-minute lesson lasted eighty minutes. I didn't regret it or feel like I made a mistake... I saw that my students needed more time to work on the task at their own pace and I wanted to give them the time without telling them what to do. (EPI, lines 114-121)

As this excerpt suggests, giving students extra time to work was a conscious effort on Pina's part to avoid doing their thinking and allowing them space to productively struggle on the task. Finally, she acknowledged that balancing the mathematical challenge needs time and perseverance on the part of the teacher to improve (EPI, lines 131-133).

Overcoming the motivational barriers faced by students from immigrant backgrounds. Another contextual challenge faced by Pina was the indifference of students with an immigrant background to their learning, who, despite not having learning difficulties, often prioritized other aspects of their lives over academic achievement (VCS9, lines 568-590):

> [The problem is] their indifference, not because of their learning abilities. [...] They do not care, they do not try, they just sit on the chair. *If they try, they could achieve something, and I could provide more support*. Other factors affect their learning. For instance, in my school, there are a lot of children who migrated to Cyprus from Syria. [...] They have more important problems to be concerned about, so it is difficult for me to intervene. [As a teacher] *you always do your best, but it is not easy.* It is not a typical class with let's say low, mid, and high achievers. When a child's concern is to survive, secure food for lunch, or be fed before going to bed, the last thing this child wants to deal with is mathematics. (EPI, lines 558-585, emphasis added)

This excerpt echoes Pina's anxiety to support students with a first-generation immigrant background who do not prioritize learning. Pina felt that she could not support them properly, as she felt that factors beyond her control were driving their behavior.

Becoming more judicious as to when she should interrupt autonomous work at critical lesson junctures. Another challenge encountered by Pina pertained to appropriately handling the time needed to be allocated to student autonomous work. In several video-club sessions, Pina described that she felt that she should have shifted to the whole-class discussion earlier in her lessons:

> A question that I brought up in one of the sessions was: when do we stop the autonomous work and bring students back to the plenary? In a videotaped lesson that focused on classifying triangles [i.e., L3], I noticed that neither more advanced nor less advanced students had comprehended certain ideas… so, *I gave a lot of time* [to autonomous work] and this lesson lasted eighty minutes [instead of forty]. In a discussion with [teacher educator name], my questions were "When should I interrupt autonomous work? Should it be interrupted earlier?" because I felt that the lesson was stretched. I left students working on the task for too long, and some of them got lost and couldn't make any progress. I was thinking: "Should I tell them [the correct solution]? Should I make the task more specific for them? When should I shift to the whole-class discussion, especially when I see that my students do not seem to understand? Should I let them wander?" This was an issue that troubled me. Allowing some time for struggle is good, but when and how much time should I leave for the students struggling? (EPI, min 379-399, emphasis in the original)

Describing her dilemma about when it is appropriate to transition from one phase to the other, Pina acknowledged the fine line between productive and unproductive struggle. She also shared the frustration she experienced when wondering how much time should be devoted to autonomous work, especially when students were working on the task unproductively. Even by the end of the program, she remained torn between interrupting the autonomous work and providing some hints that would reduce the demand.

Working more closely with less advanced students at the expense of also supporting more advanced students. Pina frequently expressed her need to focus on supporting students who struggle to meet the demands of a task and at the same time, extending the thinking of students who complete the task earlier and need extra challenge. For example, in L1 she argued that she wanted to support "less advanced students to become more actively involved in various activities and concurrently, give motivation to students who are capable of more so that they are not bored while working on easy tasks" (POI1, lines 28-38). Her desperation was expressed in the following question: "What should I do when they [less advanced students] can't understand even the simplest thing?" (POI2, line 38). Realizing her difficulty in scaffolding less advanced students, Pina tried to offer support to them by creating homogeneous groups in terms of their readiness level:

> I will group less advanced students to help them together during autonomous work. They will receive an enabler to work on because the way the task is presented in the student textbook will be challenging for them. The remaining students are medium to high ability, most of whom can work independently. (PRI4, lines 10-16)

In L4 (on the multiplication patterns of 2, 4, and 8), Pina grouped six less advanced students (PRI4, lines 10-17) and provided them with an enabler (see Figure 28). Pina developed multiple enablers, one for each question of the main task, drawing on three strategies generated and codified during the video-club sessions: breaking up an open question into single-step questions; modifying the wording of the task questions; and using representations.

The enablers turned out not so helpful for students: some were using subtraction instead of multiplication; others were stuck and did not know how to start; and others were using box sizes not mentioned in the task instructions (POI4, lines 5-6). Consequently, Pina spent more time with this group and rarely visited the early finishers, who surprisingly were raising their hands for more than five to eight minutes and Pina never visited them; apparently, there was no routine as to what early finishers should be doing (L4, min 7:22). After approximately 20 minutes, Pina visited those students, recognizing that they had been waiting for her for quite a while. She posed some clarifying questions and then asked them to check their mathematical sentences. Reflecting on the implementation of the first enabler, Pina explained:

> Some students got confused because of the enabler. They were using it incorrectly. Although I explained to them that they should circle the cupcakes in groups of two, four, or eight, they could not understand what they were expected to do. So, I asked them to figure out solutions using only one box size, instead of mixing the boxes. (POI4, lines 35-40)

Figure 28

Pina's Enablers of Lesson 4

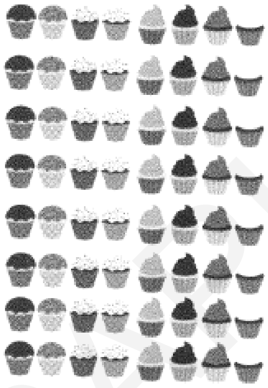
Enabler 1 (For the first question of the task in Figure 27):

1. **Circle any boxes you need to get 24 cupcakes.**

Boxes of 2

Boxes of 4

Boxes of 8



2. **Complete the table.**

Boxes of 2	Boxes of 4	Boxes of 8

3. **Write the mathematical sentence that fits your drawing on p.49 of your textbook.**

Enabler 2 (For the second question of the task in Figure 27):

- **In how many boxes of 2 can 32 cupcakes fit?**

Mathematical sentence: ______________________________

- **In how many boxes of 4 can 32 cupcakes fit?**

Mathematical sentence: ______________________________

- **In how many boxes of 8 can 32 cupcakes fit?**

Mathematical sentence: ______________________________

In her first experimentation with this tool, Pina realized that the way the cupcakes were presented on the enabler might not be very supportive. Specifically, she realized that there was some confusion with putting the cupcakes in groups, while a slightly modified iconic representation could be better (EPI, lines 62-64). To help the students overcome this difficulty she asked them to use only boxes of the same size instead of mixing them, an appropriate simplification for less advanced students. Furthermore, she noticed that less advanced students as individuals had different needs and should be provided with *individual support* (EPI, lines 124-129). Hence, she could not help them *as a group*, as intended, resulting in spending a lot of time "to figure out what each student hadn't understood, provide explanations to continue working on the task; hence, no time was left to visit the early finishers" (EPI, lines 135-140). Having discussed Pina's strengths and challenges in her teaching, the attention now turns to the evolution of her conceptual understanding around the two axes and their interplay, delving into how her understanding has transformed over the course of the PLD program.

5.5 Learning Level: Evolution of Pina's Conceptualizations Around the Two Axes and Their Interplay

Pina considerably progressed in her conceptual understanding, especially in the interplay of both axes. Initially, her grasp was characterized by a solid familiarity with cognitive activation, coupled with nascent ideas and some misconceptions regarding differentiation. As she progressed, a notable shift occurred, moving from viewing these axes in isolation to a more integrated perspective. By the program's conclusion, Pina's conceptual understanding of these axes had matured significantly, with a marked reduction in focusing on either axis in isolation. Instead, her reflections and practices centered around the nuanced interplay between them.

5.5.1 Initial conceptualizations

Familiarity with Cognitive Activation. Equipped with a master's degree in mathematics education and a previous PLD experience on issues of cognitive activation, Pina appeared to be familiar with challenging tasks from the beginning of the program (EPI, lines 920-925). During VCS1, as the participating teachers discussed strategies to enhance student engagement in deeper thinking, Pina drew upon her background to highlight key characteristics of challenging tasks. She emphasized the importance of *selecting* tasks that resonate with students' experiences and interests by situating them within realistic scenarios (VCS1, lines 45-46), avoiding tasks that could be solved through rote or predictable methods, and presenting problems in formats that diverge from students' customary approaches (lines 74-76). Additionally, she pointed out that such tasks might necessitate visual representations or the use of manipulatives (line 79) and typically require students to articulate their thinking and reasoning (lines 98).

Moreover, she acknowledged the abundance of challenging tasks within the Cypriot student textbooks (VCS1, lines 100-101). Engaging in an activity to classify a set of tasks based on their demands, she adeptly distinguished between high- and low-demand tasks, elucidating why certain tasks offered greater or lesser opportunities for critical thinking. For example, upon identifying tasks with low cognitive demand that consisted merely of completing a series of multiplications, Pina noted that these tasks were not thought-provoking "beyond the mechanical execution of the multiplication procedure" (VCS1, line 58). However, she recognized the importance of such tasks in developing students' ability to follow mathematical procedures accurately, swiftly, and effortlessly (VCS1, line 64).

Besides the characteristics of the task itself, Pina added the parameter of *students' age and grade* in deciding the level of challenge offered to students:

> For higher-grade students, this task [i.e., the one on solving a set of multiplications] is simple and requires algorithmic thinking. When used with lower-grade students, this task requires focusing on the underlying concept of multiplication. Let's say, second- and third-grade students must consider what "two times three" means to find the product. (VCS1, lines 66-69)

This contribution directly speaks to the idea of encouraging students (especially when introducing a new idea) to use some cognitive effort to engage with the underlying concepts behind the procedures and reveal why they work. This approach aligned with her previous teaching experiences; she recalled an instance in which, while teaching subtraction with regrouping, she consistently directed students' focus towards understanding the connection between the procedural steps and their representation using Base Ten Blocks (VCS1, lines 94-96).

Drawing from her own teaching experiences and a videoclip reviewed by the PLD group, Pina recommended certain teacher actions that could support not doing the thinking for students during different lesson phases. Notably, she recommended that students read the task instructions themselves to sustain cognitive activation during task launching (VCS1, line 148). She also highlighted the significance of moderating the amount of guidance provided when presenting a task, suggesting that explanations should be delivered in "stages" tailored to students' ability levels to avoid overwhelming them with information (VCS1, line 139). For student autonomous work, Pina proposed giving students adequate time to engage with the task and explore different approaches before the teacher steps in (VCS1, lines 129-132). During whole-class discussions, she emphasized the value of having students articulate their reasoning and proposed solutions, thereby deepening their understanding and extending their thinking (VCS1, line 208). Moreover, Pina's strategy of initiating discussions with a specific solution, possibly an incorrect one, illustrated a deliberate approach to sharing and sequencing student responses to maximize learning opportunities (VCS1, line 199). Her suggestions showed that she was aware of some ideas of how teachers can promote cognitive activation across different lesson phases.

In L1, which focused on the relationships between the multiples of 2, 5, and 10, Pina chose a task that would engage students' interest and planned to use probing questions to both extend the challenge and gauge their understanding during whole-class discussions (PRI1, lines 11-13; VCS2, lines 59-60). Reflecting on it, she expressed satisfaction with the active participation of students, including those typically less engaged or vocal (VCS2, lines 60-64). However, she encountered unexpected challenges, such as students struggling with concepts, she assumed they would grasp easily (VCS2, lines 74-75). Pina identified the concurrent focus on multiple multiplication patterns as a source of confusion for the students, prompting her intervention to encourage a focus on individual patterns (VCS2, lines 66-74). This experience highlighted for Pina the complexities of engaging students in challenging tasks and the need for strategic support to help them navigate these challenges effectively.

In summary, Pina entered the PLD program with a relatively advanced understanding of cognitive activation and challenging tasks. Her prior participation in a related research project, which included videotaping her mathematics lessons, contributed to her solid foundational knowledge in this area. Despite this, her first lesson revealed some gaps in her ability to anticipate and address potential student difficulties in a manner that maintains the task's demands.

Initial Ideas, Concerns, and Misconceptions About Differentiation. Pina entered the program being sensitive about the *different groups of students existing in her classroom*, according to their abilities and pre-existing knowledge. Her sensitivity to these differences was evident in L1—which preceded VCS2 in which issues of differentiation were discussed. In that lesson, she set *clear learning objectives* and she was *expecting less advanced students to work towards some progress on one of them*: "The goal is even the least advanced student to be able to identify at least the multiples of 10" (PRI1, lines 14-16). Hence, Pina had a *developing* understanding of differentiation.

Pina's strategy also involved anticipating *differentiated outcomes* from different student groups, each meeting the learning objectives in their own distinct ways. Her approach to differentiation was further illustrated by the anchoring activities and exit cards used in the lesson:

> Each student received a colored card with a different [anchoring] activity on it. Everyone was excited! Some students were asked to write numbers in which the green bulb would light up and others to write numbers that would light up the blue bulb. More advanced students were asked to write numbers that wouldn't light up any bulb. The easiest task was suggesting numbers to which both bulbs would light up. I also gave exit cards with three-digit [instead of two-digit] numbers to some student pairs asking them to figure out which bulb would light up if we hit that number. No student reacted negatively to receiving something different. (VCS2, lines 252-259)

The activities developed by Pina were *within the lesson objectives* and were both *differentiated by outcome* (by using an open-ended task and expecting very different outcomes). The different card color was used to camouflage differentiation:

> *When the differentiated tasks do not differ that much in appearance, differentiation works*. For example, this works when I give less work to some students, let's say two problems instead of three, or eight multiplications instead of twenty. *The students do not notice the difference*. (VCS2, lines 269-271, emphasis added)

This excerpt admits multiple interpretations. Firstly, the use of distinct colors to differentiate activities could be viewed as an effective planning strategy to facilitate differentiation. However, it might also highlight the prevailing classroom culture, which may work as a hindrance to promoting differentiation—it is implicitly expressed by Pina's concern regarding the possible negative reaction of students, once they realize that they are working on something different. Additionally, for Pina, differentiation might entail assigning a reduced volume of work to less advanced students, indicating a particular approach to accommodating diverse learning needs.

In sum, Pina was aware of the diverse levels within her classroom and understood that not all students would achieve complete mastery of all lesson objectives. She also incorporated some differentiation practices in her first lesson and expressed concern about whether the differentiation would be noticed by the students while teaching, finding ways to hide it.

Bringing the Interplay of Cognitive Activation and Differentiation to the Fore. Pina's interest turned towards integrating cognitive activation and differentiation after her experimentation in L1 (on the multiples of 2, 5, and 10). Upon reflection, Pina discerned different student cohorts within her classroom: a group of more advanced students for whom the challenge provided was just right; students with limited mathematical skills; and other student groups who were not challenged for various reasons, such as students with migrant backgrounds who had not mastered the Greek language (POI1, lines 13-22).

This insight led Pina to propose a focus on "differentiated cognitive activation" to better address the diverse needs of her students (POI1, line 27). This concept encapsulated her aspiration to customize her teaching approach to cater to the varied challenges and support levels required for *all* students, fostering the development of profound understandings. She expressed a keen interest in discovering strategies to scaffold the learning of both more advanced students, who "can cope with more challenging tasks [...] and shouldn't be held back by working on things that are quite easy for them", and less advanced students, whom she wished to encourage "to exert more effort and actively participate in the tasks" (POI1, lines 28-38).

This desire was further demonstrated in her approach to L2 (videotaped before bringing cognitive activation and differentiation together), focusing on generating a general rule for the number of dots in a V-formation pattern. She organized her students based on their readiness, forming a group of more advanced learners and several mixed-ability groups (PRI2, lines 12-17). However, this practice did not yield the desired effect in challenging the less advanced students, prompting Pina to seek additional guidance on how to scaffold these students more effectively to facilitate their engagement and understanding (POI2, lines 27-31).

Pina focused heavily on *handling diverse groups of students*, as evident in an incident in VCS3. The PLD group analyzed a video clip from a whole-class discussion of an algebraic task in which fifth-grade students would translate the visual pattern into a mathematical sentence to deduce a general rule. The teacher in that video clip struggled to handle different student responses, ending up doing most of the thinking for the students. Pina drew parallels between this episode and her own experiences in her second videotaped lesson, especially regarding the shared content and her challenges in balancing cognitive activation with differing student learning paces:

> I faced a similar issue in my [second] lesson. […] Some students grasped the idea. I was worried because this student had already discovered the rule… Should I encourage them to share it in plenary or leave more time for the others to work individually on the task and think of it? (VCS3, lines 314-321)

From this post-reflection, multiple issues arise. Generally, Pina problematized the tendency to ask all students to complete the task within a single time frame irrespective of their learning pace. As a result, some more advanced students were held back; meanwhile, at the other end of the spectrum, less advanced students found it impossible to keep up. She was looking for solutions to this conundrum for which she received some suggestions from the PLD group on how to flexibly use the available teaching time, such as by employing asynchronous work.

In sum, Pina's experimentation during her first two lessons and her participation in the first three video-club sessions brought to the surface various initial ideas, concerns, and needs that she had upon joining the program. These included responding to different student ability levels and learning paces. Her references to concurrently handling multiple student groups, as well as explicitly seeking ways to differentiate cognitive activation indicated Pina's focus henceforth.

5.5.2 Evolution of Pina's Conceptualizations

Consolidating Existing Ideas and Surfacing Alternative Ideas About Cognitive Activation. As it turned out along the way, the utilization of video clips from her own and her colleagues' lessons proved invaluable for Pina, not only in reinforcing and uncovering additional strategies to foster cognitive activation but also in acknowledging teaching behaviors that might inadvertently limit student thinking. For instance, upon reflecting on a lesson excerpt from another teacher participant (i.e., the third case, Michelle), Pina highlighted two practices that supported cognitive activation: encouraging students to formulate explanations because of their academic payoffs ("It helps them to talk about math and justify their answers", VCS2, lines 145-147) and contrasting multiple solutions to identify and understand errors (see the "why" behind the mistakes, VCS2, lines 147-149). This reflection indicated Pina's appreciation for leveraging student thinking and errors as learning opportunities (VCS2, line 151).

Moreover, Pina contributed to discussions about what constitutes *'stealing students' thinking.'*[38] During a conversation about a video-clip from a teacher participant's teaching, in which students were asked to identify what was given and what was asked by the task during its launching, there was debate on whether this approach might undermine student autonomy. Pina suggested that the impact of such actions might depend "on the level of the students you interact with. If you steal the thinking from one or two students, the majority will do the thinking and start working on the task" (VCS2, lines 109-110), implying that while it might restrict the thinking of a few, it could potentially support the majority to begin engaging with the task. This insight hints at Pina's ongoing challenge to balance support across diverse student groups without diluting the cognitive demand.

Additionally, Pina admitted instances in which she *felt compelled to lower the challenge for the sake of completing the lesson*, reflecting on a colleague's post-lesson reflections that resonated with her own experiences: "[The teacher] wondered if he had done more thinking than the students themselves. We all do this under time pressure" (VCS3, lines 351-352). Pina began to worry about how her teaching actions during the available teaching time could affect the learning of different student groups. She recounted her deliberate efforts during L3 to avoid providing direct answers to struggling students, especially as the lesson progressed and time became scarce ("The lesson took so long because I tried as much as possible to withhold from telling them the answer. […] What would be the point of telling them this beforehand?", POI3, lines 88-89; 106). This reflection illustrates Pina's commitment to preserving the integrity of the learning process, even in the face of time pressure.

In VCS4, the PLD group discussed that research has shown that teachers tend to inadvertently take over the thinking while supporting less advanced students. Pina confirmed that and admitted that teachers tend to do so "to feel sure that they [i.e., the

[38] Teachers used the expression "stealing students' thinking" to describe doing the thinking for their students. This expression was preserved in contexts where teachers used it.

students] have understood" (VCS4, lines 268-269). This admission marked a significant moment of self-reflection for Pina, as she recognized additional factors that contributed to her tendency to provide excessive guidance to her students, including seeking affirmation of students' understanding.

In VCS5, the PLD group revisited the concept of classifying tasks according to their demands. As in the first session, Pina effectively pinpointed specific characteristics that determined their level of challenge, using the TAG (see Table 1). During the analysis of a subtraction task involving three-digit numbers and the use of Base Ten Blocks for regrouping, Pina was the only teacher who was confident in telling the level of challenge offered by this task:

Michelle: Let's discuss Task 1 [i.e., the one with subtraction with regrouping]. I'm not sure which category it should be classified. [She oscillated between procedures without connections or procedures with connections.]

Pina: *No, it is not classified as procedures without connections* because there is a challenge in connecting the concept [of subtraction with regrouping] with the algorithm. It requires comprehension on the part of the students to see this relationship. They must exchange the hundreds for tens. (VCS5, lines 258-262, emphasis in the original)

Contrary to Michelle who considered this task as controversial, Pina recognized that this task intends to connect this procedure to meaning (lines 5-7), confirming her knowledge of challenging tasks.

The evolution of her conceptualizations around challenging tasks and their demands illustrates a significant maturation in her understanding of cognitive activation, and the teaching practices that either foster or hinder this challenge. This growth provided her with insights into moments when she might inadvertently lower task demands. Despite this progress, her explicit focus on cognitive activation had diminished somewhat from the program's onset. Moreover, Pina devoted even less attention to differentiation as an isolated concept. Yet, her engagement with the interplay of cognitive activation and differentiation became increasingly pronounced, as described below.

Focusing on the Interplay of Cognitive Activation and Differentiation.[39] As Pina was experimenting with different ideas in her teaching practice and reflecting

[39] Pina appeared to have placed less emphasis on issues of differentiation and focused more on the interplay of cognitive activation and differentiation—hence, differentiation is not discussed in isolation.

upon her experimentation, her understanding of the interplay of both axes underwent significant development. Her active participation in discussions around three key themes highlighted this evolution. These themes included the strategic use of *question-posing*, the application of *enablers and extenders*, and the navigation of challenges and opportunities in teaching *different levels of students*. Her reflections and adaptations in these areas, especially evident during the video club sessions, illustrate her growing sophistication in these domains. More elaboration on how her thinking around these themes developed over time is provided next.

Question posing. Pina gradually began to consider *the kind of questions she would pose to different groups of students*, especially to those who hit a stumbling block. For example, in VCS4, the PLD group focused on questioning to enhance the harmonious functioning of cognitive activation and differentiation. Toward this end, the PLD group read some narratives that described how different interactions took place at different lesson phases in a fifth-grade lesson focusing on exploring the relationship between the fractional part and the whole. At different junctures, the teacher participants were asked to either consider the teacher's interactions with students and her questioning or suggest what questions they would pose to challenge different levels of students. As it turned out, this activity increased Pina's involvement in the session, by explaining the *usefulness of particular open-ended questions* posed by the teacher (e.g., "By asking the students to share what they see, the teacher can understand what each student sees because each notices something different; then, by asking them what the task asks them to do, (s)he activates them to think", VCS4, lines 136-137) and *identifying incidents where the teacher lowered the demands for all students* ("At the end [of the narrative], the teacher shares some remarks herself; she could let a student explain it", VCS4, lines 169-170). The following excerpt is typical of how Pina participated in this discussion. We enter the episode when the group discusses a narrative showing how the teacher interacted with three different groups (i.e., more advanced, less advanced, and students in the middle) during autonomous work:

Teacher Educator (TE): Think about what you would do and how you would interact with each of these three student groups. [The teacher participants discuss in pairs for a couple of minutes.] Let's begin with how you would handle the student group who feel "lost".

Georgia: We could start by focusing on the concept of fractions to see if they have realized what 4/4 means.

TE: Let's go a step back.

Pina: *We could ask the students to say, "What are we asked to do here?" or even go another step back and ask: "What have you understood?".*

Georgia: Wouldn't that disorient them [i.e., the students]?
TE: We must ensure that they have understood what they are asked to do.
Georgia: We could ask them "Why couldn't you understand this?"
Michelle: … and they will answer that they don't know! [laughs]
TE: Typical answer. What do we do in that case? How can we handle it?
Stella: We can reread the task and they can explain it in their own words.
Nancy: We can analyze what the task asks them to do, what the term fraction means, what ¾ means, and so forth.
TE: Sometimes when trying to clarify things we are trapped in giving them ready-made answers.
Pina: *We can ask "What makes it difficult for you?"*
Michelle: They will answer "everything"! [laughs]
TE: At this point, we must insist on seeing what troubles them the most. It makes them think and at the same time, it gives us access to their difficulties.
Pina: *But it is not easy for the student who has a limited understanding of the task instructions to realize that and specify what they do not understand. It is a metacognitive skill and only more capable students can do it.*
TE: That is why we must insist on encouraging them to explain what the task asks them to do. (VCS4, lines 175-197, emphasis added)

In this vignette, Pina shared some open-ended questions that could be posed while interacting with the group of students who were "stuck" (lines 8-9; 21), sequencing the questions in an order she considered logical to be posed to elicit students' thinking (line 9). She seemed to take the discussion very seriously and tried to come up with solutions on how to handle this group of students, unlike Michelle who rather playfully thought that no matter what a teacher tries, students will not respond productively. At the same time, she doubted whether less advanced students would be able to respond to "metacognitive" questions (lines 26-29). This incident implies that she was considering ways to vary the questions she would pose to different groups of students.

Different levels of students. Pina realized that some stumbling blocks to learning are often predictable and can be determined by the language used in the task; other personal or cultural student characteristics, or content-related factors. For instance, drawing on the discussions of VCS4, Pina stressed the importance of asking students "to define the difficult but central-to-understanding words" of the task (VCS5, line 27): "[In L3] The words "classifying" and "criterion" are difficult for third graders. So, I asked them to define them in multiple ways […] My classroom includes many students who are foreign language speakers, so I had to ensure that they somehow had understood what they were expected to do" (VCS5, lines 30-32). Pina identified two reasons for asking students to discuss and clarify the wording of the task: *students' grade level* and *language proficiency*.

In the following sessions, Pina began to *deconstruct the concept of students' "level"*. For example, in VCS5, the PLD group reflected on the students that the teachers tend to consider as "more advanced". To her surprise, Pina noticed that depending on the *lesson content*, the *ability level* of her students differs:

> I noticed that more advanced students, who are strong in numbers and operations, have the same difficulties as less advanced students in geometry. […] So, the lesson content may play a role. (VCS5, lines 111-118)

This realization conflicted with her earlier belief that students are grouped as more or less advanced regardless of the *mathematical content* they work with. This is an interesting conceptualization of the idea of student level.

Enablers and extenders. In addition to question posing, the tool of enablers and extenders was another tool embraced by Pina; however, in using it, Pina was looking for ways to strike the right balance between the level of challenge and the level of scaffold she would offer to different groups of students. When the idea of enablers and extenders was first introduced in VCS5, Pina was trying to figure out what could be considered as an enabler (e.g., "Aren't the tables [included in the original task] themselves enablers? […] They organize student's thinking", VCS5, lines 529; 547).

After experimenting with the idea and reflecting on how the enablers and extenders she had incorporated into her last lesson worked, Pina explained that she had anticipated that students would face some difficulties with various parts of the task because of its abstract nature and therefore, she developed multiple enablers (one for each part of the main task) considering ways to visualize the task (note that creating multiple enablers was not an emphasis when introducing the idea of enablers, VCS7, lines 496-509). Pina came to re-examine the appropriateness of her developed enablers: "Some students needed more explanations on how to work on the enablers. The diagram [of the enabler] was confusing. I should have given something else, like boxes of cubes." (VCS7, lines 532-536). She realized that her students who received the enabler could be better supported by enactive representations rather than iconic ones (VCS7, line 537). Discussing the use of her enablers with the rest of the PLD group, she admitted that she deliberately represented more cupcakes than the number of cupcakes that was mentioned in the task instructions *to maintain some of the task challenge* (VCS7, line 578).

As an extender, she planned on posing constraints for "more advanced students" or early finishers (VCS7, lines 514-516). The next day, she decided to work on the extender *with the whole class*. However, as it turned out, "it was difficult for the students" and she ended up designing the cupcakes in the boxes *herself* (VCS7, lines 519-520). This shows that Pina had not yet understood that extenders are not tasks used by the whole class.

Due to her growth mindset, her first implementation of enablers and extenders did not discourage her but made her think of possible ways with which she could improve her teaching practice, acknowledging that the implementation of any new idea takes time to become established: "Using enablers and extenders needs more work and practice because this was my first experimentation" (VCS7, lines 541-543). She realized that the implementation of new ideas is not a one-time thing but needs *systematic experimentation* so that she can *understand* the new idea of enablers and extenders and use them successfully.

5.5.3 Final Conceptualizations

Evolving Conceptualization of the Interplay of the Two Axes. Drawing attention to cognitive activation or differentiation issues in isolation was greatly eliminated toward the end of the PLD program; therefore, this section focuses on how Pina's final conceptualizations around the interplay of both axes evolved. In particular, she highlighted the importance of both axes for student learning; pointed out how she saw the relationship between the two axes; referred to the teacher's role in promoting them; and discussed her pertinent gains from the PLD program and how her teaching skills could be further developed. These themes are analyzed below backed up with supportive excerpts from her end-of-program interview.

Endorsing the interplay of the two axes. After all the experiences she gained from the PLD program, Pina explicitly expressed her *trust in the interplay of the two axes*:

> Both axes are very important for teaching because they help each student to reach their differentiated learning goals. […] Not all students can achieve the lesson goals to the same extent. Both axes can motivate the students to go one step further than where they are (EPI, lines 649-653; 656-661)

In this excerpt, Pina considered how valuable both axes are for students' learning and motivation to learn. She also realized that *all* students would work on the lesson's objectives but the depth to which they will reach varies according to the student's level.

Adjusting the interrelated axes of cognitive activation and differentiation. Pina argued that cognitive activation and differentiation are two "*interrelated*" constructs (EPI, line 614). When asked to provide some examples that would support her argument, she further explained that one of the main reasons for differentiating teaching is to cognitively activate *all* students (EPI, lines 619-621), by differentiating the process; the starting point of each student; or the task context to match students' interest (EPI, lines 621-624). Therefore, for Pina, differentiation was the vehicle for promoting cognitive activation—her already well-known concept. She argued that the level of challenge that is offered to students is up to the teacher:

> If the teacher makes [the task] way too easy, the students will not be challenged because the answer will be somehow given. This depends on us [i.e., the teachers]. We need to adjust the challenge and also withhold from telling. This requires balance. It is not easy. […] We need to be careful during lesson planning, because in trying to enhance one axis, we may lose the other. (EPI, lines 625-640)

Herein, Pina spoke to the idea of *adjusting* the challenge to meet all students' needs. She seemed to consider the two axes like two dimmer switches (each corresponding to one axis) that the teacher must balance to function properly together. However, she realized that balancing the level of challenge and the scaffolding offered to students, so that the level of challenge is just right for the students, requires careful consideration.

Considering the role of student background factors in preserving the challenge. Reflecting on whether she believed that the challenge can be maintained at a high level for all students, Pina argued that this is possible, subject to certain conditions. Specifically, maintaining the mathematical challenge for all students requires good teacher preparation in advance with proper lesson organization and the inclusion of differentiation techniques (EPI, lines 534-536). The cognitive level can be maintained even for students who are less advanced in mathematics, as long as, the challenge offered to them is *just right* (EPI, lines 598-603). However, she believed that the level cannot be maintained for some cases of students, who despite having the mental capacity to work with mathematics, refuse to think and work, due to multiple reasons that are unrelated to school (EPI, lines 539-545). But even for these students, Pina argued that she does not give up trying to challenge them; rather she is troubled by the students' resignation while still keeping her expectations high (EPI, lines 555-556).

Handling these students, especially those with an immigrant background, was a main contextual challenge for Kate, as analyzed in Section 5.4.2.

Reflecting on her gains from the PLD program. Pina realized that she had consolidated various ideas because of her participation in the PLD program, including: (a) *being less talkative so as not to steal students' thinking* (VCS9, line 232-235); (b) *posing appropriate questions during different lesson phases* (EPI, line 233; 322); (c) *sharing multiple solutions during whole-class discussion* (EPI, line 323-324); (d) *appropriately organizing the lesson so that all students are cognitively activated* (EPI, lines 210-214); I *introducing a task and scaffolding them to make observations without telling them the answer* (EPI, lines 317-318); and (f) *keeping in mind that the challenge must be maintained for all students* (EPI, lines 798-803). She claimed that these tools are important for working at the interplay of both axes. Moreover, she argued that most of these ideas are also applicable to other subjects (EPI, lines 758-760).

Although she felt that she benefited in all three lesson phases (i.e., TL, SAW, WCI), Pina believed that she was better at promoting the interplay during the phase of student autonomous work because she could better monitor students, identify their challenges, and scaffold them according to their needs by using the appropriate questions (EPI, lines 338-347). She also benefited in the phase of task launching—though to a lesser extent than in student autonomous work (EPI, lines 346-347). On the other hand, she believed a teacher could do the least in terms of promoting the interplay or maintaining the level of challenge for all students during the phase of whole-class interactions (EPI, lines 348-349). When asked to justify her opinion, she explained that even though a teacher could challenge the students with appropriate questions in that particular phase, it is generally harder to "check" the extent to which each student would be cognitively activated because of the need to *simultaneously* attend to students of different readiness and ability levels (EPI, lines 354-356; 364-366).

Consolidating ideas requires continuous experimentation. Although improving in certain aspects of her teaching, Pina described that she would like to have more opportunities to improve in both axes in the future since "the issues discussed cannot be exhausted [in a single PLD program], there are still many important and big topics to learn" (EPI, lines 858-860). This illustrates a teacher who was constantly seeking learning opportunities to further develop her cognitive schemas around different aspects of her teaching. Furthermore, Pina wanted to experiment more with specific practices, such as *flexible grouping* (EPI, lines 497-501) and *enablers and extenders* (EPI, lines 502-505). For her, it was a matter of continuously trying and improving:

> With frequent experimentation, some things can become quite routine. So, it's not about whether I faced some challenges or not in the specific lessons. The more often we use the practices in our lessons, the easier it is for us to implement them. It requires constant experimentation. This is something I realized from my participation in the program. Of course, some practices may not work in some instances, the more you try, the better you get. (EPI, lines 467-481)

In sum, Pina kept approaching her learning with a growth mindset. Specifically, she had a clear focus on improving; she embraced "failure" as a chance to learn; and she took responsibility for improving her teaching practice.

Finally, in both the initial (VCS1) and the concluding sessions (VCS9), teachers were requested to position their teaching within a two-dimensional diagram represented in Figure 29. This positioning aimed to reflect the extent to which they incorporated challenging tasks in their teaching and/or catered to the varying readiness levels of their students.

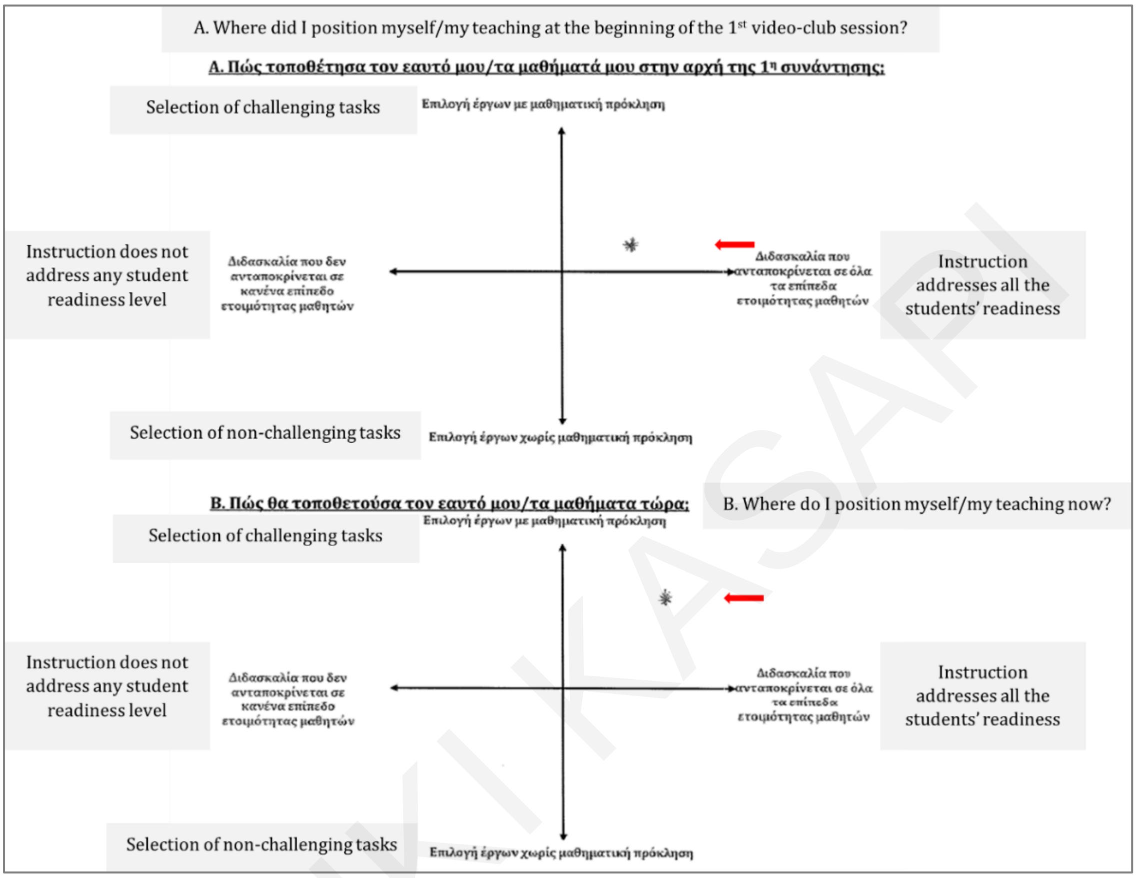
Notably, in VCS9, Pina altered the positioning of her teaching compared to VCS1, demonstrating an upward shift on both axes. Her improvements were somewhat greater in terms of using challenging tasks. Pina attributed these differences to EDUCATE:

> The difference is attributed to the knowledge I gained from the [PLD] program and the various practices that I implemented in my classroom (e.g., differentiation ideas, use of enablers and extenders, creating homogeneous groups). (TRC9, question C)

Pina identified differences in both her *conceptualizations* and *practices*. Her reference to *acquiring theoretical knowledge* on the one hand and *experimenting with various teaching practices* on the other hand suggests a recognition of two distinct areas of learning and development. These can be considered separate but interconnected dimensions of professional growth. Pina's conceptualizations reveal some awareness of her performance in the classroom. In the next section, the changes in her conceptualizations and teaching practices are compared.

Figure 29

Four-Quadrant Diagram Depicting Pina's Positioning of Her Teaching (A) Before and (B) After Participation in the Program Concerning Cognitive Activation and Differentiation, With Its Translation in English



5.6 Learning and Behavior Levels: A Comparison Between Pina's Conceptualization and Practice

The analyses of Pina's conceptualizations and practice regarding the two axes under study and their interplay revealed different paths of change. First, it illuminated that she was observed to steadily (rather than suddenly or erratically) improve in her teaching practice in the two axes and their interplay (despite minor setbacks); this means that she was either willing, open, or ready to improve her teaching. On the other hand, over time, she increased her conceptual focus on the interplay of the two axes while gradually reducing the emphasis she placed on each axis in isolation. This implies that she was a case of a teacher who shifted her attention towards exploring how these axes work together, instead of attending to them separately.

When comparing the changes in Pina's practice and conceptions, sometimes they were in agreement, while at other times, differences or inconsistencies were noticed. In particular, Pina's conceptualizations and teaching practice regarding the interplay of cognitive activation and differentiation were fairly consistent since they changed in a similar and coherent way. Some inconsistencies were observed between the changes in Pina's conceptualization and her teaching practice related to cognitive activation; she initially placed more emphasis on cognitive activation in her conceptualization, but over time this emphasis gradually decreased, while her actual practice showed a steady rise in focus on cognitive activation. A similar pattern was observed in Pina's conceptions and practice about differentiation, where there was a decrease in conceptual emphasis on differentiation despite a steady increase in her practice.[40] In a nutshell, in the case of each axis in isolation, the two patterns were not very well aligned. The magnitude to which the changes in her conceptualization were compatible with the changes in her teaching practice for each axis is discussed further below.

Starting with the axis of the interplay of cognitive activation and differentiation where the greatest coherence was observed, Pina was able to bring to the surface her emerging perceptions and needs around the interplay from the very beginning of the PLD program. Initially, she focused on the need for "differentiated cognitive activation" to scaffold diverse student groups. Over time, she deepened her understanding of these issues and emphasized specific tools, such as question-posing and enablers and extenders, to achieve the synergy of cognitive activation and differentiation. Throughout the PLD program, Pina became an advocate for the interplay and saw improvement in her teaching practice, particularly in the phase of student autonomous work—the lesson phase where she believed that a teacher could better attend to the needs of different students. Although there were some setbacks as she experimented with these practices, her overall trajectory was upward. In essence, Pina's focus on and commitment to understanding and implementing the interplay between cognitive activation and differentiation were key factors in her improvement as a teacher. It also implies that her experimentation and willingness to try new practices were important in helping her overcome setbacks and continue to make progress. The fact that the changes in her conceptualizations and her teaching practice were found to be fairly consistent (with some minor inconsistencies) suggests that she was able to integrate new understandings into her teaching practice in a meaningful way, and vice versa.

As far as cognitive activation is concerned, Pina's conceptualizations did not always align with her actual teaching practice, and there were some discrepancies

[40] The decline in Pina's attentiveness to issues of cognitive activation or differentiation separately could be due to the fact that over time, she was attending to the interplay more.

between the two. Upon her entry into the program, Pina's prior knowledge and experience in teaching mathematics (i.e., her additional qualification in mathematics education and her earlier PLD experience with the use of challenging mathematical tasks) were evident in how she identified and analyzed challenging tasks and her understanding around some factors that facilitate or inhibit cognitive activation. While Pina entered the program with confidence and skill, her first lesson showed that maintaining high cognitive activation was more complex than she anticipated. However, as she continued to experiment with cognitive activation practices, she showed significant growth in her practice, solidifying her understanding of what makes a task challenging and how to handle factors that could lower its demands. Despite reducing her focus on cognitive activation towards the end of the PLD program, Pina's steady and substantial improvement in her practice demonstrated that her initial knowledge on this axis provided a good foundation for her to build on. Essentially, although she did not remain preoccupied with cognitive activation, Pina was given the opportunity to bring back earlier ideas and experiment with them systematically across her four videotaped lessons while trying to understand the new concept of the interplay of cognitive activation and differentiation. Therefore, the opportunity for experimentation, as well as the enrichment of her arsenal of teaching tools with additional ideas (including ideas from the other two axes) helped her demonstrate great growth in her actual practice, even if she no longer focused exclusively on the axis of cognitive activation.

Regarding differentiation, while Pina's practice steadily improved over time, her conceptual understanding thereof did not necessarily progress at the same rate. Pina's practice generally improved steadily (with notable progress in the phase of student autonomous work), but her conceptualizations followed a relatively downward path. It might be the case that her conceptualization was evolving but we did not have the means to observe it, without Pina externalizing her thinking. Also, the evolution of her conceptualization of the interplay might suggest that her ideas about differentiation were also evolving. Specifically, at the beginning of the program, she had some misconceptions about differentiation, such as giving less and easier work to less advanced students and trying to camouflage differentiated outcomes so that differentiation is not noticed by students. Although over time she did not provide any further evidence of how her conceptualizations of differentiation had changed, as she gained more experience, she implemented various strategies and practices of differentiation without disguising them from students and without giving less or easier work to less advanced students. This implies that Pina had developed a more accurate and nuanced understanding of differentiation, and she was implementing it more effectively in her practice.

Overall, the gradual decrease in conceptual emphasis on the two axes in isolation (i.e., cognitive activation and differentiation) suggests that Pina prioritized different aspects of teaching regarding the interplay of the two axes based on the needs of her students; her existing knowledge and needs; and her classroom realities. Furthermore, the gradual upward shift in her teaching practice may indicate that the two axes and their interplay operate *synergistically*, promoting one another in actual practice. This stems from the fact that Pina did not place equal conceptual emphasis on each of the two axes of cognitive activation and differentiation but saw steady improvement in all axes and their interplay. In addition, her strong initial knowledge of mathematical content and cognitive activation, as well as her more interpretive stance and growth mindset, may have helped her to process and implement the ideas more effectively, and emerge as a "successful" case of the program (in terms of her teaching).

The program helped Pina to build upon her existing knowledge and teaching skills to some extent and to address some challenges she faced, with some of them unresolved, showing that there was still room for improvement. It is important to point out that these nuances were not necessarily revealed by quantitative analysis and hence the importance of qualitative analysis becomes clearer. Specifically, the quantitative analysis did not help us surface these challenges, but these were real and important, illustrating how complex it is to support teachers' work on cognitive activation and differentiation. In sum, Pina's learning trajectory and the coherence (or lack thereof) between her conceptualizations and her teaching practice can provide insights into how PLD programs focusing on cognitive activation and differentiation can support teachers towards this end.

In sum, Pina emerges as a compelling case of relatively stable professional growth. Initially, she viewed cognitive activation and differentiation as separate axes of teaching. Over time, she shifted towards a more integrated approach, focusing on the interplay between these axes. In addition, her teaching development across the axes of cognitive activation, differentiation, and their interplay showcases the cumulative benefits of ongoing PLD. Her teaching constituted a set of teaching ideas, which, like in a musical symphony, each practice harmonized with the next, creating an effective outcome. Despite her successes, her experience also speaks to the challenges inherent in implementing new teaching practices. Her case illustrates the importance of persistence and resilience in the face of challenges, with some being overcome and others persisting as areas for further growth. The next chapter delves into the second case, Kate, who, in contrast to Pina, was trying to keep all her balls in the air over the PLD program.

CHAPTER 6. THE CASE OF KATE: HIGHS AND LOWS, EBBS AND FLOWS

Kate is the second case. Although she gradually built her conceptual pathways around cognitive activation, differentiation, and their interplay, when synchronously experimenting with various ideas, she either abandoned them to experiment with other ideas presented in the sessions or her experimentation with these ideas faded in subsequent lessons. Hence, the gradual shift in her conceptualizations was not very well aligned with the ebbed and flowed teaching practice observed across her four videotaped lessons. This trend of trying to juggle multiple ideas at the same time while attending to both cognitive activation and differentiation was more apparent in the detailed qualitative analysis of her practice. In brief, Kate gained certain aspects from the PLD program whilst also facing certain challenges inherent in concurrently attending to both axes.

6.1 Kate's Background

Kate was an experienced teacher who was fan of differentiation and the idea of asynchronous work. During the EDUCATE PLD program, she was in her second year of teaching Grade 6, having taught for 18 years most of which were in lower grades (14 out of 18 years). Apart from her bachelor's degree in Elementary Education, Kate held a master's degree in Language and Cultural Education.

Her decision to participate in the program was informed by her previous engagement with differentiation issues alongside her curiosity to understand what cognitive activation is:

Kate: As soon as I saw the [EDUCATE] announcement, it sparked my interest because I have dealt with issues of differentiation in teaching in the past, so I thought it might help me to deepen my knowledge a bit.
Interviewer: So, for differentiation, not so much for cognitive activation.
Kate: Yes, for differentiation. Perhaps I hadn't understood the concept of cognitive activation well and I wanted to see what we mean in practice by this term.
Interviewer: So, to see a different perspective on differentiation.
Kate: Yes. (EPI, lines 21-32)

In the past, she participated in a two-year-long school-based PLD program focusing on theoretical aspects, strategies, and techniques of differentiation, combined with enacting differentiated lessons co-planned with teacher educators. "The first year was an introduction to what differentiation is, techniques, and strategies. The subsequent year was about planning and implementing lessons" (EPI, lines 34-40). Her enthusiasm over her previous PLD experience was obvious in her final interview, during which she expressed that "the idea of differentiation convinced" her by the time it was first introduced to her, and she kept her focus on differentiation issues even after the completion of that PLD program (EPI, line 224). In a nutshell, Kate entered the program with a strong focus and background in differentiation ("I was quite involved with differentiation, both theoretically and practically" EPI, lines 48-49).

6.2 Reaction Level: Kate's Evaluation of the EDUCATE PLD Program

In her end-of-program interview, Kate's attitude towards attending the EDUCATE PLD program reflected a positive disposition. She expressed that she "consistently attended the sessions with a sense of pleasure and eagerness" (EPI, lines 479-480). Her assertion, "Honestly! I was coming very pleasantly [to the sessions]" underscored the intrinsic value and enjoyment she derived from participating in them (EPI, line 482).

Pressed to mention negative aspects of the program, Kate did not identify any, stating that "everything [we worked with] was useful […] I never felt that I lost my time" (EPI, lines 449; 475). She enjoyed that the program allowed her to explore and discuss topics that were previously overlooked or undervalued in her teaching practice (EPI, lines 449-451). Also, she emphasized its role in "refreshing and deepening" her understanding of various teaching practices, especially in differentiation and challenging students in mathematics: "Although I had some theoretical foundations [on differentiation], I significantly and substantially deepened my understanding of teaching issues, addressing various challenges I had in cognitively activating students, and in dealing with a difficult subject [i.e., mathematics] in various ways." (EPI, lines 52-59)

Another significant aspect of the program for Kate was the focus on designing mathematics lessons and analyzing tasks (EPI, lines 451-454). Also, the printed EDUCATE materials offered the practicality and convenience for revisiting and consulting the information as needed (EPI, lines 488-490). Moreover, she valued the collaborative aspect of the program, which involved planning and designing with other teachers (EPI, lines 460-461). This process of exchanging ideas and discussing with peers offered her opportunities to learn from others' perspectives and practices (EPI, lines 461-465). The thought-provoking discussions of the video-club sessions made her "think and improve" (EPI, lines 499-502). The small group size facilitated more focused and meaningful exchanges, especially between teachers who were teaching the same grade (EPI, lines 535-541). Kate appreciated "the warm and supportive atmosphere of the group, which fostered a safe environment for open and honest communication, vulnerability, and the sharing of challenges and mistakes without the fear of judgment" (EPI, lines 541-545).

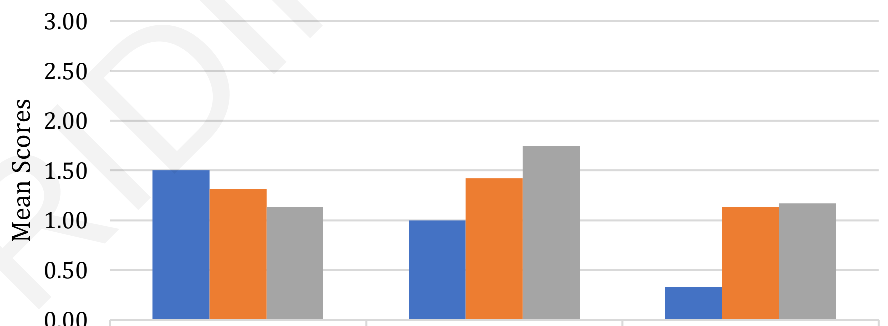
In general, she described the PLD program as "a resource for teachers seeking self-improvement beyond their formal education" (EPI, lines 388-389). She emphasized the program's role in providing updated research on the latest theories and developments in education (EPI, lines 931-395) and supporting her "to manage anything that arises in the classroom—from the most unlikely mistakes to the most expected ones" (lines 400-401).

6.3 Results Level: Kate's Teaching Performance in Her Concluding Lesson

Figure 30 illustrates Kate's performance in cognitive activation differentiation, and the interplay of the two axes during her concluding lesson, across the three lesson phases.

Figure 30

Kate's Performance in Cognitive Activation, Differentiation, and Their Interplay in Each Phase of Her Concluding Lesson



	Cognitive Activation	Differentiation	Interplay
TL	1.50	1.00	0.33
SAW	1.31	1.42	1.13
WCI	1.13	1.75	1.17

Note.
TL: Task Launching; SAW: Student Autonomous Work; WCI: Whole-Class Interactions.

The graph portrays Kate's performance across the axes as low to moderate, with mean scores fluctuating near the median of the 0 to 3 scale. Kate's pinnacle of performance in differentiation was observed during the phases of student autonomous work and whole-class interactions—her only score over the median of the scale (1.75), coupled with her adeptness in fostering cognitive activation during the task launching (and somewhat during the student autonomous work) phase—scoring the median of the scale (1.50). In contrast, her performance was most challenged in the interplay of the two axes, especially during the task-launching phase, indicating areas ripe for pedagogical refinement (less than 1.20). A notable observation is that most scores predominantly reside below the scale's midpoint, indicating a pronounced potential for improvement by the program's culmination.

Considering the observed quantitative outcomes, one might posit that while Kate experimented with some practices from each axis, the persistent low to mid scores indicate that the PLD program's impact was perhaps not fully optimized for her specific developmental needs. This suggests that the PLD program might require further customization or additional support mechanisms to address the challenges Kate faced.

In all, the graph depicts Kate as a case who did not significantly benefit from the program. However, this graph does not allow for discerning any patterns of evolution in Kate's engagement with the PLD program. A comparative analysis of mean scores across various timepoints might facilitate a different delineation of her teaching behavior. Such an analysis would illuminate significant shifts or consistencies in her teaching, serving as a barometer for the PLD program's effectiveness in fostering her professional growth. An in-depth examination of Kate's teaching practice *across* her lessons is presented next.

6.4 Behavior Level: Evolution of Kate's Teaching Performance

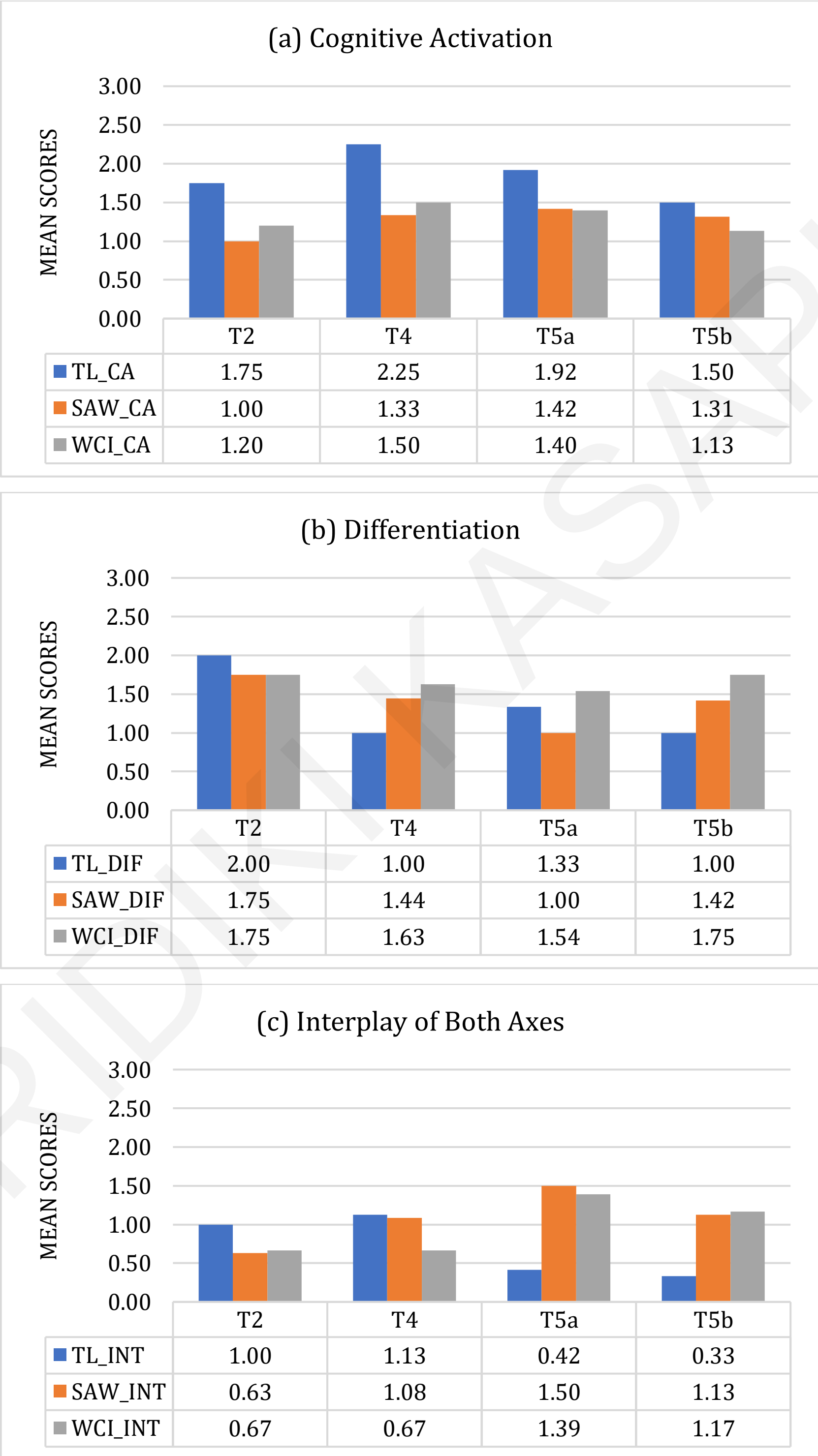
This section delves into the evolution of Kate's teaching performance through both quantitative and qualitative lenses. The quantitative analysis draws on aggregated mean scores from her videotaped lessons, offering a structured view of her progress in cognitive activation, differentiation, and their interplay. Complementarily, the qualitative analysis sheds light on her experimentation with certain teaching practices, highlighting ebbs and flows in their implementation. Together, these analyses paint a comprehensive picture of Kate's teaching behavior during the PLD program.

6.4.1 Quantitative Analysis

Figure 31 depicts a not-so-promising picture in terms of improvements in Kate's teaching across the four lessons.

Figure 31

Kate's Performance in (A) Cognitive Activation, (B) Differentiation, and (C) Their Interplay, Per Phase Across Her Lessons



Note.
TL: Task Launching; SAW: Student Autonomous Work; WCI: Whole-Class Interactions.

The mean scores for cognitive activation showed fluctuations across different teaching phases. There was a notable peak in the mean score for task launching during Kate's second lesson (2.25), which was the highest score across all axes and phases.

The scores for differentiation were higher overall compared to cognitive activation, with all scores starting above the median of 1.50 in her first lesson. The score of differentiation during whole-class interactions not only showed growth in her final lesson compared to other phases but also was the only phase that ended on an upward trend, indicating improvement. The mean scores for the interplay of cognitive activation and differentiation were generally low across all phases, with none reaching the median score of 1.50. The interplay during student autonomous work showed the most improvement from the second to the third lesson, which could imply an increased effectiveness in this respect, but this was not sustained in the concluding lesson.

Despite the growth in the quality of Kate's work in some respects, in her last lesson, her mean performance was near or below the scale median (1.50) in all three axes. In her first lesson, the teacher performed above average on differentiation (around 2.00, see Figure 31b), at the median on cognitive activation (around 1.50, see Figure 31a), and significantly below the median on the interplay (below 1.00, see Figure 31c). In her subsequent lessons, little improvement in cognitive activation (L2 and L3) was observed in contrast to her mean performance on differentiation, which gradually dropped even below average. In addition, her mean performance on the interplay of the two axes in the phases of student autonomous work and whole-class interactions (around 1.00) and the decrease in her mean performance in task launching (around 0.50) illustrated significant room for improvement (see Figure 31c).

In sum, the quantitative analysis gave the impression of a teacher who made almost no progress in promoting the dual goal of cognitive activation and differentiation or whose performance fluctuated across her lessons. The trends suggest variability in Kate's performance across different teaching phases and axes. While there were peaks in performance, they were not sustained, and in some cases, there was a downward trend, indicating areas where improvement could be targeted.

6.4.2 Qualitative Analysis

From the qualitative analysis of Kate's lessons, it is apparent that she experimented with different ideas while trying to address the dual goal of cognitive activation and differentiation (as the number of check marks in Table 13 suggests). A pattern that is easily discernible in Table 13 pertains to the *relative lack of stability* regarding the quality of her experimentation across the four lessons. At first glance, the case of Kate illustrates a teacher with ebbs and flows in her experimentation who seemingly gained not much from the PLD program—a finding that is in line with the quantitative results.

The shading of the cells of cognitive activation and the interplay of the two axes in Table 13 suggests a general pattern of inconsistency in the frequency and the quality of Kate's experimentation, shifting from light green to some bright greens and back to light green or even reds. In particular, she implemented most practices that promote either cognitive activation or the interplay more *sporadically* (e.g., "discussing key mathematical ideas to the task and/or the goal of the lesson without reducing the demand" or "using flexible grouping" appears only once in L2 and L4, respectively, see Table 13) or *inconsistently* (e.g., "using enablers" appears in her last three lessons with variations in terms of quality as evidenced by the successive changes in the color shades—moving from light green to red and back to light green).

Admittedly, the differentiation practices with which she had already been familiar from her previous PLD experience were more systematically implemented and in fairly good quality (see Table 13). The table also reveals that there were no substantial shifts in the quality and frequency of implementation of differentiation practices. In all, the findings suggest that Kate seems to have experienced specific gains from the PLD program (e.g., "using extenders" gradually improved—moving from light green to bright green color) despite also facing certain challenges (e.g., the quality of "bringing the class to the plenary at appropriate checkpoints" worsened—going from light green to red).

Table 13

Kate's Experimentation with Practices That Promote (A) Cognitive Activation, (B) Differentiation, and (C) Their Interplay, Across Her Lessons

A. COGNITIVE ACTIVATION				
Task Launching	**L1**	**L2**	**L3**	**L4**
• Selecting mathematically challenging tasks	✓	✓	✓	✓
• Maintaining the cognitive demands of the task as presented to students during task launching	✓	✓	✓	✓
• Asking students to read the task instructions and implement the think-pair-share strategy	✓		✓	✓
• Discussing (key) mathematical ideas (to the task and/or the goal of the lesson) without reducing the demand		✓		
• Asking students to explain/make sense of mathematical symbols, sentences, or representations				
• Handling unexpected student interference which could probably steal students' thinking				
• Relaunching the task ending up initiating a whole-class discussion				
Student Autonomous Work	**L1**	**L2**	**L3**	**L4**
• Providing mathematical prompts to students to help them make some progress on the task and take up the challenge	✓	✓	✓	✓
• Pressing students for explanation/meaning, for making conceptual connections, or engaging them in mathematical reasoning		✓	✓	✓
• Posing mostly closed-ended questions	✓			
• Monitoring students' work and being more directive than needed	✓	✓		
• Allowing another student to steal his pair's thinking				
• Asking for explanations that focus on describing the procedure used				
• Telling the students precisely *how* to work (step-by-step) on the task				
• Pointing out errors in students' work and remediating with procedures				
Whole-Class Interactions	**L1**	**L2**	**L3**	**L4**
• Eliciting student reasoning or providing opportunities for mathematical reasoning and meaning-making without reducing the challenge	✓	✓	✓	✓
• Rephrasing student ideas to address key mathematical ideas related to the task at hand (in interaction with students)	✓	✓	✓	✓
• Synthesizing and extending student contributions to address key mathematical ideas related to the task at hand (in interaction with students)	✓	✓	✓	✓
• Presenting and discussing the shared solutions	✓	✓	✓	✓
• Comparing and evaluating the shared solutions	✓	✓	✓	✓
• Bringing the class to the plenary at appropriate checkpoints	✓	✓	✓	
• Having a clear direction during the discussion				
• Extending the discussion by posing a question that is more challenging than the task-at-hand				
• Handling unexpected student solutions		✓		✓
• Handling unexpected student interference which could probably steal the thinking	✓	✓		
• Handling alternative conceptions around mathematical ideas				
• Providing directive hints or ready-made answers	✓			
• Enacting IRE interactions when checking in plenary	✓			
• Introducing important mathematical ideas very early		✓		
B. DIFFERENTIATION				
Task Launching	**L1**	**L2**	**L3**	**L4**
• Having the resources or materials available to be used by the students	✓	✓	✓	✓
• Providing a clear way of working on the task	✓	✓	✓	✓

• Providing materials to students and not ensuring that they have understood how they can be used while working on the task				
Student Autonomous Work	**L1**	**L2**	**L3**	**L4**
• Implementing asynchronous work	✓	✓	✓	✓
• Using learning aids	✓	✓	✓	
• Using an entry card	✓	✓		
• Monitoring students' work and more explicitly assessing their needs	✓	✓	✓	✓
• Using a tic-tac-toe board as a final assessment activity	✓		✓	
• Using an exit card		✓		✓
• Using anchoring activities	✓	✓	✓	✓
• Encouraging multiple expressions of content, process, and/or product			✓	✓
• Grouping students according to their proficiency levels				
Whole-Class Interactions	**L1**	**L2**	**L3**	**L4**
• Highlighting important mathematical ideas	✓	✓	✓	✓
• Sequencing student solutions in a reasonable progression		✓	✓	
• Allowing students to start explaining any method they want				✓
THE INTERPLAY OF COGNITIVE ACTIVATION AND DIFFERENTIATION				
Task Launching	**L1**	**L2**	**L3**	**L4**
• Explaining potentially unfamiliar non-mathematical aspects of the wording of the task or difficult words (context- or scenario-wise)		✓		
• Clarifying mathematical aspects of the task		✓		
• Activating relevant existing mathematical knowledge and strategies	✓	✓	✓	✓
• Posing questions that indicate the level of support that students need in order to engage in the task without reducing the level of challenge				
• Spending no time clarifying the task instructions during task launching	✓		✓	✓
Student Autonomous Work	**L1**	**L2**	**L3**	**L4**
• Directing different types of questions to different (groups) of students	✓	✓	✓	✓
• Circulating within all the groups attempting to attend to all students				✓
• Using enabler(s)		✓	✓	✓
• Using extender(s)		✓	✓	✓
• Sharing a general strategy devised by a student group with the rest of the class, to support them make progress on the task and taking up the challenge			✓	
• Using flexible grouping				✓
• Maintaining the demands of more advanced students and trivializing the thinking of less advanced students	✓			
• Devoting more time to scaffolding students who are facing difficulties	✓	✓	✓	
• Not establishing a routine for what the early finishers could do once they complete the main task	✓	✓		
• Facing difficulties in supporting less-advanced students to make progress on the task	✓	✓	✓	
• Directing the exact same questions to all (group of) students				
Whole-Class Interactions	**L1**	**L2**	**L3**	**L4**
• Using incorrect or incomplete student solutions as resources for all students' learning	✓	✓	✓	
• Holding students accountable to attend to their classmates' thinking	✓	✓	✓	✓

Notes.
1. L#: Lesson ordinal number
2. The check mark illustrates that the teacher experimented to some extent with the practice in that lesson.
3. The color scale demonstrates the quality and frequency of the implementation (see Section 3.7 in Chapter 3).
4. Bright green: the practice was implemented in high quality and the demand was not reduced; Light green: it was implemented in medium quality and the demand was somehow reduced; Red: it was implemented in low quality and the demand was reduced; Grey: this practice hinders the particular axis.

A closer examination of Table 13 points to three main themes, regarding the changes or lack thereof in Kate's teaching: (a) *trying to juggle multiple balls at the same time while attending to both cognitive activation and differentiation*; (b) *consolidating certain aspects of concurrently attending to both cognitive activation and differentiation; and (c) struggling with certain ideas inherent in concurrently attending to both axes*. In what follows, the findings are discussed in light of these three themes and are accompanied by indicative examples.

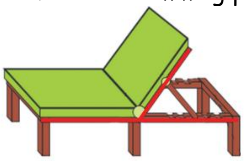
Trying to Juggle Multiple Balls at The Same Time, While Attending to Both Cognitive Activation and Differentiation. The first theme pertained to Kate's experimentation with various ideas simultaneously, something that appears to have led her to two different subthemes during teaching: (a) *in some cases, she enacted a new practice in one of her videotaped lessons and then abandoned it in the next lessons*; (b) *in other instances, she implemented a new practice in a lesson, and in every subsequent lesson her experimentation with this practice was fading as she was implementing new practices*. These behaviors appeared to have been used as coping mechanisms by Kate, as she was apparently trying to manage the cognitive load embedded in the complexity of dealing simultaneously with multiple teaching ideas (although these ideas were introduced incrementally during the PLD program, see Table 7 in Chapter 3).

Enacting a new practice in one lesson and then abandoning it in the next lesson. A key example regarding the first subtheme was observed in the phase of task launching. Particularly, in L1 (focusing on complementary and supplementary angles, see Figure 32), Kate spent no time discussing the task instructions with her students, asking them to "read the instructions and then work on the task individually; once you finish, check your answers in pairs" (L1, min 48:50). Hence, the demands were not lessened at any time (see Table 13). By not having potentially unfamiliar (non-)mathematical aspects of the task being explained to them, students did not understand the context and the mathematical ideas of the task to render it more accessible. This led to student difficulties during autonomous work with a notable number of students not being able to decontextualize the task and identify the angles created by the backrest and the base of the sunbed.

Figure 32

Kate's Main Task in Lesson 1

The backrest of the sunbeds used on beaches and swimming pools moves to different positions.



Photini drew the diagrams below, to show angles that are formed when the backrest is in different positions.

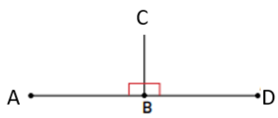


Diagram 1

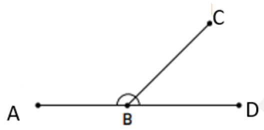


Diagram 2

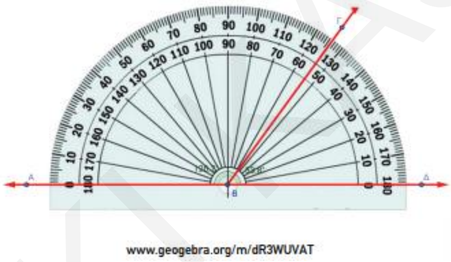
(a) Describe how the angles $A\hat{B}C$ and $D\hat{B}C$ change, as the backrest moves in different positions.

____

____

____

(b) Photini used an application to measure the angles $A\hat{B}C$ and $D\hat{B}C$, when the backrest was in different positions.



(c) Move point C at different points using the application and complete the table.

Measure of angle $A\hat{B}C$	Measure of angle $D\hat{B}C$	Sum of the two angles

(d) What do you observe?

(e) Write the relationship that connects the two angles.

(Mathematics Curriculum, 2016, Grade 6, Unit 4, pp. 12-13, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

During VCS3, a discussion revolved around the importance of having students explicate to a peer what the task is asking them to do and discussing key mathematical ideas to the task without reducing the demands during task launching. Drawing on this discussion, Kate experimented with this idea in L2 (focusing on calculating the sum of the polygon angles, see Figure 33), which resulted in a notably different quality of her work during this lesson phase compared to what was observed in the previous lesson:

Kate: [Student 1], please, *explain the task instruction in your own words*… Don't try to solve it yet! Let's all hear the shared ideas to ensure that you have understood the task correctly.
Student 1: We must divide the polygon into triangles that are not visible yet.
Kate: Oh! You mean triangles that are not yet formed. So, we must partition the polygons into triangles, right?
All students: Yes.
Kate: *What does the phrase "non-overlapping triangles" mean?*
Student 2: Each triangle must be next to the other, not cover the other triangles.
Kate: Next to each other, not one on top of the other. So, they should not overlap. Nice. This is *the second important detail* that you must consider while working on the task. And *the third important detail* is… how are we going to achieve this?
Student 3: By bringing all the diagonals from a vertex of the polygon.
Kate: *What is a diagonal?*
Student 4: A straight line.
Kate: Good. Is this a diagonal? [she *draws a counterexample of a diagonal* starting from and ending on two opposite sides]
Student 1: No.
Kate: Why?
Student 5: It should start from a vertex and end at another vertex.
Kate: Nice. Is the task clear to you? [most students seem to agree] Good. Start working on the task. (L2, min 2:50, emphasis added)

In this excerpt, Kate attempted to elicit student understanding and engage them in both the context and the mathematical ideas of the task: (a) she was holding students accountable for attending to and understanding their classmates' sharing so that all students would understand the task (lines 1-3); (b) she was aware of key important ideas that could pose challenges for students when reading the task instructions ("important details") and spent time trying to enable students to identify them (lines 10-13); and (c) she supported students to develop a common language of what the task was asking, using representations (lines 17-18).[41] In her subsequent lessons (L3 and L4), however, she no longer experimented with this idea; rather, she returned to her familiar practice of asking students to implement the think-pair-share strategy as in L1 ("Please, read the task instructions and work individually on the task. Then, if you wish, discuss your work in pairs", L4, min 4:00) and placed more emphasis on the use of enablers or extenders (see Figure 33), which were introduced in VCS5, as discussed later.

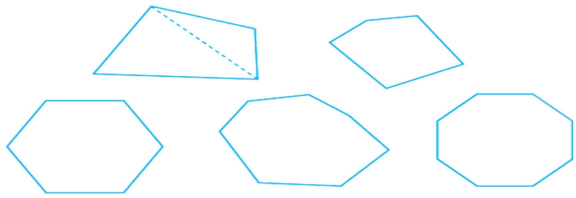
[41] Another possible factor that encouraged Kate to try this practice was the close collaboration she had developed with her colleague, Georgia, who taught at the same grade level. The two teacher participants happened to be videotaping a lesson on the same content and their lesson plans shared many similarities. Often, Kate and Georgia shared and compared their ideas and experiences from their daily teaching practice, during and after the program's conclusion (EPI, lines 548-550).

Figure 33

A Representation of (A) The Main Challenging Task Used in Kate's Second Lesson Along With (B) Its Enabler and (C) Extenders

(A) <u>Main task</u>

(a) Divide each polygon into non-overlapping triangles, by bringing all the possible diagonals from one vertex of the polygon.



(b) Complete the table.

Polygon	Number of polygon sides	Number of triangles in which the polygon is divided	The sum of polygon angles
Quadrilateral	*4*	*2*	$2 \cdot 180^{o} = 360^{o}$
Pentagon			
Hexagon			
Heptagon			
Octagon			
Decagon			
n-gon			

(c) Compare the number of sides of a polygon with the number of triangles in which the polygon is divided. What do you observe?

(d) How can you calculate the sum of the angles of any polygon? Write down a general rule.

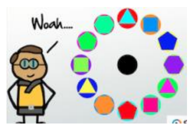
(Mathematics Curriculum, 2016, Grade 6, Unit 6, p. 15, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

(B) <u>Enabler</u>

Divide each polygon into non-overlapping triangles, by bringing all the possible diagonals from one angle of the polygon. You can start from the vertex marked below.

(c) <u>Extenders</u>

Find the number of the sides of a polygon whose angles' sum is 4140°.



ii. Solve the problem:

Mr. John will build a pool in the shape of a polygon in his cottage. If the sum of the angles of the pool's bottom is 1980°, what does the bottom of the pool look like?

Implementing a new practice in a decrescendo way. The second subtheme suggested that Kate was implementing a new practice in a lesson and in every subsequent lesson, her experimentation with this practice was fading as she was implementing new practices. An example pertaining to this subtheme was noticed in the phase of whole-class interactions (see Table 13). During this phase in her first lesson (on complementary and supplementary angles, see Figure 32), Kate posed close-ended questions aiming at correcting student answers or errors in the tasks. For example, the students had worked on measuring complementary angles and calculating their sums during autonomous work, and Kate asked them to share their answers:

Kate: What kind of angle is that?
Student 1: Acute.
Kate: [talking to the rest of the class:] Do you agree? [then, talking to Student 1:] Are you talking about the red angle?
Student 1: I have changed my mind. It is obtuse.
Kate: How many degrees is it? Read the protractor clockwise to see where the side of the angle crosses the number scale.
Student 1: 125°.
Kate: You are very close. What kind of angle is the second angle?
Student 2: Acute.
Kate: How many degrees is it?
Student 3: 55°. Miss, together they are 180°!
Kate: *Good observation. We will talk about it in a minute*. Let's see the measure of these angles. How many degrees is the first angle? [The class continues to present the measures of the remaining angles in the same way for about 10 minutes and then Kate asks:] What are the sums of the angle pairs?
Student 3: All the sums are 180°.
Kate: Did you all make the same observation? *Listen to me. The sum of a pair of complementary angles is 180°. Write your observations in your textbook.*
Student 4: They form a straight angle.
Kate: Bravo! Now read the theory [elaboration of mathematical ideas as presented in the textbook] and discuss the text in pairs. Explain it to your peer. (L1, min 23:41, emphasis added)

In this vignette, the teacher enacted Initiate-Response-Evaluate (IRE)[42] interactions when checking their understanding in plenary and superficially treated her students' errors (lines 1-5). At the same time, although she tried to prevent a student from doing the thinking of his classmates (lines 12-13), she ended up over-guiding students (lines 6-7) and providing directive hints (lines 16-17) and ready-made answers (line 19-21).

A major part of VCS4 focused on discussing effective teaching practices for promoting both cognitive activation and differentiation during the phase of whole-class interactions, such as deliberate selecting, sharing, and sequencing student

[42] The IRE model is a *repeated* teacher-led discourse pattern in teacher-student interactions, in which the teacher *initiates* a question, a student *responds*, and the teacher then *evaluates* the response, barely allowing students the opportunity to express and expand their thinking (Cazden, 2001).

solutions, revoicing, explaining, or clarifying student ideas, and connecting multiple solutions. Drawing on what was discussed during this session, Kate extensively experimented with these ideas in L2 (i.e., the lesson focusing on generating a rule for calculating the sum of the angles of any polygon). While monitoring student autonomous work, Kate identified two approaches used by the students for partitioning the polygons into non-overlapping triangles (see Figure 33). It was obvious that she decided to share the selected solutions in a specific progressive sequence that made pedagogical sense, starting from the simplest (partitioning the polygons by bringing diagonals starting from the same vertex) to the more complex and unexpected (partitioning the polygons by bringing diagonals starting from different vertices). We enter the episode when the second way for partitioning the pentagon was presented:

Kate: [The first approach for partitioning the triangles had already been shared and discussed and was still presented on the smartboard.] While circulating in the class, I noticed that your biggest difficulty was with partitioning the polygons. How did you partition the hexagon, [Student 6]?
Student 6: My diagonals start from different vertices…
Kate: [Draws Student 6's method on the smartboard] Hmm. They do not start from the same vertex. [Talking to the rest of the class:] Is this way of partitioning the polygon different than the one previously shared?
All students: Yes.
Kate: Is it correct?
All students: [some students agree whilst others disagree]
Kate: Yes, or no? Do the triangles overlap? Does the number of triangles differ?
Student 1: It is correct. We have the same number of non-overlapping triangles.
Kate: How many triangles do you see with each method?
All students: Four.
Kate: Hence both ideas for partitioning the polygons are correct. (L2, min 18:52)

In this episode, Kate utilized multiple ideas for partitioning the polygons as resources for all students' learning. Contrary to the previous lesson, both solutions were not only presented, but they were also discussed, and compared.

The experimentation with this practice began to fade in the subsequent lessons. Kate attempted to develop some connections between multiple methods but very superficially, compared to L2. For example, in L3 (focusing on multiplying an integer by a fraction, and vice versa, see Figure 34), two solutions for solving problems involving the multiplication of *an integer by a fraction* were presented and explained:

multiplication (3·2/5), and some links were made to repeated addition (2/5+2/5+2/5). With minimal interaction with the students, the teacher briefly mentioned that the method of multiplication is a more practical and quick solution method than repeated addition, especially when an integer is a big number. When the solution to the second problem involving the multiplication of *a fraction by an integer* was presented (see Figure 33), no connection was made between the two types of problems and their solutions, which was essential for achieving L3's core goals. Similarly, in L4 (focusing on mixed-number multiplication, see Figure 34), the three methods for mixed-number multiplication included in the task were presented and explained; students were asked to note which method was more practical, easy, and convenient for them. Kate briefly claimed that the most practical method is different for each student without allowing students to explain their arguments. It seems that she gradually started to reduce the focus on the interactions taking place in plenary and emphasized improving practices related to the phase of autonomous work (see Table 13).

Consolidating certain teaching ideas and practices. This theme illustrated that the experimentation helped Kate consolidate certain ideas, as evidenced by three specific examples of practices improved over time: (a) *posing appropriate questions and prompts to help students make progress on the task and take up the challenge; make conceptual connections; and reason mathematically*; (b) *developing and using extenders*; and (c) *circulating within all the groups, attempting to attend to all students*. Noteworthy, the qualitative analysis revealed that consolidation needed consistent and continuous experimentation with ideas in practice, as explained in the next section.

Posing appropriate questions and prompts. Firstly, Kate was observed refining different practices across her lessons, such as improving her questioning and prompting to help students make progress on the task and take up the challenge, make conceptual connections, and reason mathematically. Despite still using a few procedural questions, drawing from a discussion of VCS3 regarding the importance of posing different scaffolding questions and their characteristics for eliciting, connecting, evaluating, clarifying, and extending student thinking, Kate embedded more open-ended questions in her interactions with students (e.g., "How did you work to complete the table? Read the table carefully to identify a pattern" L2, min 8:20; "Some triangles are overlapping. What can you do to fix this?", L2, min 11:25; "What is the relationship between the number of the polygon sides and the number of triangles?" L2, min 12:10) and refrained from giving ready-made answers/hints. This notable change supported student participation since Kate did not rush to respond to their difficulties. Listed below is an indicative example of such an interaction with a student on formulating a general rule for calculating the sum of the angles of any polygon (see Figure 33):

Kate: [Reads student's answer out loud:] "They have a difference of two". You have grasped the idea, but you need to formulate it more clearly.
Student: I wrote an example: "6 – 2". What else can I write about their difference?
Kate: What did you subtract here?
Student: 6 minus 2…
Kate: What does number 6 represent in your example?
Student: The number of the polygon sides.
Kate: What does number 4 represent?
Student: The triangles formed in the polygon.
Kate: How did you get the result?
Student: It is the result of the hexagon.
Kate: I still don't get it… What is the rule for calculating the sum of the angles of any polygon?
Student: First, I find the number of non-overlapping triangles the polygon can be partitioned into. Then I subtract this number from the number of the polygon sides and multiply the result by 180.
Kate: [noticing that his answer was partly correct:] What do you subtract? Think about it. (L2, min 32:48)

Kate's questioning was intended to check the student's understanding of the underlying mathematical ideas and elicit explanations for how the rule was formulated (lines 11, 13-14). She listened to the student and interpreted his claims and questions (lines 3-6, 18), to identify his errors and difficulties (lines 13-14). This approach was typical in L2 when students were working on the most demanding task parts.

In L3, Kate further polished her questioning techniques to stimulate students' mathematical thinking. The questions posed during autonomous work required higher cognitive thinking from the students (e.g., "How did you work on that? Explain your thinking" L3, min 22:00; "What question should be posed to determine that a subtraction sentence is the most appropriate choice?" L3, min 22:09; "Why did you reject this mathematical sentence?" L3, min 22:53; "Show how the selected mathematical sentence fits the problem" L3, min 31:54). For struggling students, she also incorporated some closed-ended questions to help them give meaning to the symbols and contextualize the mathematical sentence (e.g., "What does this mathematical expression mean [8 ÷ ¼]? Can you show me what was divided?" L3, min 29:14). Similar patterns of teacher-student interactions were observed in L4, somehow less frequently than in L3.

Developing and using extenders. Kate gradually improved the design and use of extenders in her lessons. As with enablers, she experimented with multiple parameters, such as the strategy used for developing them; the number of extenders; and the combination of using extenders with flexible grouping. In particular, she developed a series of extenders for each lesson, experimenting with various strategies, including using reverse thinking (in L2, see Figure 33); completing a problem's question and selecting the suitable mathematical sentence for solving it; formulating their own problems to given mathematical sentences (in L3, see Figure 34); problem-solving and posing; writing equations; and making connections with other mathematical concepts previously taught (in L4, see Figure 35). Extenders were relevant to the lesson objectives, always extending students' thinking a step further.

Regarding their enactment, Kate was providing extenders to her students who had finished the main task early, showing an improvement in how she interacted with them from lesson to lesson. In L2, the extender was provided to five early finishers after Kate checked that they completed their work on the main task, without substantially interacting with them and asking for explanations ("Bravo, now work on this", L2, min 42:30) or after students asked for it ("Miss, can I have a worksheet [with the extender]?" L2, min 48:28). However, the teacher never returned to the students who had received an extender to check their progress on it. Similarly, in L3, the teacher provided extenders to some students who had correctly solved the second task, after ensuring that they had successfully completed the main task but without asking for explanations (L3, min 26:11). In contrast, in L3 and mainly in L4, Kate was seen asking students to *explain* their thinking around the main task, before providing them with extender(s).

The following excerpt from Kate's L3 is indicative. She circulated and monitored students' work. Upon noticing a student who finished his work on the assigned task, she asked him to explain his thinking. The student replied, apparently explaining his thinking [which was not very audible] and Kate continued, "You're awesome! You have even moved a step farther! Well done! I will get you something else to work on" (L3, min 26:09). Moreover, while working on the extender in L4, she asked a group of students to work together on a multiplication that included powers (L4, min 26:55). The combination of using extenders with flexible grouping showed that the goal of her extenders was indeed to extend the thinking for her students and promote student collaborative skills. This incident suggested that she was finally using the extenders as an opportunity for students to learn and extend their thinking.

Reflecting on the use of extenders, Kate was excited about the way the extenders had worked in her lessons and how her students had received them:

> I think the [idea of] extender[s] was particularly useful for my students. They were like 'Miss, bring me one extender, I'm done [with the main task]' […] And the next day, they were like 'Miss, can we show you how we have worked on it?' and they were coming to me to explain their thinking. (VCS7, lines 420-426)

She realized that when preparing and employing extenders, the lesson and its subsequent one ran smoothly since students were motivated to explain their thinking. Her developed extenders were on target and also headed toward transformative knowledge which was considered the main lesson goal for the next lesson. The direct links of the extenders to transformative knowledge showed that she successfully sequenced the mathematical challenge to attend to the needs of more advanced students.

Circulating within all the groups, attempting to attend to all students. The third example pertained to gradually shifting from devoting more time to scaffolding less advanced students to circulating within all student groups, attempting to attend to all students. For example, in L1, Kate would praise students who had solved the task correctly (e.g., "Keep up the good work, bravo!", L1, min 58:42) and ask them to work on something more challenging or an anchoring activity, after checking their work mostly in proforma ways (e.g., reading their answers in their textbooks, without asking students to explain their thinking or justify their answers). In contrast, more support was offered to less advanced students, usually ending up doing their thinking: Kate would stop by confirming that the student had read and understood the task instructions by posing a sequence of closed-ended questions, then provide some clarifications herself, and give directive hints on how to work on the task. A characteristic example of how she handled a situation with a student who was facing difficulties is presented next. The student raised her hand because she could not understand the task instructions; without asking for explanations, Kate headed toward her saying:

> Pay attention to the letters [at the three points of each angle] and identify the two angles. What does this line excerpt indicate? [Kate rushes to respond to her own question:] It indicates the backrest of the sunbed leaning to the right. So,

one angle has increased... what about the other angle? [Then, talking to the whole class:] What happens to the pair of angles when a change is made to the inclination of one angle? Work on it [and she heads to other students]. (L1, min 15:45)

From this vignette, it is obvious that for students who faced difficulties, Kate largely reduced the cognitive challenge, perhaps indicating that she needed more support on how to provide scaffolding to students without trivializing their thinking.

In L2, Kate made efforts to maintain the cognitive demand for the students who were struggling to solve the task, by posing an open-ended question when approaching them. Consider the following two episodes:

Kate: [She notices that a student partitioned one of the polygons into a rectangle and says:] Ok, I see some triangles in which the first polygon is partitioned. What about this polygon? Did you form triangles? You must turn it into a triangle. *What can you do?*
Student: Should the diagonal start from here? [indicating a vertex from which the other diagonals start]
Kate: Yes, you can start from there. Where should the diagonal end? *One of your vertices does not have a diagonal.*
Student: Oh! OK. [the student corrects his error]
Kate: Bravo. (Episode 1, L2, min 9:53, emphasis added)

Later, she interacted with another student:

Student: Miss, what shape is the n-gon?
Teacher: Hmm... let's discuss it. *What do you think?*
Student: It means that it has n number of angles.
Kate: Yes, this is true. The n might be any number. *Observe the completed table to identify a pattern. How did you work to complete the rows for the rest of the polygons*? [she leaves the student and heads to another student group] (Episode 2, L2, min 8:20, emphasis added)

In both episodes, Kate scaffolded students to overcome their difficulties or errors. She initiated the discussion with an open-ended question, asking them to discover ways to fix their errors (Episode 1, line 4) or explain their thinking (Episode 2, line 2). Although she maintained the level of cognitive activation in the second episode (Episode 2, lines 4-6), in the first episode, she pointed out the student's error *herself* (Episode 1, lines 2-4) and gave him a hint that directed his thinking (lines 7-8). However, in both cases, Kate allotted minimal time for the students to engage in mathematical reasoning and provide explanations.

As in the previous lessons, in L3, students who were making some progress on the task or arrived at the correct solution without any support were praised by Kate, though the *type of support and her questioning techniques* were getting better. Notably, in L4, Kate found a way to address this challenge: she was observed circulating within all the student groups trying to attend to the needs of all students and interacting with them in more substantial ways. Specifically, Kate circulated among the groups, stopped by most student pairs, while being willing to listen to or read students' thinking. When students were facing difficulties, she asked them to describe or explain what they had done up to that point on the task; then, she posed a question/suggestion to work on (e.g., "Why do you think that this product is less than $2\frac{2}{3}$? Think about it and explain your thinking." L4, min 9:05). In other cases, she provided enablers to less advanced students after letting them struggle for a couple of minutes ("Focus only on this multiplication method [i.e., repeated addition]" L4, min 16:54). Kate also headed to more advanced students, who explained their thinking without being asked or showed their answers in their textbooks. After carefully listening to or reading students' ideas, Kate posed clarifying questions to help them be more precise or identify and correct any errors ("How many eighths are there in a whole unit?" L4, min12:30) or extending questions (e.g., "Provide suitable titles for each mixed-number multiplication method, considering the process followed to be solved" L4, min 15:32).

Struggling with certain ideas inherent in concurrently attending to both fronts. Like Pina, Kate faced different *observed* or *perceived* challenges when working at the nexus of cognitive activation and differentiation. These challenges were clustered into two subcategories, depending on the timepoint they occurred or were treated. The first type pertained to *challenges that were mitigated or addressed* during the PLD program, such as providing extended learning opportunities to more advanced students with the use of extenders, or gradually avoiding doing the thinking for students. The second type concerned a group of *unresolved challenges*, including maintaining the cognitive demands, despite bringing the class to the plenary at appropriate checkpoints; providing adequate support to students who still experience difficulties with working on the task even after the provision of enablers; facing time constraints; developing and enacting enablers appropriately; and handling the accumulated knowledge gaps of some higher-grade students. Each type of challenge is elaborated on below.

Challenges Addressed or Mitigated During the PLD

Avoiding doing the thinking for the students. Kate progressively reduced the instances of thinking on behalf of her students during teaching. This was particularly observed during autonomous work and whole-class interactions, the phases that took up most of the lesson duration (this can be seen from the gradual reduction of the number of check marks in the grey-coloured practices of Table 13). Specifically, in L1, most of her questions elicited simple pre-determined responses, including a suggestion as to the correct answer. The IRE question sequences posed by Kate provided little wait time and lessened the opportunities that students had to participate in mathematical reasoning. For instance, students were working on investigating the relationship between two supplementary angles by moving their common side to two different positions (see Figure 32); Kate approached a student, read his response ("The two angles are equal"), and interacted with him, ending up being too directive:

Kate: *So, what happens to the second angle because of the change in the measure of the first angle*? When one angle is…
Student: …obtuse...
Kate: …the other angle is…
Student: …acute.
Kate: Do we have acute and obtuse angles in this example here? [showing two equal supplementary angles]
Student: No.
Kate: So, what happens to the pair of angles?
Student: They change.
Kate: When the one angle inclines to form an acute angle…
Student: …the other becomes obtuse.
Kate: *So, when we make a change on one angle, a change is made on the other angle of the pair. Write it in your own words.* (L1, min 17:00, emphasis added)

In this episode, Kate expected the student to fill in the missing words in her incomplete statements instead of letting him think and arrive at the conclusion himself (lines 2-5, 11-12). Equally problematic, she expressed the conclusion herself, leaving no room for the student to think, despite telling him to write the conclusion in his own words (lines 13-15).

On the same pattern, during the brief whole-class interactions of the same lesson, Kate encouraged students to share their mathematical contributions—confined to superficial observations—but she synthesized important mathematical ideas with limited student contributions. For instance, when students were asked to share their observations about the relationship of the supplementary angles formed when the backrest of a sunbed moved in two different positions (see Figure 32), many students had not arrived at a conclusion. Kate initiated a discussion with an open-ended question ("What have you discovered about the two angles?"). Once realizing that the first shared student answer was incomplete, she described the context of the problem herself:

> It seems that you haven't understood the givens of the problem. Let's say, I move the backrest to a vertical position; two right angles are created. When I want to relax, I move the backrest toward the back. What happens to the two angles? (L1, min 18:07)

She finally synthesized the key ideas herself with only minimal student contributions, telling students exactly what to write in their textbooks ("We conclude that as the measure of the first angle increases the measure of the second angle decreases. Let's complete the task… Write: As one angle increases the other angle decreases." L1, min 19:00). As she admitted during her first post-lesson interview, she had not foreseen how difficult this task/concept would be ("I considered the tasks very easy and that students would complete them without any delay", POI1, line 3-4), since students had previously worked on more difficult concepts/tasks ("Come on, you have done much more difficult thinking than that in the past, haven't you?" L1, min 18:07).

As the PLD program progressed, the excessive lessening of the demands observed in her earlier lessons was reduced, since Kate started refraining from giving explanations herself while posing more open-ended questions and providing students with fewer guiding prompts. In doing so, she provided students with more opportunities to think further about their arguments, as detailed in the third subtheme of her challenges.

Providing extended opportunities for thinking and reasoning to more advanced students. This was Kate's biggest perceived challenge. Kate supported that the program helped her address a concern she had had for years concerning how she could productively handle students who finish working on a task early, ensuring sufficient access to extended learning opportunities rather than becoming disinterested due to repetitive content ("I always had this concern –even as a novice teacher— why do we let these 3-4 students feel bored?", EPI, lines 78-81; "What can I do with those students who finish early and have nothing to do? [When this happens,] I lose valuable teaching time, so I must have something ready for them", VCS2, lines 16-18).

She was preoccupied with this concern until the introduction of the idea of extenders. After she experimented with this tool, as a classroom routine for early finishers, she was positively surprised ("I was like 'Wow! I had never thought of it!", VCS9, line 450), as doing so was aligned with her need to attend to this student cohort. Hence, this challenge was somewhat alleviated toward the end of the program with the design and use of appropriate extenders, as observed in her lessons. This was not the case with the following challenges which remained till the end of the PLD program.

Unresolved Challenges

Reducing the demands despite bringing the class to the plenary at appropriate checkpoints. Although Kate would bring the class to the plenary at appropriate checkpoints, she still faced struggles with maintaining the cognitive demands. This difficulty remained until the third lesson.[43] Specifically, in L1, the phases of autonomous work and whole-class interactions alternated *too* quickly, resulting in not leaving enough time for students to make observations and draw conclusions about supplementary and complementary angles. In L2, Kate appropriately interrupted autonomous work and initiated a whole-class discussion at a critical point of the lesson (*after* students had worked on partitioning the polygons into non-overlapping triangles and *before* formulating the relationship between the number of sides of a polygon with the number of triangles in which the polygon is partitioned into, see Figure 33). However, when the class was working on calculating the first polygon (the quadrilateral), Kate herself introduced the idea of the sum of the angles of a triangle ("What do you remember about the sum of the interior angles of a triangle? We have two triangles in the quadrilateral, which equals 180^o+180^o", L2, min 17:06). This idea was surfaced very early without letting students identify, share, and explain how the angles of the two triangles correspond to the angles of the quadrilateral.

The discussion continued with completing the rest of the table to the decagon and extracting the general rule for the n-gon (see Figure 33). A student observed a pattern vertically, but the teacher urged students to read the table horizontally. However, another student jumped in, saying "It is minus 2" (L2, min 17:55) and Kate wrote the rule algebraically on the smartboard, asking the class to formulate it verbally. Kate recognized that this had reduced the demands: "We went beyond the octagon and reached up to the calculation of the sum of the interior angles of the decagon. I should have stopped the discussion earlier because a student jumped in and presented the rule!" (VCS7, lines 429-431). Similarly, in L3, Kate initiated whole-class discussions at points that lowered the demands (e.g., L3, min 27:40).

[43] The main task used in L4 consisted of a central question rather than a series of sub-questions. Hence, it was not needed to bring the class to plenary at certain checkpoints.

Providing adequate support to struggling students. At several junctures during the program, Kate was confronted with a dilemma: how to ensure that the lesson goals would be achieved for all students while at the same time sustaining the cognitive challenge for different groups of students:

> What can I do with the students who cannot think of at least a way to start working on the task? Giving them some help does not imply that I am stealing their thinking. […] I try to help them in several ways, by providing some hints or posing questions–but this kind of support leads nowhere, …hence, I *have to* tell them the answer; there is nothing else left to do! (VCS2, lines 10-13, 38-40).

Looking somewhat troubled and frustrated, Kate admitted that she did not know how to handle struggling students. Although the program had offered her some concrete ideas on how to deal with this issue (e.g., enablers), this challenge remained till the end of the PLD program ("I shared my concern *again*: what can I do when students get stuck? Of course, some ideas were discussed…but if students get stuck and you run out of time? EPI, lines 101-103, emphasis added).

In fact, the idea of enablers did not appear to be very convincing to her. She felt that enablers do not always work with less advanced students in conjunction with *time pressure*, which further complicated the picture. Particularly, time constraints when trying to balance between keeping the mathematical challenge high and handling students' difficulties were amongst the challenges that popped up as Kate experimented with the program ideas. Kate acknowledged that in her interactions with struggling students, she had increasingly become over-guiding—something she attributed to time pressure: "Time pressure is an issue. This task took much longer than expected. […] What should I do when students do not make any progress or do not seem to get the mathematical ideas? What else can I do when they are stuck?" (POI3, lines 119-124).

Developing and enacting enablers appropriately. Kate used enablers in her lessons, facing different difficulties each time (see Figures 33, 34, and 35). Despite endorsing this idea, she still had issues and difficulties that were not captured by the quantitative analysis. In her last three lessons (see Table 11 in Chapter 3), Kate attempted to adjust the task in response to her students' particular needs, by developing and using enablers, each time varying different parameters such as (a) the number of enablers planned; (b) the strategy used for developing them; and (c) the degree of interaction with students who received them.[44]

For instance, the enabler in L2 presented a starting point by marking a vertex from where students could draw the diagonals to form the triangles (see Figure 33). The development of this enabler was supported by her detailed analysis of the task and students' artefacts/reactions from previous years: "My ex-students faced a lot of difficulties with bringing diagonals so, my decision was based on experiences I had with them" (VCS7, lines 291-292)—thus, bringing the perspective of students for the issues raised. The enabler could help students who struggled with partitioning the polygons access the task and make at least some progress on it without trivializing the challenge.

In the next lessons (L3 and L4), Kate developed *multiple enablers*, one for each challenging task. Her enablers were of two types (a) *using representations* for visualizing the mathematical notions (for L3: a picture showing the quantities for making the abstract concrete, and a diagram illustrating the multiplication of a fraction by an integer see Figure 34; for L4, a representation of the concept of mixed numbers multiplication, see Figure 35) and (b) *reducing the work students were expected to do* (for L4: asking them to focus on only one of the three methods, see Figure 35).

However, some of her enablers that involved representations had some limitations: they were either doing the thinking for the students or were mathematically incorrect. For example, the diagram of the second enabler used in L3 showing ¼ of 8 was too leading for the students (see Figure 34) and the diagram of the first enabler of L4 was not mathematically appropriate for representing the concept of multiplication (see Figure 35). Kate also admitted difficulties in developing appropriate enablers for her students (e.g., she expressed her concerns about the suitability of the diagram used as an enabler for representing the concept of fraction multiplication in L4, PRI4, lines 44-51). Perhaps she needed more support in developing enablers that would be supportive, mathematically appropriate, or would *not* do the thinking for the students.

[44] All enablers were planned and used according to the anticipated student difficulties.

Figure 34

A Representation of (A) the Main Challenging Tasks of Kate's Third Lesson, Along With (B) Their Enablers, and (C) Extenders

(A) Main task 1

(a) In the adjacent image, each bottle contains $\frac{2}{5}$ L of juice. Calculate the total quantity of juice included:

(i) in 3 bottles

(ii) in 5 bottles

(iii) in 7 bottles

(b) Describe a method to calculate the product of a whole number and a fractional number.

(Mathematics Curriculum, 2018, Grade 6, Unit 8, p. 11, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

(B) Main task 2

Anna had 8 L of juice and used $\frac{1}{4}$ of it. How much juice did she use?

(a) Which mathematical statement best describes this problem? Explain your thinking.

A. 8 – ¼ B. 8 + ¼ C. 8 ÷ ¼ D. ¼ · 8 E. ¼ ÷ 8

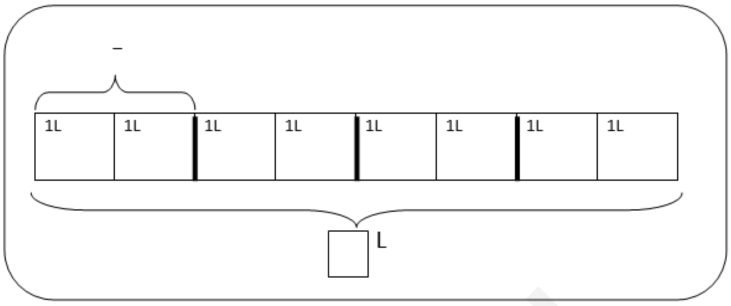
(b) Use words, drawings, or mathematical symbols to solve the problem.

(Mathematics Curriculum, 2016, Grade 6, Unit 8, p. 12, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sports, and Youth of Cyprus)

(C) Enabler 1 (For Main Task 1)

Enabler 2 (For Main Task 2)

(b) Use the following representation to solve the problem.



(D) Extenders

(a) Complete the question of the problem and select the appropriate mathematical sentence that fits the problem.

- Maria used $\frac{5}{6}$ of a dozen eggs to make a cheese pie. ______________________________
____
____

$12 - \frac{5}{6} = n$ $\quad\quad$ $\frac{5}{6} + 12 = n$ $\quad\quad$ $\frac{5}{6} \div 12 = n$ $\quad\quad$ $12 \div \frac{5}{6} = n$ $\quad\quad$ $12 \cdot \frac{5}{6} = n$

(b) Write a mathematical problem that corresponds to the mathematical sentence $6 \cdot \frac{2}{3}$ and then solve it.

____
____
____

(c) Write a mathematical problem that corresponds to the mathematical sentence $\frac{2}{3} \cdot 6$ and then solve it.

____
____
____

Figure 35

A Representation of (A) The Main Challenging Task of Kate's Fourth Lesson, Along With (B) Its Enabler and (C) Extenders

(A) Main Task

Explain each child's thinking.

The product of the mathematical sentence $\frac{7}{8} \cdot 2\frac{2}{3}$ is less than $2\frac{2}{3}$.

The product of the mathematical sentence $\frac{8}{7} \cdot 2\frac{2}{3}$ is more than $2\frac{2}{3}$.

Aris

Phaedra

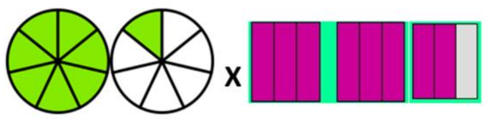
(Mathematics Curriculum, 2016, Grade 6, Unit 8, p. 33, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

(B) Enabler

Aris

$\frac{7}{8}$ x $2\frac{2}{3}$

Phaedra



$\frac{8}{7}$ x $2\frac{2}{3}$

(C) Extenders

1. Complete the inequalities with the symbols <, > or =.

$6 \cdot \frac{3}{10}$ …… $7 \cdot \frac{3}{10}$ $\frac{11}{12} \cdot 16$ …… $\frac{11}{12} \cdot 2^4$

$8 \cdot \frac{2}{3}$ …… $2 \cdot 3 \cdot \frac{2}{3}$ $4\frac{4}{9} \cdot 4\frac{1}{2}$ …… $4\frac{2}{9} \cdot 4\frac{3}{6}$

2. Write a problem which is suitable for the mathematical sentence $3\frac{2}{3} \cdot 4\frac{5}{6}$ and then solve it.

Beyond her difficulties in *designing* appropriate enablers, Kate faced some difficulties related to their *enactment* and specifically to the inappropriate monitoring of students' work to assign enablers. Specifically, the selection of students who would take an enabler was deliberate (i.e., less advanced students), according to her knowledge of her students and their difficulties. However, students received the enablers too early, without the teacher having ensured that they had understood the task instructions and clarified the difficult points or verified that they were still stuck; therefore, the extent to which the enablers helped these students was not clear. In all three lessons, she provided enablers sparingly to around six to eight students *right after* facing initial struggles, without first asking students to share their understanding or checking their progress on the main task (e.g., she immediately headed to students and without seeing their work or asking them to explain their thinking, she gave them an enabler and said "use this –this might help you because I included a starting point to divide the polygons", L2, min 5:23).

In L3 and L4, Kate was more flexible to provide enablers to other students who struggled with the task after letting them work for some time on the task or left it on them to determine if they would use it (L3, min 20:50, 21:49). Moreover, she usually checked whether the enabler had supported students' work in proforma ways (e.g., "Was that of any help?" L2, min 14:15) slightly increasing her interaction with the students in L3 and L4 (e.g., "What is the relationship of this part with the whole? You wrote that it is ½. Is it half of the whole? Think about it" L3, min 25:43; "Did you understand this method?", L4, min 20:32).

In both L3 and L4, she immediately assigned the enablers to two students (the same in both lessons) asking them to ensure that they worked on them—during the POIs, she clarified that she had expected those students to face significant difficulties on the task. She also assigned enablers to four additional students (again, the same in both lessons) once realizing initial student struggles, without, however, interacting with them to figure out their difficulties; unlike with the first couple of students she left it open to the students to use the enablers ("I leave this here with you—use it if you think that it might help you", L3, min 21:55). Interestingly when a student asked for an enabler (in L3, min 20:57), Kate quickly monitored the student's work, replying that "you don't need to be given an enabler". In other cases, when students were having some difficulties, she would ask some quick questions to help them rethink their work (e.g., "I noticed that you subtracted; what should the task be asking you to do, to subtract?", L3, min 21:00), without giving them any enablers. She was trying to be flexible while also deciding who truly needed enablers.

Reflecting on her experimentation with the tool of enablers Kate was kind of uncertain about the extent to which her implementation of enablers supported her work with students: "In all honesty, I am not sure about the extent to which the enabler helped [my students]" (VCS7, lines 395-396). One possible reason could have been that this was her first experimentation with this tool and that her students needed to be familiarized with using it. Furthermore, designing inappropriate enablers may also be associated with any deficiencies in her mathematics knowledge for teaching.

Handling the accumulated knowledge gaps of higher-grade students. Kate's difficulty in enhancing cognitive activation for upper elementary school students who have cumulative knowledge gaps and cannot follow the curriculum was a concern that was not alleviated until the end of the PLD program, despite her efforts to support these students in different ways (POI4, lines 159-163): "Whatever I do, as a teacher, fails; [When students reach Grade 6 with these knowledge gaps], I think we [teachers] have no responsibility. The responsibility is on the system or somewhere else" (EPI, 176-190). Although initially feeling desperate about these students' lack of required prior knowledge (VCS2, lines 126-129), Kate was relieved when the PLD group recognized the extra difficulties imposed when teaching students in upper grades in one of the sessions ("It is practically impossible to fill these knowledge gaps in the sixth grade", POI4, lines 165-167). Kate's concerns around attending to the needs of less advanced students epitomize the multiple dilemmas a teacher faces while trying to support students without doing the thinking for them; the teacher needs to weigh in different considerations, including students' needs, their progress on the task, and the perceived or actual time pressure.

6.5 Learning Level: Evolution of Kate's Conceptualizations Around the Two Axes and Their Interplay

This section explores the transformation of Kate's conceptual understanding regarding cognitive activation and differentiation, alongside their interplay. Initially grappling with limited knowledge on cognitive activation, heavily favoring differentiation, and unaware about their interplay, Kate's path reflects significant growth. Over time, she shifted her focus, deepening her grasp of each axis and their synergy. Her final conceptualizations revealed a firm belief in the importance of cognitive activation, and differentiation as a fundamental teacher duty, and a nuanced understanding of their interplay, marking a profound evolution in her educational philosophy.

6.5.1 Initial conceptualizations

Inadequate Knowledge Around Cognitive Activation. At the commencement of the PLD program, Kate's knowledge of cognitive activation was underdeveloped as she herself acknowledged (EPI, lines 28-30). Kate was not very talkative during VCS1 which focused on cognitive activation issues. Yet, she entered the discussion twice to superficially agree that there were instances in which the teacher of the video-clip the PLD group had watched became over-guiding (VCS1, lines 7-9) and then, to suggest that asking for explanations would help the teacher refrain from doing the thinking for the students (VCS1, lines 20-23). However, her idea of "explanations" largely expected students to simply describe the steps of a procedure rather than justify their thinking.

Another critical incident that emerged during VCS1 was when Kate pointed out that in that session the focus was mainly on cognitive activation and asked whether, during the upcoming sessions, the group would also work on differentiation (VCS1, lines 26-27). This remark as well as the nature of her participation probably illustrated that at least at the beginning of the program she was more interested in differentiation and would like to get more support in that direction or that she was not very comfortable with this new idea introduced to her. Some traits of her character, such as her introversion, could also justify to some extent the nature of her participation in the video-club sessions at least in the initial stages (she acknowledged herself, "I needed some adjustment time" EPI, line 618). When asked by the interviewer to explain the extent to which she had felt uncomfortable when sharing her ideas and video-clips from her teaching with the rest of the group, she provided a short account of her personal traits: "Initially, yes. Because I am a bit introvert as a person. So, I do not easily externalize my thoughts and feelings. I am more reserved. So, at first, I felt uncomfortable" (EPI, lines 654-657). Over time, her contributions to the video-club discussions began to escalate as will be explained in the next sections.

Her inadequate initial conceptualization of cognitive activation was also manifested in Kate's difficulties in identifying the cognitively demanding part of the task of her first lesson on complementary and supplementary angles. Specifically, when asked which part of the lesson would pose the greatest mathematical challenge to students, she did not answer directly and was somehow hesitant to respond—compared to her response time to the other questions— explaining that she could not know which lesson part would be more challenging for her students before teaching. When pressed further, she replied with difficulty that constructing complementary and supplementary angles would be hard for students, because of the relationship between angle pairs (PRI1, lines 38-39, 47).

Also, Kate was unaware of ways she could activate her students during the lesson, as suggested by the frequent pauses she had made when the discussion was coming to issues of cognitive activation. She repeated twice her uncertainty about how she could promote student cognitive activation using the lesson tasks ("I'll see how cognitive activation will work. I am not sure" PRI1, line 60). Asked to describe how she would cognitively activate her students in L1, she said that she would implement her usual classroom working routines (individual work—pair-work—whole-class discussion) because students would explain the concepts and the solved examples to each other (PRI1, lines 54-59).

Reflecting on how her first lesson unfolded, she disappointingly saw students facing difficulties with tasks she considered "easy, obvious, clear, not related to pre-requisite knowledge" (POI1, lines 2-9). To handle this unexpected event, she departed from her original lesson planning by devoting more time passing by all students to provide individual support and decided to reteach the content in the coming lesson (POI1, lines 10, 60). Kate had previously revisited the idea of *reteaching* the concept to help students consolidate new mathematical ideas: "The more they hear it, the better" (VCS1, lines 42-44). In brief, her conceptualizations of cognitive activation were at a rudimentary stage. Contrastingly, Kate entered the program very confident about incorporating differentiation into her teaching, as will be explained next.

A Superfan of Differentiation. Kate's initial systematic references to differentiation and other related terms revealed her passion and relatively deep knowledge of and engagement with differentiation. As she explained, her interest in learning more about differentiation had been aroused by a certain incident with a school inspector:

> [The school inspector] observed my lesson and told me: "I haven't seen differentiation in your lesson". I remember it characteristically! I asked her: "What do you mean by differentiation?" Her answer was "to produce different teaching material for the less advanced students", which is far from what we call

> differentiation! […] When I got the chance, I decided to receive PLD on differentiation and see what the school inspectors were asking for… It was something completely different! What convinced me about differentiation is that we help both charismatic students and less advanced students. They all have an equal right to learn and do what they can! (EPI, lines 227-242)

Interestingly, driven by the practical and ethical concerns associated with learning and teaching, Kate wanted to learn more about differentiation. Although in the first place, she wanted to improve her teaching for teacher evaluation purposes, she did not get disappointed when the PLD was not what her school inspectors expected. As she explained, she had bought into the concept of differentiation "from the first sessions [of that PLD program]" (EPI, line 224), acknowledging that it can support the two "extreme" student groups in terms of readiness and ability level.

During VCS2, Kate was still reserved, but obviously, more willing to speak when the focus of the discussion was shifted to ideas and practices of differentiation. Although the first half of VCS2 focused on issues of cognitive activation, Kate persistently tried to get the discussion to differentiation with various questions and comments (e.g., "There are two or three sixth-grade students in my class who do not know how to do mathematical operations. So, I can't help them –what should I expect from them?" VCS2, lines 44-45). This gave the facilitators the chance to shift the topic of the discussion to differentiation, by asking teacher participants to share their initial perceptions around it.

Kate confidently took the floor first, explaining three basic ideas about how she conceptualized differentiation: as a way with which all students can (a) be cognitively activated and feel that they succeed (VCS2, lines 58-59); (b) achieve a core learning goal to different shades and degrees (lines 64-65) and (c) get equal opportunities so that they acquire the new knowledge to some extent (line 67). Elaborating upon her last statement, she provided the following account: "Providing equal opportunities so that students certainly conquer the core knowledge. Also, giving opportunities to exceptionally able students to move forward with transformative knowledge or to less advanced students to consolidate the knowledge" (VCS2, lines 69-73).

Several key points can be identified in Kate's initial conceptualization of differentiation. First, she pointed out that differentiation is about helping all students learn something new to different extents, depending on how far they went from that at the beginning of the lesson. Second, Kate argued that all students need to feel challenged so that they invest effort and their right to learning is met—somehow prioritizing differentiation over cognitive activation. Finally, turning to how differentiation can be materialized, she kept insisting on the two extreme groups, bringing knowledge from her previous PLD experience (she even introduced the term "transformative knowledge" without the facilitators mentioning it).

Additionally, Kate was versed in different strategies (i.e., using asynchronous work; and setting clear core, pre-requisite, and transformative lesson goals) and techniques (i.e., rank-ordering lesson tasks according to their difficulty; developing entry cards for activating pre-requisite knowledge and exit cards for student assessment; and providing learning aids) for differentiating teaching, which she planned on using during L1 (PRI1, lines 11-20). Reflecting on how differentiation worked in L1, Kate seemed very satisfied because "even the less advanced students managed to construct or *at least* recognize the pairs of angles" and some of them worked on "transformative tasks" (POI1, line 50). It is worth noting that L1 was implemented at Timepoint 2, but the teacher was already very familiar with various differentiation terms.

Being Clueless About the Interplay of Both Axes. Initially, Kate did not have in mind the concept of cognitive activation, let alone the interaction between the two concepts (i.e., cognitive activation and differentiation). Despite her inadequate knowledge of or discomfort with cognitive activation, throughout her teaching career, Kate was concerned about identifying ways to appropriately challenge more advanced students, who usually constitute "the minority of the classroom" (EPI, line 80), and support them in moving a step farther:

> All these years, I was worried that we were paying more attention to less advanced students to help them acquire knowledge while neglecting the charismatic ones. It is very pivotal that the latter do not waste their time and get bored because the rest of the class keeps repeating the same things until everyone solidifies the knowledge. […] Cognitively activating those students to move on to something more challenging has always been very important to me. (EPI, lines 62-76)

In this excerpt, the teacher emphasized her concern and interest in identifying and discussing ways that would extend the challenge for her "charismatic" students who usually happen to be the early finishers and are engaged in unnecessary repetitive work.

Trying to attend to the needs of these students, in L1, she proactively planned on using "more complicated and harder tasks for more advanced students, focusing on transformative knowledge" (PRI1, lines 24-25). However, she seemed to have had an alternative idea of what *transformative knowledge* was, defining it as something that *all* students can work on after presenting their solutions in plenary (VCS1, line 41). Over time, a shift in her perceptions around working on both axes in tandem was noticed, which is discussed next.

6.5.2 Evolution of her Conceptualizations

Analyzing and Reflecting on the Cognitive Level of the Task and the Lesson. Although still puzzled over the cognitive level of the tasks, Kate gradually started horizontally rank-ordering, and reorganizing the different *subtasks* of a given task, according to their mathematical challenge (VCS6, lines 340-341), and sharing her concerns with the rest of the group. For instance, in VCS5, the PLD group worked on classifying a set of tasks as high or low challenging. Teacher participants were not sure about the level of a Grade-3 task on subtracting three-digit numbers with regrouping. While some teacher participants emphasized that problem-solving is a challenging area, another teacher group supported the idea that solving a problem does not render the task as high-level since students can simply execute operations without making substantive connections to the content. Still unsatisfied by the ideas shared, Kate continued pressing the group to decide on the task classification ("So, at what level would we classify this task, as such?", VCS5, line 364). At this point, Kate focused on how the task itself could solved, ignoring how other classroom-based factors (e.g., the scaffolding behavior exhibited by the teacher) may contribute to the cognitive level at which the task is finally implemented; over time, Kate has been supported to identify and modify such factors along the way by the PLD program, as highlighted below.

Specifically, Kate was increasingly becoming familiar with cognitive activation issues and was consciously aiming to challenge her students ("I usually have in mind that I should maintain the demands, but I can't always achieve it", POI3, lines 87-89). Gradually, Kate identified and described the cognitively demanding part of her lessons with greater ease. For example, the main task of her second lesson expected students to bring all the diagonals from one vertex to partition the given polygons into non-overlapping triangles and arrive at a general rule for calculating the sum of the angles of any polygon (see Figure 33). In this lesson, she confidently acknowledged that the solution process and the use of prior knowledge were the cognitively demanding part of the task ("Figuring out a pattern that connects the number of the polygon sides with the sum of its interior angles entails working inductively to discover the rule and considering the prior knowledge on the sum of the triangle angles", PRI2, lines 39-41). She was more satisfied observing that her students were quite challenged because they worked on their own, tried, and ended up with multiple ways of dividing the polygons into triangles (POI2, lines 13-15).

Moreover, she started voicing her teaching dilemmas and critical moments which had presumably led to a decrease of mathematical challenge in her lesson. For instance, Kate consciously opted to show how to draw a diagonal from one polygon vertex to *all* students during the launching of the abovementioned task used in her second lesson (see Figure 33). Although she knew that the diagonal was a new concept, she considered it important to introduce it so that the students would not deviate from the intended goal ("I supported all students by clarifying what a diagonal is, so to avoid misunderstandings... I may have stolen their thinking though", POI2, lines 25-27). This decision shows some flexibility on her part: based on students' work and unexpected challenges, she considered the clarification of keywords to be very important for students to understand the task instructions and thus, avoid ending up engaging in unsystematic exploration.

Over time, Kate began to discern at which point in the lesson a possible intervention by the teacher could reduce the cognitive requirements of the task. For example, there was an episode in which she was reflecting with a colleague, Georgia, on a lesson they had both implemented (on the sum of the interior angles of a polygon, see Figure 33). Kate seemed to have developed an understanding of when teacher actions can have an impact on the challenge:

Georgia: I formatively evaluated my students by asking them to implement the rule in other polygons. I am wondering whether doing so helped them consolidate the new knowledge or is considered stealing their thinking…
Kate: Did you formatively evaluate the students *after they themselves* drew a generalization?
Georgia: Yes.
Kate: So, *this was not* stealing! (VCS7, lines 364-369, emphasis added)

Reflecting on her colleagues' experience, Kate gradually improved her noticing skills and knowledge with respect to certain teaching aspects discussed in the sessions (lines 4-5, 7).

Similar discussions, during VCS7, reinforced her progress in understanding that challenging tasks are not necessarily experienced as such, because of how a task gets implemented (facilitated by the teacher). For instance, in the same lesson (on calculating the sum of polygon angles), both Kate and Georgia dealt with a similar situation and felt that they allowed the discussion to unfold for longer than it should have, resulting in students jumping in and giving away the rule. In Kate's case, she was interacting with the students in plenary, in order to reach a conclusion, when some students interfered and gave away the rule. Kate built on the students' interference and then asked another student to express the rule in symbols. She was wondering whether she had dealt with this situation efficiently, illustrating that she had started questioning her moves and handling of unexpected teaching events during teaching. Kate was torn on this incident: on the one hand, she noticed that sometimes "this is inevitable to happen" (VCS7, line 447); on the other hand, she was judgmental towards herself: "I feel that I made a mistake and I want to point it out [...] I should have stopped the discussion earlier because a student jumped in and gave away the rule!" (VCS7, lines 429-431). It seems that Kate began to delve into issues of cognitive activation, leaving differentiation issues aside for a while, as discussed below.

Gradually Eliminating the Focus on Differentiation While Raising Open Issues. After Timepoint 2, Kate eagerly presented the differentiation practices she was using. For example, she confidently referred to asynchronous work, which "already works very well in her lessons" (VCS2, lines 120-124) and emphasized the use of entry and exit cards (VCS2, lines 29-30, 32). Moreover, she enthusiastically shared her conceptualizations of the *tic-tac-toe technique*, explaining that (a) students can work individually and asynchronously on it, by definitely solving the task/exercise in the central box which focuses on the "core knowledge"; (b) it is used "for refreshing the pre-requisite knowledge" or even "checking if the students have established the new knowledge"; (c) it includes tasks of "increased difficulty" aiming to "differentiate the challenge" for students; and (d) its structure promotes "both horizontal and vertical differentiation" because each student can solve the selected line or as many lines as possible at their own pace (VCS3, lines 67-96). Her final comment spoke for itself: this technique is "clearly, in favor of differentiation" (VCS3, line 96). The terms she used in this excerpt show her deep knowledge of the differentiation techniques she was applying.

However, despite her confidence in differentiation, she raised some emerging concerns. For instance, she provided an imaginary classroom scenario in which she gives an entry card and realizes that the pre-requisite knowledge of most of the students is not in place and wondered: "What would I do? Would I reteach the same content?" (VCS2, lines 126-129). This was the first time Kate's idea of *reteaching* the same content to ensure that all students would get it (which she had advocated earlier in VCS1) was challenged by herself. Kate gradually eliminated her references to the concept of differentiation probably trying to manage the complexity of the interplay, as explained next.

Moving Into Depth Regarding the Interplay. Despite the emphasis of VCS2 being on differentiation, Kate repeatedly took the opportunity to bring back what puzzled her the most: how to handle different student groups during autonomous work. For example, in collaboration with Georgia, Kate brought up specific suggestions, such as making good use of the time students spend working autonomously, asking more advanced students to come up with alternative solutions; providing supporting tools to less advanced students; and then asking the former to discuss and compare their multiple solutions (if any), while the rest of the class works on the core task (VCS2, lines 111-112). This was in line with the idea of enablers and extenders, which was officially introduced in a forthcoming meeting (second half of VCS5). It is worth noticing that Kate was describing examples of enablers (focusing students' attention on a less sophisticated method) and extenders (asking for generalizations) that she was using outside the context of the PLD program (VCS5, lines 157-160).

In VCS3, for the first time, teachers were asked to consider whether differentiation and work on mathematically challenging tasks go together or if one hinders the implementation of the other. Kate briefly supported that "student cognitive activity is enhanced *by differentiation* because the students know that they have to explain or present their work on the task…" (VCS3, lines 153-155). For her, when differentiation was implemented, cognitive activation and student accountability were indirectly promoted; it seems that Kate was still prioritizing differentiation—maybe because of her background and familiarity with differentiation—in the pursuit of offering sustained and on-level mathematical challenge to students.

Later in this session, she was invited to refer to teaching practices she had included in L1 that helped her in promoting both axes. She referred to asynchronous work; devoting plenty of individual work time; planning what the early finishers would do next; and asking students to explain or check the correctness of their work/solutions in pairs (VCS3, lines 117-126). Interestingly, in the previous sessions and interviews, she considered these ideas as practices promoting each axis separately but at this point, she brought them together.

When the idea of enablers and extenders was officially introduced, Kate raised some important concerns, illustrating that she focused on the essence of the new idea. Specifically, she explicitly attended to (a) the number of enablers ("Should we have single or multiple enablers per task?", VCS5, lines 538-541)—an aspect that could be differentiated, aiming to respond to the learning needs of different student groups—and (b) the relation between extenders and the core learning objectives (VCS5, lines 576-585). In the next sessions, Kate took the opportunity to share an interesting question: "What is the difference between transformative knowledge and extenders?" (VCS6, lines 272-273). She also suggested the use of enabler(s) by specific students as a key criterion for flexibly grouping students (VCS7, lines 266-278).

Her questions and ideas were probably influenced by her concern for attending to the needs of more/less advanced students, which was also consistent with her tendency to think deeply about the content of each session and draw links between the concepts. Based on this, we can assume that the implementation of the ideas received in earlier sessions informed the discussion and Kate moved into more depth; it, therefore, provided real opportunities for her educational development. We assume that she became more skilful in figuring out how to introduce or combine the ideas discussed in real-classroom situations.

Also, Kate was very excited about getting involved in generating a list of steps for lesson planning, in light of all the things that have been discussed in the sessions (this activity was implemented in VCS6, see Table 7)—a request that she had directly or indirectly made in previous sessions. She was very eager to prepare a draft lesson following these steps. As expressed in her reflection cards, Kate was repeatedly asking for it: "Designing specific lessons" (TRC1); "Designing lesson plans" (TRC2); "Designing specific lessons in the context of differentiation" (TRC3); "Considering cognitive activation and differentiation during lesson planning" (TRC4 & TRC5). The way she had gradually shifted her focus from planning *specific* content to lessons that attend to *differentiation* and finally, to lessons that promote *both axes*, indicates a change in her conceptualizations toward bringing together the two axes. This activity adjusted to her perceived needs and provided the opportunity to think deeply, along with her colleagues, about how to approach the lesson content to maintain the challenge while differentiating teaching (EPI, lines 448-460).

6.5.3 Final Conceptualizations

Endorsing the Idea of Cognitive Activation. At the end of the PLD program, Kate became a proponent of cognitive activation. She supported the idea that cognitive activation can be achieved when the teacher prepares and uses the appropriate tools that can potentially cognitively activate students and works diligently on the lesson goals ("The lesson must have a starting point and the teacher should keep in mind the endpoint", EPI, line 149). Moreover, she listed several practices she had learned because of her participation in the EDUCATE PLD program. She admitted that she learned how to analyze a task during lesson planning; avoid doing the thinking for the students; and formulate open-ended questions to enhance cognitive activation (EPI, lines 414, 426, 452-453).

Particularly, through her questioning, she felt that she started allowing greater autonomy to students, compared to how guiding she had been in previous years, especially during the phase of student autonomous work (EPI, lines 344-356). For example, she saw a change in herself: "I used to tell students: 'The task asks us this and that'. Now, I use different wording: 'What does the task ask us to do? How we can you get started working on it?' to cognitively activate them" (EPI, lines 419-421).

In addition, reflecting on the complexity of orchestrating whole-class discussions during the sessions, she emphasized that focusing on this phase during the sessions was of particular importance for her. For instance, she believed that she was now filtering what she was sharing or asking the students to do; was weighing students' thinking to select and sequence students' solutions in the synthesizing part of the lesson; and implementing specific actions (such as, revoicing, repeating, and adding on) to help students extend their solutions (EPI, lines 333-337).

Although there was a shift in how Kate gradually acquainted herself with the task analysis, when prompted to identify and explain the challenging part of her final lesson, she focused mainly on describing the steps and the final answer that students were expected to achieve (PRI4, lines 19-22, 64-66). As in her first post-lesson reflection, not anticipating the exact kind of thinking or parts of the task that would be challenging for students and also, taking for granted that students would easily activate their prior knowledge made her disappointingly realize that her students had more difficulties than expected in her last lesson: "I thought that the lesson would flow easily [...] because this was something we had discussed again and I expected that they would get it more easily" (POI4, lines 5, 102).

Finally, whereas at some points during the sessions, she accused herself after considering incidents in which she had stolen students' thinking, in her EPI, she became more realistic. She realized that there are moments in which part of the demands will decline but the teacher's role is to reduce such situations: "Avoiding stealing students' thinking is important. I cannot always achieve this—a student may jump in and share her solution and I might not prevent this. We saw this happening in the shared clips from our videotaped lessons—it happened to me, as well. The important thing is to reduce stealing students' thinking" (EPI, lines 426-436). This statement shows a sensible realization on her part of what can be expected in real classroom situations without implying that she had become more complacent with the idea of sometimes doing the thinking for the students as part of actual teaching practice.

In essence, Kate embraced the idea of trying to maintain the demand as much as possible; at the same time, she realized that sometimes a teacher may need to strike a balance between the level of cognitive activation and the support provided to students, by sacrificing part of the task demands (depending on the lesson goals and students' difficulties), in order to help the students make some progress on the task. Despite her meticulous lesson planning, as each lesson unfolded, unexpected incidents arose, which required in-the-moment decisions based on her interpretations of students' work and understanding (EPI, lines 92-94).

Differentiation Is "a Moral Responsibility". During her EPI, prompted to reflect on the extent to which differentiating teaching is challenging to be accomplished in a mixed-ability classroom, Kate insisted on her initial ideas, arguing that differentiation is a "moral responsibility" for the teacher and at the same time, "a right" for students (EPI, line 248). In contrast to how simple she thought that differentiation was at the beginning of the PLD program, in the end, she argued that it is not easy work, because it takes considerable planning and preparation on the part of the teacher (EPI, lines 105-106).

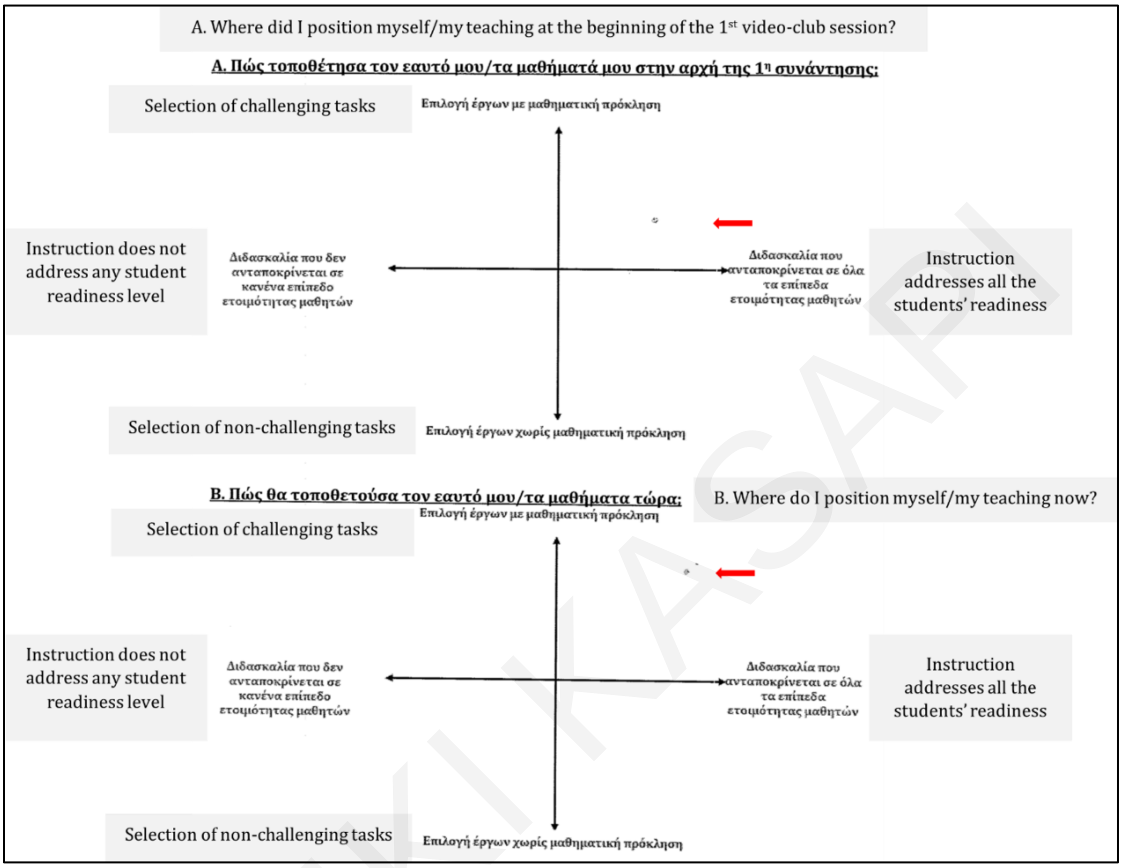
Despite remaining still excited about differentiation (EPI, line 201), Kate's arsenal of differentiation tools was not enriched after she participated in the PLD program because she had already entered the program knowing several differentiation strategies and techniques from her earlier PLD experience. Yet, she was given the opportunity to refresh those practices and further experiment with some of them, such as employing entry/exit cards; formatively assessing students; asking students to generate their own mathematical problems; or using asynchronous work (EPI, lines 199-212). Compared to the beginning of the PLD program, Kate reduced her references to differentiation, thus indicating that she possibly focused more on the interplay of the two axes.

Consolidating Certain Aspects of the Interplay. After all the experiences she had gained, Kate supported that the cognitive challenge can be maintained for *all* students but to different degrees, promoting positive changes in their academic achievement and attitude toward challenge (VCS9, lines 253-254; EPI, lines 252-261). As she admitted, the PLD program supported her in refreshing her prior knowledge which became more profound with respect to differentiation. Her knowledge base was also enriched with issues of cognitive activation: "I did have some theoretical foundations on differentiation, but I went deeper into teaching, dealing with the various difficulties that students face in being cognitively activated while being engaged with a difficult subject [i.e., mathematics] in different ways" (EPI, lines 53-59). For her, the greatest gain from the PLD program was that she learned *how* and *why* it is necessary "to activate all students and care for everyone" (EPI, line 62), especially *more advanced students*. Contrary to the first time she was asked to describe the relationship between the two axes (see VCS3), in her EPI, she did not prioritize one axis over the other but gave them equal weight. Kate believed that the two axes serve each other: they advocate the same desired result "that is, each student to get the appropriate support, according to their readiness level and needs" (EPI, line 284).

Conclusively, in the first (VCS1) and last (VCS9) sessions, teachers marked their teaching approach on a two-dimensional graph (see Figure 36). The vertical axis showed their use of challenging tasks, while the horizontal axis represented addressing varied student readiness. In comparison to her initial teaching stance, Kate estimated improvement along both axes, as evidenced by the diagonal upward-right shift of the spot within the same quadrant. A more careful observation of the shift reveals that there was a slightly greater improvement in the selection of cognitively demanding tasks compared to addressing different students' levels.

Figure 36

Four-Quadrant Diagram Depicting Kate's Positioning of Her Teaching (A) Before and (B) After Participation in the Program Concerning Cognitive Activation and Differentiation, With Its Translation in English



Requested to identify factors contributing to her transition, she offered the subsequent explanation:

> [I noticed] an enhancement in the learning process, by placing emphasis on aspects such as questioning; avoiding stealing the student's thinking; providing learning aids; using enablers and extenders; delivering individualized assistance when necessary; and orchestrating whole-group discussions. (TRC9, question C)

Kate's response indicated her prioritization of students' learning experience (process), rather than the learning outcomes (product). Interestingly, she provided greater specificity regarding practices she implemented during student autonomous work, while she referred to the whole-group discussion in more general terms. When prompted to provide further detail about this in her EPI, she clarified that she learned to discern which solutions to present and deliberate upon in a whole group setting; and how to articulate her utterances and questions in ways that stimulate critical thinking and respect students' cognitive processes (lines 338-342).

While the enumerated practices showcased her attempt to encompass diverse pedagogical aspects, the rigorous scrutinization of the depth of their implementation can provide a more comprehensive picture of Kate. Thus, having considered the changes in Kate's conceptualization, we now turn to comparing them with the changes in her teaching practice.

6.6 Learning and Behavior Levels: A Comparison Between Kate's Conceptualization and Practice

By comparing the changes in Kate's conceptualization of cognitive activation, differentiation, and their interplay with the changes in her practice in terms of these three axes, relatively strong links in the axis of differentiation and little coherence in the other two axes are identified. Specifically, relative stability and consistency were found in the conceptualization and practice of differentiation ideas during the PLD program. In the other two axes, her conceptualization was constantly moving forward while some conflicts between her previously established concepts arose (e.g., she questioned her earlier idea of reteaching the lesson content to help students assimilate it). However, in her practice, she experimented with multiple practices, albeit not systematically, thus rendering herself a case of a teacher who made some progress in some respects and at the same time, had several setbacks. In what follows, we elaborate more on the extent of the alignment of the two (i.e., conceptualization and practice).

Apparently, Kate was solid on differentiation (in terms of conceptualization) and her practice reflected that. Besides, she had the fewest back-and-forths with respect to this axis in both her conceptualization and practice. She was a characteristic example of a teacher who entered the program feeling confident and knowledgeable about ways with which she could incorporate differentiation into her classroom; this was manifested in her practice, as well (see Lesson 1, Table 13). Remarkably, although she did not remain preoccupied with the concept of differentiation until the end of the PLD program, she had the opportunity to bring back earlier ideas and systematically experiment with them across her four videotaped lessons. Reviewing, reflecting on, and experimenting with differentiation ideas and practices further solidified her learning. In her four lessons, she experimented with a similar number of practices mostly at a fairly good level of quality (Table 13). Therefore, regarding differentiation, her conceptualization and practice were largely aligned throughout the program.

On the other hand, Kate's conceptualization of cognitive activation seemed to be mostly ahead of her respective practice. At the commencement of the PLD program, her preliminary conceptualization of cognitive activation was in its infancy, as evidenced by her difficulties in analyzing and discussing the mathematical challenge of tasks. As time went by, Kate began to identify some challenging aspects of the lesson; recognize contextual factors that could influence the challenge; bring up difficulties and dilemmas she faced; and finally, embrace the idea of trying to maintain the demand as much as possible. However, the changes in Kate's conceptualization of cognitive activation were not consistent with the changes in her teaching practice. Although she experimented with various ideas promoting cognitive activation, this was not done in a systematic and consistent way. Oftentimes she would engage in some practice in one lesson and then abandon it or focus on another practice in the next lesson. Despite making some improvements on both fronts (conceptualization and practice), Kate seemed to be better at conceptualizing cognitive activation rather than promoting it in her lessons. Hence, the changes in Kate's conceptualization and practice in this axis do not reflect very close correspondence.

A similar low correspondence was also found in her conceptualization and practice regarding the interplay of both axes. Whilst Kate was placed in a learning context (i.e., the PLD program) where she was expected to equally weigh issues of cognitive activation and differentiation (which was quite innovative for her), she finally supported a reciprocal relation between cognitive activation and differentiation. She thereby made significant advancements in her conceptualization regarding the interplay. The novel concept of cognitive activation and its interplay with differentiation piqued her interest and curiosity, deciding to concurrently experiment with multiple new ideas (see Table 13), gradually leaving aside the familiar issues of differentiation. However, working on various ideas and fronts simultaneously in terms of the interplay seemed to become overly complicated for Kate, who generally ebbed and flowed in her practice, following a comparable trend as in her practice with cognitive activation ideas.

In all, the lack of close correspondence between her conceptualization and practice in the two axes (cognitive activation and the interplay) suggests that her learning was not linear. Moreover, it reminds us that teaching requires concurrently attending multiple levels and that each teacher as a learner (in this case, Kate) has her own starting point when entering a PLD program. For example, Kate was already a fan of differentiation and unaware of cognitive activation; one could not expect massive changes in her practice, given her underdeveloped conceptualization of cognitive activation issues upon joining the program. Perhaps effectively managing the workload and balancing multiple ideas was pretty intense for Kate, who might have needed *more time* to better “digest” the new concepts and more scaffolding on setting her priorities or focusing on improving the enactment of certain practices.

What should be acknowledged is that in both axes (cognitive activation and the interplay), she became *aware* of the improvements in her teaching (e.g., advancement in avoiding doing the thinking for the students; posing more open-ended questions, etc.) and also of the difficulties she faced (e.g., challenges in preventing students from telling despite bringing the class to plenary at appropriate checkpoints; developing appropriate enablers, etc.). Therefore, if she had been given the opportunity for more video-club sessions and videotaped lessons, we might have seen more improvement in these practices, since she might have become increasingly more capable of dealing more effectively with her unresolved difficulties.

In conclusion, Kate represents a case of a teacher who, while showing promise and improvement in certain areas, notably differentiation, struggled to cohesively integrate and sustain advancements in cognitive activation and its interplay with differentiation across her practice. In particular, she had a solid foundation and consistent growth in differentiation, where her conceptual understanding and practical application remained aligned with her practice and showed progressive refinement throughout her participation in the PLD program. For cognitive activation and the interplay between cognitive activation and differentiation, her journey was more tumultuous. While she exhibited an evolving conceptual understanding of cognitive activation, her practice did not consistently reflect these advancements. Similarly, her grasp of the interplay between the two axes, despite showing initial promise, did not find a sustained and systematic expression in her teaching. Overall, her evolution underscores the complex, non-linear nature of teaching and professional learning, highlighting the challenges of balancing multiple pedagogical goals simultaneously.

CHAPTER 7. THE CASE OF MICHELLE: (OVER)SHARING IS CARING

Michelle is the third case. She was a well-intentioned teacher who, in her zeal to support her students, tended to 'overshare' educational guidance, inadvertently taking over their thinking process. This over-involvement stemmed from a caring intention, mirroring the "sharing is caring" ethos, but it crossed into providing excessive help. This approach, while supportive, risked stifling students' ability to think independently and engage in productive struggle, essential components of effective learning. Still, Michelle's journey in the PLD program showcased her transition from unfamiliarity and avoidance of challenging tasks to embracing them as tools for cognitive activation. However, her understanding of differentiation remained underdeveloped, predominantly aiding less advanced students and lacking comprehensive approaches to meet the varied needs of her entire classroom. This highlighted the intricate balance required in teaching between offering support and fostering autonomy and the potential challenges teachers face in this endeavor.

7.1 Michelle's Background

Michelle was a fifth-grade elementary teacher with a bachelor's in elementary education and 13 years of teaching experience (four of which were in upper grades). Before EDUCATE, her only exposure to PLD on cognitive activation or differentiation issues pertained to attending sporadic seminars on differentiation (EPI, line 370). She decided to participate in EDUCATE after recognizing her need for PLD on the emphases of the PLD program (EPI, line 184). She admitted that she had limited knowledge of how the two axes are defined or what they entail:

> When I first read the description of the PLD program, I thought that participating in it would be quite helpful to get some ideas, [to learn] what differentiation is, [how] to actively engage all levels of students… Such opportunities are scarce in schools. (EPI, lines 20-22)

Specifically, she mentioned having come across certain ideas related to the two axes before the PLD program. However, she recognized her limited understanding thereof (EPI, lines 43-44), which led to feelings of insecurity regarding these matters ("Before participating in the PLD program, I wasn't entirely confident if some things I had read or considered implementing were indeed correct." EPI, lines 45-46). In addition, her attitude towards the students was that it would be difficult to reach out to them or provide support in terms of cognitive activation and differentiation if they lacked interest. This attitude can be summed up in her motto: "If they are not interested in learning, no matter what you do, they won't learn!" (EPI, line 235), which, as will be discussed next, characterized her thinking throughout the program.

7.2 Reaction Level: Michelle's Evaluation of the PLD Program

In her end-of-program interview, Michelle acknowledged the value of her participation in the EDUCATE program, attributing it to various program characteristics. Although, initially, she was unfamiliar with the concepts of cognitive activation and differentiation (EPI, lines 186), some valuable elements of the program helped her grasp the concepts and convinced her of the feasibility of maintaining the cognitive level for all students. These elements included the video-club component (EPI, line 260; 301) in which she observed and discussed excerpts from her teaching with other teachers (EPI, lines 264-269); the exchange of ideas with them (EPI, line 267); the emergence of practices through discussions without being imposed by the teacher educators (EPI, lines 270-273); and the organization of those practices into meaningful and coherent learning schemas (EPI, lines 271-274). These program elements, as Michelle envisaged, played a significant role in her professional growth.

Furthermore, she argued that the gradual introduction of new concepts and practices in the program allowed for incremental experimentation with them across the three lesson phases (EPI, lines 187-189) and prevented her from feeling overwhelmed by the bombardment of new ideas (EPI, lines 191-192). As she explained, the gradual learning and step-by-step implementation of practices have contributed to her confidence in developing a lesson plan ("I now feel safer"[45], EPI, lines 195-197). She highlighted that this process mirrors a method that she would like to follow in her classroom to "foster a learning culture [similar to the learning culture developed during the EDUCATE sessions] to sustain the cognitive demands for all students" (EPI, lines 162-163; 166-167). In other words, the culture of the PLD program acted as a blueprint for the type of learning environment she aimed to cultivate within her classroom.

Despite the positive aspects of the program, she expressed a need for individualized feedback. It was challenging to obtain feedback promptly because of her constrained schedule and limited free time, which prevented her from consulting with the program facilitators, despite their availability (EPI, lines 308-309). She believed that receiving such individualized feedback could enhance her ability to closely

---

[45] This phrase was repeatedly used in her end-of-program interview.

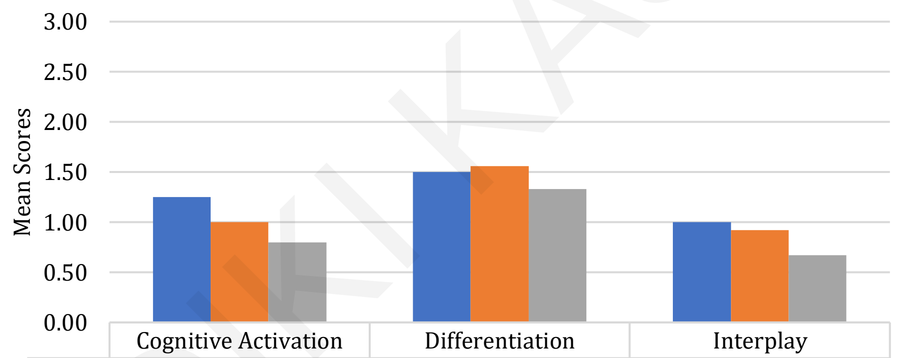
observe specific aspects of her teaching and effectively address any challenges she encountered. The next section examines Michelle's teaching performance in the culminating lesson, to determine if it reflects a positive picture, consistent with her overall positive reaction to the program.

7.3 Results Level: Michelle's Teaching Performance in Her Concluding Lesson

Figure 37 provides an overview of Michelle's teaching performance in her concluding lesson, showcasing the mean scores across three key axes: cognitive activation, differentiation, and their interplay, within each phase of the lesson.

Figure 37

Michelle's Performance in Cognitive Activation, Differentiation, and Their Interplay, in Each Phase of her Concluding Lesson



	Cognitive Activation	Differentiation	Interplay
TL	1.25	1.50	1.00
SAW	1.00	1.56	0.92
WCI	0.80	1.33	0.67

Note.
TL: Task Launching; SAW: Student Autonomous Work; WCI: Whole-Class Interactions.

Across all axes, Michelle's performance in her concluding lesson exhibited modest mean scores, ranging from just below 1.00 to the midpoint of the scale at 1.50. Her performance in differentiation achieved relatively higher scores compared to cognitive activation and the interplay of the two axes, indicating a greater emphasis on tailoring her teaching to accommodate the diverse needs of her students, especially during the student autonomous work phase, in which she attained her highest score of 1.56. In cognitive activation, the scores suggest that while Michelle began her lessons with tasks that engaged students at a higher level, scoring 1.25, maintaining this level of cognitive challenge appeared to be less robust during in student autonomous work and whole-class interactions phases, with scores of 1.00 and 0.80 respectively. This might be due to the nature of work in these phases, which could make it more difficult to sustain cognitive activation and attend to different student needs. Her performance in the interplay of cognitive activation with differentiation was the least successful, with scores ranging from 0.67 to 1.00. This suggests that integrating these axes throughout the lesson remained a significant challenge for Michelle, even though the practices for each axis independently seem to be applied more effectively.

Michelle's performance across all three axes was consistently higher during the task-launching phase. The student autonomous work phase—as already explained—maintained a stronger focus on differentiation compared to the other axes. In contrast, the whole-class interactions phase, although still addressing differentiation, displayed the lowest scores. This may be indicative of the inherent challenges in orchestrating the dynamics of whole-class interactions while simultaneously providing differentiated, cognitively challenging experiences to each student.

Michelle's teaching patterns, as indicated by the scores from her culminating lesson, show that she was somewhat more adept at differentiating her teaching, notably during student autonomous work. Nevertheless, these patterns also highlight areas ripe for PLD, particularly in enhancing cognitive activation and the interplay of both axes across all lesson phases, not solely at the beginning. From this outcome, one could claim that the EDUCATE program did not significantly help her improve her practice. Yet, without access to her mean performance scores from earlier lessons, it is not possible to conclusively determine whether her performance has stayed consistent, declined, or even experienced a minor improvement throughout the program. This is the point to which we turn next.

7.4 Behavior Level: Evolution of Michelle's Teaching Performance

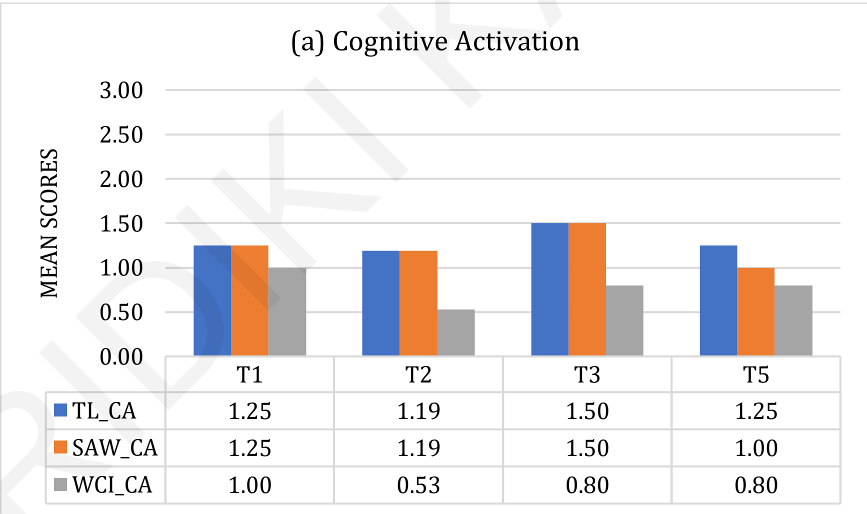
Following Kirkpatrick's model (2007), Michelle's teaching behavior was examined through both quantitative and qualitative methods. The quantitative analysis revealed that her mean performance across all axes fell below or neared the midpoint of the scale. Although there were slight enhancements in cognitive activation and differentiation—with differentiation showing a marginally better level of progress—her performance lacked consistent improvement across the board. The qualitative assessment shed light on a blend of moderate success in the application of certain differentiation practices, alongside challenges in maintaining cognitive activation and their interplay. Michelle's tendency to over-guide her students was seen as a manifestation of her "care" for her students but paradoxically limited their learning potential. Further elaboration on the findings is provided in the following sections.

7.4.1 Quantitative Analysis

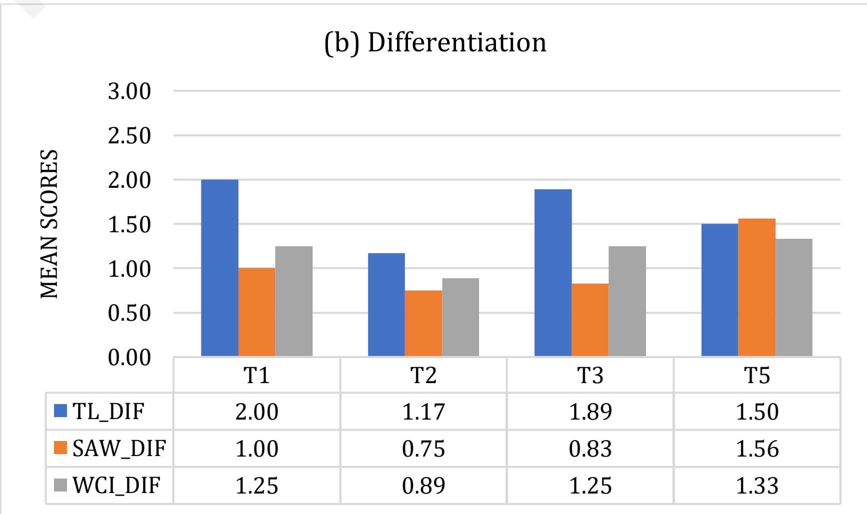
Figure 38 presents the results of the quantitative analysis of Michelle's coded videotaped lessons, focusing on cognitive activation, differentiation, and the interplay between the two. A preliminary observation highlights that her overall performance was consistently below or close to the mean value of the scale (1.50) across all aforementioned axes and throughout the entirety of the three lesson phases. Perhaps the only exception was a slightly upward trend indicating a minor, albeit somewhat erratic, improvement in student autonomous work.

Figure 38

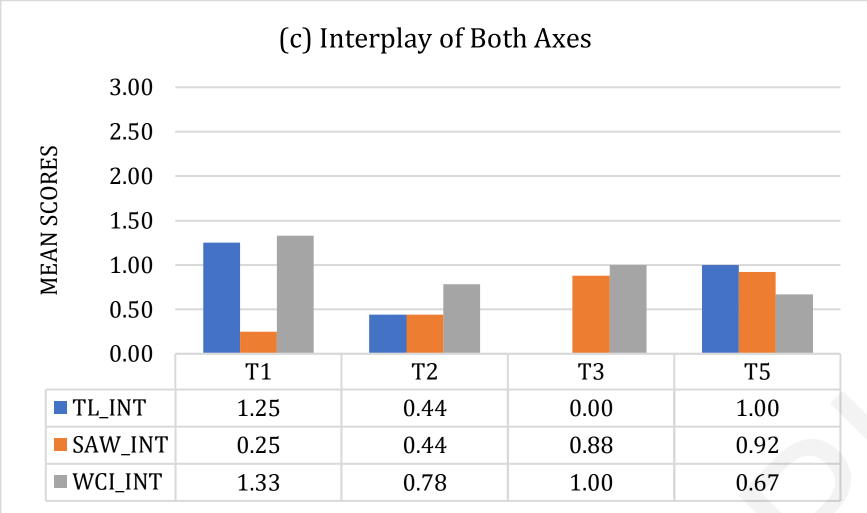
Michelle's Performance in (a) Cognitive Activation, (b) Differentiation, and (c) Their Interplay, Per Phase Across Her Lessons



	T1	T2	T3	T5
TL_CA	1.25	1.19	1.50	1.25
SAW_CA	1.25	1.19	1.50	1.00
WCI_CA	1.00	0.53	0.80	0.80



	T1	T2	T3	T5
TL_DIF	2.00	1.17	1.89	1.50
SAW_DIF	1.00	0.75	0.83	1.56
WCI_DIF	1.25	0.89	1.25	1.33



Note.
TL: Task Launching; SAW: Student Autonomous Work; WCI: Whole-Class Interactions.

Concerning cognitive activation (see Figure 38a), her performance displayed relative stability during the task-launching phase, hovering around the value of 1.20. An exception surfaced in her third lesson, where her performance in task launching reached the scale's mean (1.50). While there was some marginal amelioration in her performance in student autonomous work, this progression lacked consistency, especially considering that in her final lesson, it dropped to 1.00. In terms of the phase of whole-class interactions, her performance again exhibited notable variability but largely remained well below the mean of the scale (ranging from 0.53 to 1.00).

Addressing differentiation (see Figure 38b), the graph suggests that Michelle demonstrated a higher competence in differentiation in contrast to cognitive activation. This contrast runs contrary to her steadfast conceptualizations regarding differentiation and her evolving perceptions of cognitive activation across her participation in the PLD program, as detailed in the previous section. If compelled to provide a more nuanced assessment within the realm of differentiation, it can be postulated that her performance underwent a modest improvement concerning student autonomous work, along with a subtle progression in the phase of whole-class interactions.

Moreover, the dataset underscores a consistency in her performance within the interplay of the two axes (see Figure 38c), coupled with a marginal improvement in student autonomous work. Notably, the graphs depicting differentiation and the interplay of these axes (see Figures 38b and 38c) unveil a perplexing observation: a conspicuous higher score in her performance during the task launching and whole-class interaction phases in her initial lesson compared to the subsequent lessons. This intriguing peculiarity can be ascribed to the content and structure of her first videotaped lesson, which was exclusively centered around solely one challenging main task, encompassing only one occurrence of each lesson phase.

Overall, Michelles' teaching profile displays inconsistency across the three axes in terms of improvement. While she displayed a commendable effort to integrate the practices and concepts learned during the EDUCATE program, her performance suggested areas for further refinement. Deeper qualitative investigations can shed light on the nuances and intricacies of her teaching practice.

7.4.2 Qualitative Analysis

Table 14 portrays the outcomes resulting from the qualitative analysis of Michelle's videotaped lessons. The checkmarks in Table 14 signify her attempts to explore and experiment with certain ideas discussed in the PLD program. The implementation of multiple practices was mostly of moderate quality, as indicated by the light green color (e.g., "Presenting and discussing multiple solutions" or "Monitoring students' work and formatively assessing their needs"), while only few differentiation practices were gradually applied in high quality, as shown by the bright green color (e.g., "Implementing asynchronous work" or "Sequencing student solutions in a reasonable progression").

Michelle had been familiarized with certain tools yet faced many challenges as evidenced by the multiple grey- and red-coloured checkmark boxes. Despite systematic experimentation with certain practices, difficulties persisted, attributed to her tendency to spoon-feed students and steer them to the "right" answer. Prominent hurdles emerged in encouraging autonomous work and facilitating effective whole-class discussions. For instance, the practice of "Providing mathematical prompts to students to help them make some progress on the task and take up the challenge" during autonomous work was recurrently implemented but fraught with obstacles (see Table 14a). While she systematically made efforts to scaffold students without doing the thinking for them, struggles arose in avoiding excessive guidance (as indicated by the red color in her first three lessons). In her fourth lesson, she restrained her excessive intervention during the main task, lasting roughly half the lesson (depicted by the light green color). However, throughout the rest of the lesson, she reverted to grappling with the dilemma of whether to let students arrive at conclusions independently or provide them with too much guidance. This difficulty seemed to stem from her urge to chime in and tell them what to do. This observation could also explain the overall trend of inconsistency that emerged from the quantitative analysis of her lessons (see Figure 38).

In sum, the qualitative analysis suggests that Michelle utilized certain tools and improved in their use during the program; however, her overall progress did not demonstrate a specific direction in terms of change (either upward or downward). Specific themes emerging from the qualitative analysis of her lessons are further examined below.

The analysis of Michelle's lessons revealed three main themes pertaining to changes or lack thereof in her teaching: (a) *refining some practices*; (b) *experimenting with ideas while still overly guiding students*; and (d) *facing challenges in promoting the interplay of cognitive activation and differentiation*. The emerging themes are discussed below along with illustrative examples from her lessons.

Table 14

Michelle's Experimentation with Practices that Promote (a) Cognitive Activation, (b) Differentiation, and (c) Their Interplay, Across Her Lessons

D. COGNITIVE ACTIVATION				
TASK LAUNCHING	**L1**	**L2**	**L3**	**L4**
• Selecting mathematically challenging tasks	✓	✓	✓	✓
• Maintaining the cognitive demands of the task as presented to students during task launching		✓	✓	✓
• Asking students to read the task instructions and implement the think-pair-share strategy	✓	✓	✓	✓
• Discussing (key) mathematical ideas (to the task and/or the goal of the lesson) without reducing the demands	✓			
• Asking students to explain/make sense of mathematical symbols, sentences, or representations	✓			
• Handling unexpected student interference which could probably steal the thinking				
• Relaunching the task ending up initiating a whole-class discussion				
STUDENT AUTONOMOUS WORK	**L1**	**L2**	**L3**	**L4**
• Providing mathematical prompts to students to help them make some progress on the task and take up the challenge	✓	✓	✓	✓
• Pressing students for explanation/meaning, for making conceptual connections, or engaging in mathematical reasoning	✓	✓	✓	✓
• Posing mostly close-ended questions	✓	✓		
• Monitoring students' work and being more directive than needed	✓	✓	✓	✓
• Allowing another student to steal his pair's thinking	✓			
• Asking for explanations that focus on describing the procedure used		✓	✓	
• Telling the students precisely *how* to work (step-by-step) on the task	✓	✓	✓	✓
• Pointing out errors in students' work and remediating with procedures		✓	✓	✓
WHOLE-CLASS INTERACTIONS	**L1**	**L2**	**L3**	**L4**
• Eliciting or providing opportunities for student mathematical reasoning and meaning-making without reducing the challenge	✓	✓	✓	✓
• Rephrasing student ideas to address key mathematical ideas related to the task at hand (in interaction with students)	✓	✓	✓	✓
• Synthesizing and extending student contributions to address key mathematical ideas related to the task (in interaction with students)	✓	✓	✓	
• Presenting and discussing multiple solutions	✓	✓	✓	✓
• Comparing and evaluating multiple solutions		✓	✓	
• Bringing the class to the plenary at appropriate checkpoints				
• Having a clear direction during the discussion				
• Extending the discussion by posing a question that is more challenging than the task-at-hand				
• Handling unexpected student solutions				
• Handling unexpected student interference which could probably steal the thinking				
• Handling alternative conceptions around mathematical ideas				
• Providing guiding hints or ready-made answers	✓	✓	✓	✓
• Enacting IRE interactions when checking in plenary	✓	✓	✓	✓
• Introducing important mathematical ideas very early		✓		
DIFFERENTIATION				
TASK LAUNCHING	**L1**	**L2**	**L3**	**L4**
• Having the resources or materials available to be used by the students	✓	✓	✓	✓
• Providing a clear way of working on the task	✓			
• Providing materials to students and not ensuring that they have understood how they can be used while working on the task	✓		✓	

STUDENT AUTONOMOUS WORK	**L1**	**L2**	**L3**	**L4**
• Implementing asynchronous work	✓	✓	✓	✓
• Using learning aids				
• Using an entry card		✓		
• Monitoring students' work and formatively assessing their needs		✓	✓	✓
• Using a tic-tac-toe board as a final assessment activity				
• Using an exit card				
• Using anchoring activities				
• Encouraging multiple expressions of content, process, and/or product	✓	✓	✓	✓
• Grouping students according to their proficiency levels	✓	✓	✓	✓
WHOLE-CLASS INTERACTIONS	**L1**	**L2**	**L3**	**L4**
• Highlighting important mathematical ideas	✓	✓	✓	
• Sequencing student solutions in a reasonable progression	✓	✓	✓	
• Allowing students to start explaining any method/solution they want	✓			✓
F. THE INTERPLAY OF COGNITIVE ACTIVATION AND DIFFERENTIATION				
TASK LAUNCHING	**L1**	**L2**	**L3**	**L4**
• Explaining potentially unfamiliar non-mathematical aspects of the wording of the task or difficult words (context- or scenario-wise)	✓			
• Clarifying mathematical aspects of the task	✓			
• Activating relevant existing mathematical knowledge and strategies				
• Posing questions that indicate the level of support that students need in order to engage in the task without reducing the level of challenge				
• Spending no time clarifying the task instructions during task launching		✓	✓	✓
STUDENT AUTONOMOUS WORK	**L1**	**L2**	**L3**	**L4**
• Directing different types of questions to different (groups) of students				
• Circulating within all the groups attempting to attend to all students		✓		
• Using enabler(s)				✓
• Using extender(s)			✓	✓
• Sharing a strategy devised by a student group with the rest of the class, to support them make progress on the task and taking up the challenge				
• Using flexible grouping				
• Maintaining the demand for more advanced students and trivializing the thinking of less advanced students	✓	✓		✓
• Devoting more time to scaffolding students who are facing difficulties	✓		✓	✓
• Not establishing a routine for what the early finishers could do once they complete the main task	✓			
• Facing difficulties in supporting less-advanced students to make progress on the task	✓	✓	✓	✓
• Directing the exact same questions to all (groups of) students	✓	✓	✓	✓
WHOLE-CLASS INTERACTIONS	**L1**	**L2**	**L3**	**L4**
• Using incorrect or incomplete student solutions as resources for all students' learning	✓	✓	✓	✓
• Holding students accountable to attend to their classmates' thinking			✓	

Notes.
1. L#: Lesson ordinal number
2. The check mark illustrates that the teacher experimented to some extent with the practice in that lesson.
3. The color scale demonstrates the quality and frequency of the implementation (see Section 3.7 in Chapter 3).
4. Bright green: the practice was implemented in high quality and the demand was not reduced; Light green: it was implemented in medium quality and the demand was somehow reduced; Red: it was implemented in low quality and the demand was reduced; Grey: this practice hinders the particular axis.

Refining some practices. The first theme pertains to the improvement that Michelle demonstrated in certain practices, including (a) *implementing asynchronous work*; (b) *encouraging multiple expressions of content, process, and/or product*; (c) *using extenders*; and (d) *posing questions to press students for explanation, meaning, conceptual connections, or reasoning.* The first two practices stem from the axis of differentiation, and in their application, Michelle has notably improved her work, as depicted by the color transition from light to bright green in Table 14. The last two practices come from the axes of cognitive activation and the interplay with differentiation. Despite a mediocre implementation of the latter practices in her lessons (as seen by the shift from red to light green color in Table 14), she grasped onto these tools and gradually refined their use, mitigating some unfavorable aspects that previously compromised their efficacious enactment. The nuances and shifts observed in each of the practices are detailed next.

Implementing asynchronous work. Michelle consistently implemented asynchronous work across her lessons. Yet, there was a discernible progression in both the depth and breadth of its application from lesson to lesson. In L1, which focused on negative numbers, the application of this practice seemed rather cursory. For instance, upon completion of the main challenging task, some early finishers relayed their progress to Michelle. In response, Michelle briefly announced to the entire class that they could proceed to the next tasks in the student textbook ("Whoever finishes can move on to the next tasks", L1, min 11:43). However, only a few students responded to her call and most of the students were waiting for the rest of the class to complete the main task.

The concept of "asynchronous work" was introduced and discussed in VCS2 and VCS3, before the commencement of Michelle's L2, focusing on the commutative and associative properties of addition. On five separate occasions in this lesson, Michelle proactively signaled to the students in the plenary that they could proceed to the next tasks of their student textbook once they had tackled the main challenging task (e.g., "You have two to three minutes to complete the remaining task. If anyone finishes *before everyone else's allotted time*, work on Exercise 1." L2, min 15:38, emphasis in the original). Moreover, she introduced an additional, more complex task involving problem-solving. She encouraged students who completed the textbook tasks early to take on this added challenge ("I will give you another task for those of you who fill in the numbers in the equations. It requires more thinking than the previous task." L2, min 40:25). Thus, in L2, it was evident that Michelle took a more intentional and proactive approach toward the use of asynchronous work, even going as far as to introduce supplementary tasks.

In the last two lessons, Michelle persisted in the implementation of asynchronous work. She not only bolstered it through the incorporation of extenders and enablers but also encouraged diverse expressions of content, process, or product (the improvements in these practices are presented next). In L3, which was focusing on division, students were encouraged to use several methods to tackle a division task. Many ventured beyond traditional solutions, seeking and pinpointing alternative strategies. During L4, which delved into ratios and proportions, students worked simultaneously but at their pace on a tiered task (comprising of the enabler, the main task, and the extender). Furthermore, in both lessons, as students engaged with tasks that applied their recently developed knowledge, they were either at varied stages of the same task or worked in entirely different tasks.

Michelle's journey in the implementation of asynchronous work showcases a dynamic evolution. Starting with a rudimentary application in L1, she progressively deepened her approach. As she mentioned in her end-of-program interview, “Allowing them to work asynchronously is something I've adopted and do quite frequently” (EPI, line 80). By encouraging diverse problem-solving strategies, integrating more structured and layered tasks, and emphasizing individual pacing, she demonstrated her adaptability and responsiveness to multiple student needs and learning paces.

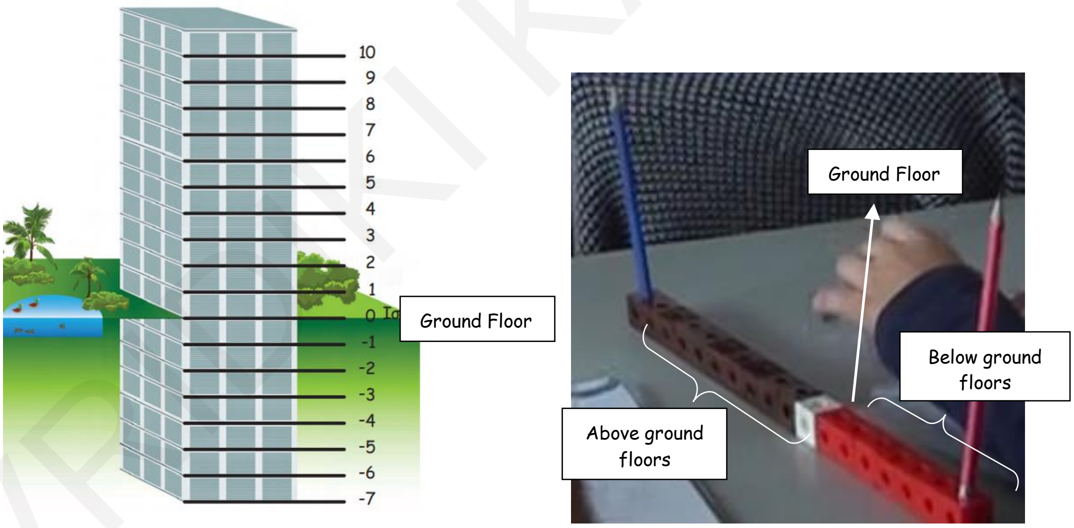
Encouraging multiple expressions of content, process, and/or product. Throughout Michelle's teaching trajectory, a clear emphasis emerges on encouraging students to express their understanding in multiple ways, be it through content, process, or product. This section delves into how she consistently and progressively honed on this practice, adapting it to address the diverse needs of her students. Particularly, in L1 (on negative numbers), Michelle encouraged students to use the diagram of the apartment building included in the main task or Unifix® cubes to solve the main task (see Figure 39):

> You can use the apartment building diagram. Also, I will give each student pair the cubes if you want to use them. The cubes above the white one represent numbers greater than zero, and the cubes below the white one represent numbers less than zero. (L1, min 2:43)

The use of the apartment building diagram could serve as a metaphorical visual representation for understanding positive and negative numbers (with floors above and below a reference point that corresponded to the 'ground level'). The cubes, on the other hand, offered a tangible way to visualize and manipulate these numbers. Notably, all students ended up using the same tools, the cubes. The interpretation of this outcome is twofold: First, students found the cubes beneficial for understanding the topic. Second, students might have felt peer pressure to follow what the majority of their classmates were doing or believed there was only one "correct" way to solve the task. If so, differentiation was not entirely realized. Nevertheless, it is commendable that Michelle recognized the diverse ways students process information, especially considering that differentiation methods had not been addressed in the preceding VCS1.

Figure 39

Apartment Building Diagram and Unifix® Cubes from Michelle's Lesson 3



(Mathematics Curriculum, 2016, Grade 5, Unit 9, p. 56, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

In L3, focusing on the addition properties, Michelle developed and enacted tailored tasks for three different groups of students (i.e., less advanced, intermediate, and more advanced, see Figure 40).

Figure 40

Michelle's Modified Versions of Two Tasks to Attend to the Needs of Different Student Groups in Lesson 3

Modified Versions of Task 1	Modified Versions of Task 2
<u>For less advanced students</u>	
3 + 5 + 7 10 + __ 12 + 8 + 9 9 + (12 + __) 25 + 7 + 5 __ + 30	Andrew has 7 yellow marbles, 8 green ones, and 3 blue ones. George has the same number of marbles. How many blue marbles does George have if the yellow and green ones total 10?
<u>For students who are in the middle</u>	
215 + 28 + 15 230 + ___ 340 + 20 + 13 (340 + 13) + ___ 172 + 18 + 90 ___ + 190 450 + 30 + 70 (30 + 70) + ___	Mr. John sold 45 muffins, 28 chocolates, and 35 profiteroles. Mrs. Helen sold the same number of sweets. How many profiteroles did Mrs. Helen sell if the other sweets totaled 80?
<u>For more advanced students</u>	
782 + 13 + 8 13 + ___ 1240 + 1310 + 400 (1240 + 400) + ___ 865 + 320 + 15 320 + ___ 3445 + 2100 + 45 ___ + 2100 10500+345+125+20500 (__+__) + (345+__)	Chris bought a tablet for 245 euros, a mobile phone for 127 euros, and a laptop for 415 euros. Alice paid the same amount to buy a camera, a tablet, and a mobile phone. How much did the camera cost if the other two items were priced the same as Chris's?

In Figure 40, it is evident that she varied both the number of equations, their numerical complexity, and the complexity of the word problems for each group. However, the differentiated versions of the tasks demonstrated some areas for improvement in terms of the mathematical content. In the first task, students were asked to complete the equations by filling in the missing numbers. However, the omission of the equality symbol, which is paramount for the accurate representation and solution of equations, could lead to ambiguity. Additionally, incorporating written instructions within the tasks could prove beneficial, as it would offer students additional support and direction.

Michelle's approach to differentiation by content in this lesson was praiseworthy. By adjusting the complexity and context of the two seemingly uniform tasks, to ensure that the challenge was customized to students' skill level, she was attempting to promote their engagement and understanding. For advanced learners, she increased the number complexity, while for those who might have found the main task overwhelming, she made the context and numbers more approachable. In doing so, Michelle struck a balance, attempting to ensure that every student was engaged at an appropriate level, neither under-stimulated nor overwhelmed.

In L3, on long division with a two-digit divisor, Michelle differentiated student work on the main division task both by process and product. Compared to her previous lessons, she *strongly emphasized* that students could choose their method of work:

> You have the freedom to represent your solution in various ways: drawing it out, explaining it in words, or any other method you prefer, using the Dienes blocks to illustrate your solution, performing mathematical operations, or even sketching a diagram. […] *Remember, using the Dienes blocks is optional;* they're available in your groups *if needed*. *All approaches are valid*. (L3, min 1:27, emphases in the original)

In this excerpt, Michelle allowed students to choose how they would process and understand the problem, by giving them options like using Dienes blocks or sketching a diagram. By allowing students to choose their method of work, she empowered them to take ownership of their learning process. Hence, the end product (i.e., the way students would demonstrate their understanding) could vary based on their choice. Indeed, seven different solutions emerged, all of which were presented and discussed in the plenary (first, two solutions provided by the main task and then, five student-crafted solutions, L3, min 27:16-43:57). She began by presenting and explaining the two solutions included in the student textbook (i.e., distributive property and long division). She then moved on to three solutions in which students used drawings and repeated addition, followed by a solution involving base-ten blocks. Finally, the formal division algorithm was discussed again (it was one of the methods that had already been presented but with different numbers). Discussing these varied solutions in a whole-class discussion further reinforced the idea that there were multiple ways to approach and solve a problem, promoting a growth mindset and valuing each student's thinking process. Yet, although she intended to give students a platform and acknowledge their efforts, a more preferable approach would have been to purposefully select some of the student-crafted solutions that would offer unique or insightful methods to discuss in the plenary. This would maintain the spirit of inclusivity and student engagement while also keeping the discussion focused and manageable.

In POI3, Michelle reflected on some unexpected yet positive outcomes from encouraging multiple expressions of process and product in that particular lesson:

> Indeed, some students were more engaged than I expected! For instance, a student who is typically very weak in math[ematics] managed to directly reach

> the division algorithm from the beginning, and I didn't even realize how. He solved it, explained it, and was very interested in understanding the other methods because he kept insisting, "Miss, I'm confused," "Miss, please explain it to me." This really impressed me. Another student solved it in four different ways! (POI3, lines 9-16)

This observation underscores the unpredictability and potential of student learning when given freedom of expression. Michelle's approach, which seemingly prioritized student autonomy and multiple solution methods, tapped into the latent abilities of students, even those traditionally labeled as "weak." The other student's ability to solve the problem in four different ways showcases the depth of understanding and flexibility in thinking that can be achieved when students are encouraged to explore various solutions. This incident serves as a testament to the power of differentiated teaching and the potential that lies within every student, waiting to be unlocked.

Using extenders. The use of extenders in L3 and L4 enhanced the asynchronous work, while also elevating the cognitive activation of early finishers. This stood in contrast to L1 (focusing on negative numbers), during which students had to wait for approximately six minutes for the rest of the class to complete the task (L1, min 13:23-18:05). Also, using extenders contrasted with L2 in which Michelle simply asked early finishers to proceed to the next tasks in their textbook, which entailed repetitive applications of the main task ("Those who finish proceed to the next task." L2, min 15:41).

In L3, which revolved around division with a two-digit divisor, Michelle urged students who completed their work early to *explore alternative solution methods and compare their solution methods in pairs* (L3, min 12:05). Although the idea of extenders had not yet been formally introduced at that point, Michelle appeared to be influenced by a brief discussion of VCS3 on how to engage students of varying levels through asynchronous work.

In her final lesson (focusing on ratios), which followed the introduction of extenders in VCS5, Michelle designed a more sophisticated extender for the main task, using the strategies of *problem-solving* and *combining different mathematical ideas* (i.e., ratios and liter conversions, see Figure 41d).

Figure 41

A Representation of (A) The Main Challenging Task Used in Michelle's Fourth Lesson, Focusing on Ratios, Along With (B) Its Enabler and (C) Extenders

a. Main task

Find out which of the following recipes will have the same flavor of mixed juice. Explain why.

Recipe A	Recipe B	Recipe C
The $\frac{1}{4}$ of the juice is apple juice and the rest is orange juice.	For every 1L of apple juice, there are 3L of orange juice.	For every 4L of orange juice, there is 1L of apple juice.

Recipe D	Recipe E	Recipe F
$\frac{3}{4}$ of the juice is orange juice and the rest is apple juice.	For every 1L of orange juice, there are 3L of apple juice.	For every 2L of apple juice, there are 6L of orange juice.

(Mathematics Curriculum, 2016, Grade 5, Unit 8 p.20, Reproduced with permission from the Curriculum Development Unit of the Cyprus Pedagogical Institute of the Ministry of Education, Sport, and Youth of Cyprus)

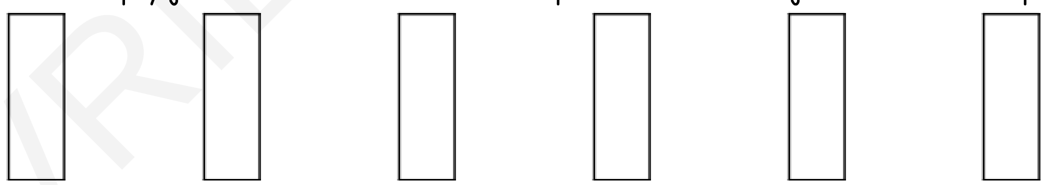
b. Enabler 1

These are empty juice containers. Use colors to present the mixed juices of each recipe.

Recipe A Recipe B Recipe C Recipe D Recipe E Recipe F

Enabler 2

These are empty juice containers. Use colors to present the mixed juices of each recipe.



Recipe A Recipe B Recipe C Recipe D Recipe E Recipe F

d. Extender

Mr. Andreas has orange juice and apple juice and wants to make 1 liter of mixed orange and apple juice. How many millimeters of orange juice and apple juice does he need to mix to get the taste of Recipe B?

Additionally, instead of generally encouraging students to find an alternative solution (as in L3)—which represents a typical extender not requiring particular preparation—Michelle offered the extenders to certain students. She distributed extenders to eight students whom she noticed had jotted down their solutions and did not require further assistance (by saying, "Work on this." L4, min 7:33; or asking them, "Are you done [with the main task]?" L4, min 9:04). Although she incorporated extenders in both of her lessons, unlike Pina, she never circled back to assess students' progress on the extenders as she was occupied with scaffolding less advanced students (e.g., “Two students had solved the extender. I feel that I didn't pay enough attention to check their progress and the accuracy of their solution. I got overly consumed with focusing on less advanced students.” POI4, lines 63-66).

Posing questions to press students for explanation, meaning, conceptual connections, or reasoning. Michelle's approach to questioning underwent notable evolution, as evidenced in her videotaped lessons. In her initial lessons, Michelle predominantly employed *closed-ended inquiries with a high degree of directiveness*. Particularly, in L1, which focused on negative numbers, a main task sub-question revolved around identifying the person positioned at a greater distance from the ground level of a multi-level building. During autonomous work, Michelle consistently delivered a comparable sequence of guiding questions to all students, regardless of their level and progress on the task. The following exchange exemplifies this type of interaction:

Michelle: Does the task ask you to find who is closer or who is farther away [from the ground floor]?
Student 1: [Finding who is] farther away.
Michelle: So, who is farther away?
Student 1: (No response)
Michelle: Where is Costas?
Student 1: (Indicating the floor labeled as -4 in the illustration of the apartment building depicted in his textbook)
Michelle: Where is Andrew? (Pointing to the floor labeled +6 in the apartment building illustration) Who is farther away?
Student 1: Andrew.
Michelle: Well done. (L1, min 6:38-7:21)

This excerpt illustrates how Michelle’s questioning style was formulated in a way that turned the task into a series of predetermined steps. She heavily guided the student (even by indicating the precise positions of the individuals, lines 9-10) and asked for monolithic answers, without allowing space for meaningful reasoning (lines 6, 9-10).

Following discussions on how to maintain the level of challenge during the task implementation, both in VCS1 and part of VCS2, Michelle's questioning underwent a transformation. Her questioning style evolved, *shifting towards more open-ended inquiries that encouraged students to provide explanations*. For example, the main task of L2 required students to compare two different methods for calculating the total value of purchased sports goods. The goal was to recognize that leveraging the commutative and associative properties of addition could facilitate their calculations. For students who arrived at conclusions supported by mathematical explanations, Michelle maintained the complexity of her questions. In contrast, when encountering students grappling with understanding, she tended to provide excessive guidance. An example of such interactions is indicated below:

Michelle: Which is the easiest method for you?
Student 1: [The method] of Alex.
Michelle: Why?
Student 1: Because he added the rackets with the shirt.
Michelle: Why is it easier to add the rackets with the shirt and not, say, add the rackets with the table?
Student 2: Because [by doing the first addition] you get a sum that can be added directly and easily with the other sum.
Michelle: Why can he easily add it to the other sum?
Student 2: Because 68 plus 12 gives us 80.
Michelle: What is going on with number 80?
Student 2: It's a rounded number.
Michelle: Bravo. Alex easily gets a rounded number.
[…]
Michelle: (Visiting a different pair of students:) Which of the two calculations is easier?
Student 3: Alex's method.
Michelle: Why do you think so?
Student 3: It is the easiest.
Michelle: Why is it the easiest?
Student 4: Because it is written.
Michelle: They did not use any paper or pencil. They both mentally calculate the sum. What is easier for you to find: adding 68 plus 184 *or 68 PLUS 12?*
Student 3: 68 plus 12.
Michelle: Yes! *Isn't it easier to add numbers that give a rounded sum?* (L2, min 9:55-10:38; 12:29-13-39, emphasis in the original)

In the first part of this episode (lines 1-14), Michelle posed open-ended questions that encouraged the students to articulate their reasoning. She typically began with an open-ended question to elicit their initial ideas followed by a series of "why" questions. The recurrent use of "why" questions was a central theme of the discussions held in VCS1 and the first half of VCS2. However, in the latter part of the episode (lines 15-26), she initially tried to sustain the level of challenge (line 18), but due to the student challenges and time limitations (immediately after, she initiated a whole-class discussion), she resorted to prompts that guided the students toward the desired answer (lines 23-24, 26). A similar pattern of interactions persisted in L3 as well, with

Michelle aiming to strike a balance between drawing out students' insights without imposing her ideas and her tendency to spoon-feed students (to be discussed next).

In L4, Michelle appeared to have done quite well with her questioning during the autonomous work of the main task. In particular, she made significant efforts to sustain the level of challenge even with students who faced difficulties, as the interaction between Michelle and a student that follows illustrates. To contextualize the episode, L4 revolved around the concept of ratios and proportional reasoning. The main task involved students deciding which recipes of mixed orange and apple juice have an identical taste (see Figure 41a).

Michelle: Are you stuck? Tell me what you're thinking.
Student: I don't understand what we're supposed to do. We need to find which recipes are the same...
Michelle: Bravo. Which recipes will have the same taste? For example, will this one be the same as this one? Read it carefully.
Student: No.
Michelle: Why?
Student: Because this [quantity] refers to apple juice and this [quantity] to orange juice. [i.e., the same number refers to a different type of juice in each recipe]
Michelle: Well done. [In this recipe], there's more apple juice, while [the other recipe] is more orangey. They won't have the same taste. [Noticing that the student was stuck:] Are you thinking of a way we can solve it? What portions of juice does the first recipe include?
Student: This recipe has 1 part apple, and the rest is orange.
Michelle: So, what part is the orange?
Student: Three [parts].
Michelle: Hmm, try to find out where else it might say that one part is apple, and three parts are orange. (L5, min 7:40-9:00)

In this exchange, Michelle attempted to assist a student facing challenges through her questioning. First, she elicited his initial ideas and understanding of the task (line 1). Then, she directed his focus to a certain recipe, offering a starting point for approaching the task (lines 4-5). Next, she prompted him for explanations (line 7) and interpretations of the ratios (lines 11-14), and finally, she assigned him a task to work on (lines 18-19). Despite Michelle's progress in questioning, in the second half of the lesson (focusing on less complicated tasks), Michelle leaned toward excessively guiding students (examples of such exchanges are provided in a subsequent theme). Perhaps, due to her need to solve the textbook tasks within a limited timeframe (i.e., a lesson), she started offering more assistance to students than needed. All in all, she made significant improvements in questioning but still faced some challenges.

Experimenting with ideas while still overly guiding students. Michelle had a genuine desire to care[46] for her students. Driven by this desire, and apparently not tolerating student struggle, she experimented with various practices discussed in the project to support students in successfully completing the tasks, removing, however, any learning barriers, and limiting their opportunities for productive struggle. This tendency towards overly guiding the students was evident in all her lessons. This was captured in Table 14, in which the application of multiple practices related to cognitive activation and its interplay with differentiation was denoted in red color, while practices with grey shading that inhibit these axes were extensively applied, especially during autonomous work and whole-class interactions.

Her well-intentioned efforts had unintended consequences. While caring was essential, Michelle failed to strike a balance between providing support and productive struggle. She inadvertently reduced the students' challenge, by (a) *telling them precisely how to work on the task during autonomous work*; (b) *pointing out errors in their work and remediating with procedures*; and (c) *going over the task solutions without much sense-making by the students during whole-class discussions*. These behaviors are explained next.

Telling the students precisely how to work on the task during autonomous work. As already discussed, during autonomous work, Michelle experimented with questioning. However, in many lesson episodes, she demonstrated (part of) the solution method to the students outright. She cared for her students and wanted to help them succeed; therefore, by providing detailed guidance or oversharing hints that made the solution obvious and procedural, she would ensure that they would not struggle or fail.

This mainly occurred when students did not directly respond to her initial open-ended question, making her feel that they faced insurmountable difficulties that should be immediately addressed. For instance, in L1 on negative numbers, she approached a student, who was using different-colored cubes to represent floors above or below ground level, to elicit her thinking on the task, leading to the following exchange:

1	**Michelle**	What did you think so far?
2	**Student:**	(Does not respond)
3	**Michelle:**	(Starts guiding the student:) Where is John’s car?

[46] The idea of “care” is adopted by Nel Noddings' work (2001), entitled "The Caring Teacher". The term “care” emphasizes the profound impact of a relational approach in education, where genuine care becomes the cornerstone of effective teaching and meaningful student engagement. True caring (in the sense that is used in Noddings’s work) might involve understanding when “the caring individual”, namely the teacher, should step back and let “the cared-for”, that is the students, grapple with challenge, as this can be a form of caring in itself.

Student: (Reads the task instructions from the beginning]
Michelle: [interrupts:] On which floor in the basement [his car is]?
Student: On the 5th [basement].
Michelle: Find and show me the 5th basement. (Points to the pink cubes representing the floors below the ground floor, see Figure 42)
Student: (Counts the cubes and points to the cube indicating basement 5)
Michelle: Great. Place a pencil on this cube. (Takes a pencil and puts it in the cube herself) Where does John want to go? He's going up using the elevator… where is he going? (Indicates the phrase of the task instructions answering the question)
Student: (Reads the phrase:) "To the 6th floor"
Michelle: Place this pencil on the 6th floor. Which is the 6th floor? (Points to the brown cubes representing the floors above the ground floor, see Figure 42)
Student: (Counts the floors)
Michelle: (Interrupts upon noticing that the student considered the cube indicating the ground floor as 1 instead of 0) No, that's the zero floor (i.e., the white cube, see Figure 42). (Places the pencil herself on the sixth floor) How many steps will he take to go from here to there? Count the steps.
Student: (Does not respond)
Michelle: Take the pencil and move it until it meets the other pencil (see Figure 42). How many steps will one pencil take until it reaches the other pencil?
Student: 1, 2, 3, 4, 5, 6…
Michelle: (Interrupts upon noticing that the student skipped the ground floor) What about this cube? John can't jump over the ground floor. Place the pencil back here and count.
Student: 11 [steps].
Michelle: Well done. That's how we find the answer. (L1, min 19:13-21:29)

Figure 42

Example of Unifix® Cubes *Usage During Student-Michelle Interaction in the Main Task of Lesson 1*

In this excerpt, Michelle initiated the discussion with an open-ended question (line 1), yet upon noticing the student's delayed response, she transitioned to closed-ended questions (line 3). To prevent wasting time because the student's response did not align with Michelle's initial query (line 4), she intervened with an even more directive question (line 5), delineating precise steps to undertake (lines 10-12, 15-16). Confronted with the student's mathematical errors (i.e., considering the ground floor as the first floor, lines 17-21; or omitting the ground floor when counting, lines 25-28), Michelle promptly provided the correct solution without giving the student time to reflect on her work. As a "caring teacher", Michelle possibly thought that the student needed this kind of support due to her errors and the problems she faced to move on. Similar dynamics persisted in L2.

In L3, which involved long division with a two-digit divisor, Michelle's tendency to excessively assist students became more pronounced. The lesson content might have contributed to this tendency. Teaching the long division method can be challenging due to the necessity of conveying multiple concepts (such as understanding multi-digit numbers, division, multiplication, and subtraction) in an easily comprehensible manner. For example, after prompting a pair of students to interpret the task information, Michelle asked them to determine their solution approach. The students chose to solve the division by using a drawing, and the ensuing interaction unfolded:

Michelle: How will you solve this by using a drawing?
All students: (No response)
Michelle: Let's say I have a box. (Takes a piece of paper and draws a box) How many cookies will I put in it?
Student 1: Twelve.
Michelle: (Writes "12" inside the box) If I add one more box, how many cookies will I put in? (Draws another box and writes "12" inside) How many cookies do we have in total?
Student 1: Twelve…twelve plus twelve…ten plus ten gives us twenty, and two plus two equals four, so twenty-four.
Michelle: So, what will happen if I add another box with twelve? We have twenty-four…plus twelve? Continue this drawing. (L3, min 4:08-5:27)

Following the students' delayed response to her initial open-ended question (lines 1-2), Michelle started drawing the diagram for solving the division (line 3) and doing all the work herself (lines 6-8). In general, Michelle perceived the students' silence or delayed response as an indication of their difficulty rather than allowing for processing time. In doing so, she seemed to probably have adopted an ill-defined approach to *caring* that did not allow much space for students for productive struggle.

In L4 (on ratios), Michelle made a notable change in maintaining the cognitive rigor of the main task (see Figure 41). As already explained under the first theme, she refrained from providing prefabricated answers and step-by-step guidance and attended more to students' thinking. Perhaps the use of enablers contributed to this, as Michelle would provide them to those who were struggling, thus avoiding giving over-guidance herself. Nevertheless, the important mathematical ideas were not highlighted during their work on the main task, and some students encountered difficulties in the remainder of the lesson, where they were supposed to apply what they had learned to other tasks. In light of these difficulties, Michelle reverted to her typical "caring" style to scaffold students. The ensuing example illustrates Michelle's guidance to a student working on a task involving the identification of fractions and ratios of the beads in a necklace:

Michelle: How many beads do you have in total?
Student: Twelve
Michelle: So what fraction are the blue ones?
Student: Four…
Michelle: Four twelfths. Write it down. What fraction are the red ones?
Student: Eight.
Michelle: Eight what?
Student: Eight beads.
Michelle: What fraction is that? You need to express it as a fraction.
Student: Eight fourths.
Michelle: Eight fourths? How many are there in total?
Student: Eight.
Michelle: There are twelve in total. So, eight out of twelve. How would you express that mathematically?
Student: Eight twelfths.
Michelle: Eight twelfths. Write it down. Now, what is it asking you here?
Student: (Reads:) "What is the ratio of the red beads to the blue beads?"
Michelle: How many red and how many blue beads are there?
Student: Four red and eight blue.
Michelle: (Noticing the student's error:) It's red to blue, not blue to red.
Student: Eight fourths.
Michelle: When it asks for a ratio, we write "eight to..."
Student: Eight to four.
Michelle: Bravo! (L5, min 27:35-29:19)

Certainly, the episode underscores Michelle's over-guidance to students, marked by the use of closed-ended guiding questions (lines 11, 18) and the provision of directive cues (lines 13-14, 20, 22). This approach, while facilitating immediate progress, raises concerns about fostering true mathematical understanding. In essence, her focus seemed to be on task completion, overshadowing the cultivation of deep conceptual mastery. Yet, her oversharing and over-guidance might be interpreted as an indication of her "care" for students by preventing them from experiencing potential difficulties or failures.

Pointing out errors in students' work and remediating with procedures. In numerous instances, Michelle's improvement was constrained by a tendency to ensure a smooth error-free lesson. In the sense of caring, such lessons were equated with students' success on the assigned tasks. However, in doing so, she might have become overprotective and have overlooked the value of errors as pivotal moments for student understanding, therefore, limiting student cognitive engagement.

Specifically, in cases in which a student made a careless or sloppy mistake due to fatigue or distraction, she promptly pointed out the error and superficially instructed them to rectify it swiftly. For example, she stated, "[This is] *incorrect*! Where did you get eighty five? This is seventy five, not eighty five." (L2, min 26:08). Conversely when addressing students who were progressing correctly, she would affirmatively state, "Well done!" or "*Correct*!" (L2, min 22:24; 26:08). Notably, she often initiated such discussions using the words "correct" or "incorrect", and she intervened in students' work by either erasing or writing hints to scaffold their progress. This approach not only lessened the cognitive demands but also shaped student perceptions of error-making.

In other instances, Michelle provided direct instructions for *accurate* problem-solving. Unfortunately, this approach often fell short of cultivating genuine conceptual understanding. This tendency was especially noticeable during L2 and L3, which revolved around numbers, operations, and algorithms. For instance, in L2 (on addition properties), she would circulate in the classroom, assessing students' work by marking their answers as either correct or incorrect, and addressing errors through procedural rectification, as evidenced below:

Student: Ma'am, I need help...
Michelle: It was expected to face challenges when you're doing it incorrectly! (Erases what the student had written in his textbook) Are you sure about your answer?
Student: (Hesitantly:) No.
Michelle: Let's see. We first add what is in the parenthesis [i.e., 60+50]. What is in the parenthesis? Six plus five?
Student: (No response)
Michelle: (Persistently asking:) Six plus five? Six plus five?
Student: Eleven.
Michelle: (Insisting:) So, sixty plus fifty?

12	**Student:**	One hundred eleven.
13	**Michelle:**	(Surprised:) A hundred-what?
14	**Student:**	One hundred one.
15	**Michelle:**	Add the units separately. How many units does number sixty have?
16	**Student:**	Zero.
17	**Michelle:**	How many units does number fifty have?
18	**Student:**	One.
19	**Michelle:**	Where did you find that? How many units does number fifty have?
20	**Student:**	Zero.
21	**Michelle:**	So, zero plus zero units. How many units will it have?
22	**Student:**	Zero.
23	**Michelle:**	Write it. How many tens does sixty have?
24	**Student:**	Six.
25	**Michelle:**	How many tens does fifty have?
26	**Student:**	Five.
27	**Michelle:**	So, in total?
28	**Student:**	Eleven. So, [the sum is] one hundred ten.
29	**Michelle:**	What should you add next? Think about it. (L2, min 24:34-26:01)

At the outset, Michelle's underscoring of the incorrect nature of the student's work appeared to be laced with exasperation (lines 2-3). The student doubted the validity of his method (lines 4-5) and Michelle embarked on providing him with a series of step-by-step instructions for the sequence of operations, compelling him to adopt a specific strategy through the repetition of the same question (line 9). In L3 and L4, Michelle's approach closely resembled that of L2, by pinpointing and elucidating students' mistakes (e.g., "Can you identify your *error*? You made an *error* because you added the two first and then added twelve again, leaving these two out. You've *made an error* in how you group the cookies." L3, min 17:58).

From the examples provided, it is evident that Michelle's understanding of "caring" may have been skewed towards protecting students from committing errors (in the future) and shielding them from challenges rather than navigating them through these struggles. By emphasizing the correct solutions only and not delving into the underlying processes or possible errors, she might inadvertently have conveyed that only the "right answers" hold value, sidelining the importance of the learning process itself.

Going over the task solutions without much sense-making by the students during whole-class discussions. As a caring teacher, Michelle had almost all students' ideas—and especially the correct ideas without pointing out any mistakes —to be presented. Despite her apparently good intentions, this practice could be seen as a misinterpretation of caring, in the sense that she was also trying to protect her students from being exposed to erroneous solutions. Genuine caring extends beyond merely ensuring that students arrive at the correct answers; it emphasizes the richness of discussions that promote understanding. However, Michelle was glossing over task solutions in whole-class discussions without actively involving students in the sense-making process.

In particular, she would initiate whole-class discussions by asking the students to present their solutions. Despite requesting explanations (using "why" questions), students were describing processes. Michelle accepted these descriptions and ultimately offered explanations herself, aiming to ensure that the accurate ones *were heard*. For instance, in L1 (focusing on negative numbers), two different solutions were shared for the task question that asked how many floors John would go up to ascend from the fifth floor below the ground to the sixth floor:

Michelle: How many floors will he go up?
Student 1: John takes the elevator up five floors.
Michelle: Five floors? *How* did you figure that out? What floor is John's car on? He is on the fifth floor below the ground. Come and circle the number on the interactive board. (She indicates -5) Where does he want to go? Where is his office?
Student 1: On the sixth floor.
Michelle: Good. Circle it [on the interactive board]. *How* will you find out how many floors he will go up from where he is to his office?
Student 1: Eleven [floors].
Michelle: *How* did you find that out?
Student 1: I counted the floors.
Michelle: One student, instead of counting one by one, did a mathematical operation. *What* operation did you do? *Explain* it to us.
Student 2: [I did an] addition.
Michelle: *Why* did you do addition? *What* did you add?
Student 2: I added minus five to six.
Michelle: You added the total number of floors below ground level to the total number of floors above ground level. So, from the ground level to the car is five floors, and from the ground level to the office is six floors, thus a total of eleven floors. (L1, min 23:30-26:15, emphasis added)

Despite presenting multiple solutions, the build-up of this discussion was primarily procedural. Michelle mostly requested accounts of how they arrived at their answers (using "how" and "what" questions, lines 3, 8, 11, 14, 16) and the students provided descriptions of steps (lines 12, 17). Posing a "why" question followed by a "what" question (lines 14, 16) resulted in getting a numerical response (line 17). Michelle accepted that and provided the mathematical explanation herself. This excerpt highlights Michelle's tendency toward proceduralizing mathematics.

In L2 (covering addition properties), Michelle guided students even more, which could be attributed to her perception of the content as being simpler or due to her attempt to attend to differentiation in this lesson (after VCS2). This contrasts with L1, where the primary focus was on cognitive activation (after VCS1). For instance, she initiated a whole-class discussion by stating the conclusion that the students were supposed to arrive at, with only minimal contribution from the students:

Michelle: *It is easier to add numbers that yield a rounded result. Can we change the order of numbers in a mathematical sentence?*
Student 1: Yes.
Michelle: Why?
Student 1: Because when you do three times four...
Michelle: (Interrupts and corrects:) When *adding* three plus four…
Student 1: It's the same again [i.e., the result].
Michelle: Exactly. It doesn't matter if we add three plus four or four plus three. Even if we have twenty numbers, we can add them in whatever order we want.
[...]
Michelle: (After presenting how they solved mathematical sentences with three addends in two different ways) What do you observe? Why are only two students raising their hands?
Student 2: Even if I change them, the result is the same.
Michelle: What are you supposed to change?
Student 2: The numbers.
Michelle: *The order of the numbers, not the numbers themselves. So, when I change the order of the numbers in addition, the result is the same. Write it where it says, "What do you observe?".*
Student 3: What should we write?
Michelle: *Write that "We observe that no matter the order in which we place the numbers, the result is the same."* (L2, min 13:34-14:10; 29:35-30:14, emphasis added)

In this episode, Michelle was driven by her tendency to lead students toward correct answers (lines 1, 8-10, 18-19). Despite experimenting with the practice of asking the students to share their observations and generalize (lines 13-14), upon realizing limited student participation in the discussion (line 15) and receiving an incomplete response (line 17), she promptly took the initiative to explain (lines 18-20) and tell them what to write in their textbooks (lines 22-24). The influence of the PLD program's ideas, particularly the extensive emphasis on employing "why" and "what do you observe" questions, was evident in this excerpt. However, the tendency to over-guide students toward providing correct answers prevailed in her approach. This could also be another manifestation of Michelle's sense of "caring"; she aimed to meticulously verify the content written in student textbooks, ensuring its accuracy and that it was error-free.

Drawing from discussions held in VCS3, in L3, Michelle experimented with presenting and comparing multiple student solutions on the division problem with a two-digit divisor in the whole-class setting. Specifically, seven different solutions were shared in the plenary (as described earlier). Throughout the sharing, she consistently prompted students to rephrase their peers' ideas, provide explanations, and compare different solutions. However, when students' explanations or connections fell short, Michelle displayed the tendency to swiftly provide the correct explanation herself.

This was typified in the following example in which the method of the long division algorithm with a two-digit divisor was presented in the plenary. The episode started with Michelle asking a student to explain the algorithm. Interestingly, the student connected this method to their already known division algorithm with a *one-digit* divisor, but Michelle left these connections to go unnoticed and pressed him to draw connections and comparisons between the distributive property and the division algorithm. Immediately, the student began to elaborate his reasoning, but Michelle constantly interrupted him, guiding him to provide the explanation she expected to hear. Despite the student's initial assertion of a difference in approach, Michelle insisted on the similarity in reasoning between the two methods, emphasizing the breakdown of numbers in both cases. Ultimately, Michelle explained the similarities and differences between the two methods to the class ("It is the same reasoning, but he did it vertically because he found how many groups fit in the hundreds, then in the tens, and finally, how many are left over," L3, min 31:11-33:14). Overall, the over-guided linking of multiple solutions in pro-forma ways, along with the abundance of solutions presented (i.e., seven in total), muddled the mathematics and appear to have caused discomfort for less advanced students, who struggled to keep up with all of them.

In L4 (which focused on ratios and proportions), there was a noticeable improvement in how whole-class discussions were orchestrated around the main task of the lesson. Unlike in the previous lessons, Michelle was actively listening to students' ideas and using meaningful questions to support and extend students' understanding and reasoning. This positive shift is evident in the following excerpt:

Michelle: Let's review the results of your work. Which recipes are the same?
Student 1: [Recipes] B and F.
Michelle: Well done. Why?
Student 1: Because recipe B states that "for every liter of apple juice, there are three liters of orange juice." Recipe F states twice as much because, "for every two liters, there are six liters of orange juice."
Michelle: Well done! Because the ratio is the same. Is there any other recipe that fits this ratio?

Student 2: No.
Michelle: Look more carefully.
Student 3: [Recipe] A matches [recipe] B.
Michelle: Why does recipe A fit in this ratio?
Student 3: Because it's one-fourth. There is one liter and three liters, which will make four when combined.
Michelle: Good job. When we combine one liter with three liters, it makes four liters [in total]. So, the apple juice is one-fourth, like in recipe A. Anything else?
Student 4: [Recipe] D.
Michelle: Why recipe D?
Student 4: Because it's three-fourths orange juice and the remaining is apple juice.
Michelle Well done, so it's the same ratio. What about the other recipes?
Student 5: [Recipe] E.
Michelle: What about this recipe?
Student 5: It's one to three. One liter of orange juice and three liters of apple juice.
Michelle: Does it match the others?
Student 5: Not so much, because there is one part of orange juice instead of apple juice.
Michelle: Way to go! This will have a strong apple taste, it has more apple juice in it, so it doesn't match. Why not Recipe C?
Student 6: Because the whole amount is five.
Michelle: Well done! Because the apple juice will be one fifth while we want it to be one fourth. (L5, min 17:08-19:45)

In this vignette, Michelle consistently prompted the students to justify matching ratios with certain recipes. She encouraged them to analyze the ratios and ingredients of these recipes to justify their conclusions (lines 3, 12, 19, and 31). Students appeared to have grasped the concept of comparing the size of quantities effectively (lines 4-6, 13-14, 25-26, 28-29). However, the presentation of ideas and solutions in an abstract manner, without incorporating any visual representation of the students' verbalized reasoning, resulted in a straightforward delivery of solutions. Annotating the ongoing discussion with visual aids, such as colors, symbols, words, or drawings, could enhance engagement and understanding for a diverse range of students.

Despite Michelle's care for her students' academic progress, her methods often leaned towards over-guidance, frequently steering students towards expected answers rather than fostering independent exploration. While she engaged students in discussions, the depth of these interactions often fell short, focusing on correct answers rather than the richness of the learning journey. This approach, though well-intentioned, might have inadvertently limited opportunities for cognitive challenge, potentially misrepresenting the true essence of caring in education.

Facing challenges in promoting the interplay of cognitive activation and differentiation. The third theme concerns Michelle's challenges (*observed* or *perceived*) in promoting both axes. As in the other teacher cases, the challenges are organized into two categories, depending on when they occurred or were treated. The first category included *challenges that were mitigated or addressed* during the PLD program, such as attending to less advanced students while neglecting more advanced ones and facing time constraints. Gathered within the second category were her *unresolved challenges*, including using enablers, nurturing a classroom culture for enhanced cognitive activation and differentiation, addressing student indifference; and lacking a clear direction toward the lesson objectives. We elaborate upon each challenge below.

Challenges Addressed or Mitigated During the PLD

Attending to less advanced students while neglecting more advanced ones. Michelle prioritized less advanced students, considering that they required constant attention and guidance to catch up with the rest of the class and grasp basic concepts (POI2, lines 54-55; POI3, lines 22-24; VCS7, line 256), while more advanced students were assumed to need less support and independently (POI2, lines 45-48; POI4, lines 31-32). This differentiation in attention and support could also be considered another indication of her sense of care. In VCS3, she desperately admitted grappling with this challenge:

Michelle: There are some students that no matter how much time you give them, they just won't focus!
TE1: What should the teacher do with those students?
Michelle: As soon as you leave them, they stop [working on the task]. They want you [i.e., the teacher] to be their tutor! Should I neglect the majority of the class for those three students? (VCS3, 106-112)

Michelle questioned the appropriateness of allocating a disproportionate amount of time and attention to the students (lines 5-6). Her concern highlighted the dilemma of supporting without compromising the learning experience of the larger group.

To provide extra support to struggling students, Michelle organized her classroom in homogeneous groups based on ability and performance (e.g., PRI2, lines 36-37; PRI3, lines 13-22) Notably, in her first three lessons, students who finished quickly would engage in conversation with one another while *waiting* for others to catch up (e.g., L1, min 11:45; L2, min 24:09; L3, min 6:53), whereas those facing challenges would experience frustration, become off-task, play with the materials, and occasionally resort to disruptive behavior while *waiting* for Michelle's aid (e.g., L1, min

16:34; L2, min 22:38; L3, min 19:53; EPI, 144-146). Therefore, while Michelle was working with a student group, the other groups always had to wait for her to *visit* them.

However, categorizing students based on pre-established measures of skill or performance implied a fixed ability level, disregarding potential variations or individual growth paths. The concept of flexible grouping, discussed in VCS2, emerged as a potential solution to the challenges faced in fixed homogeneous groupings. However, when asked whether she would consider implementing it, she disapproved it (PRI2, lines 40-42). Several factors might have contributed to her reservations about implementing flexible grouping. These encompassed *potential difficulties in classroom management* when handling varying group dynamics (e.g., "they play with the materials" POI1, lines 51-54; "they argue with certain students" VCS2, lines 315-320; "they copy from each other" POI4, lines 101-102); *resource constraints*, such as limited classroom space (VCS3, lines 175-176; VCS9, lines 422-430); and *complexities inherent in addressing the difficulties and the indifference of less advanced students in mixed-ability groups* (VCS2, lines 32-49; VCS3, lines 418; VCS4, lines 185-192). She argued that homogeneous grouping simplified their work because it increased the support that she would provide to less advanced students since more advanced ones could assist each other and progress ("Less advanced students will be helped solely because the teacher will focus on them. […] Homogeneous grouping is easier for me as a teacher." VCS7, lines 255-256; 263)

To address this challenge, Michelle gradually introduced *asynchronous work* (EPI, line 80) and used *enablers and extenders* (EPI, lines 129-132). These strategies helped mitigate the issues associated with exclusively grouping students by ability. Asynchronous work allowed students to progress at their own pace, aiding struggling students in catching up and enabling advanced students to delve deeper. Enablers offered extra support for those needing it, while extenders offered added mathematical challenge for advanced learners. Collectively, these practices helped Michelle to somehow create a more inclusive learning environment, tailoring tasks to each group.

Facing time constraints. Time constraints posed a significant challenge, as she grappled with the dual demand of effectively cognitively activating and responding to all students. This challenge was first mentioned after her first experimentation in L1 and L2 ("Time constraints and dealing with different levels of students. How can I organize time and the class to accommodate students of varying skill levels?" VCS3, lines 336-340). Michelle emphasized the *time pressure* that occurs when there is an unfavorable ratio between the required time to complete a task and the available time.

Moreover, for Michelle, *time management* was intertwined with *how the class was organized* and *the tasks used* (VCS6, lines 71-72). She recognized the complexity of ensuring that the chosen tasks were appropriate for the allocated time frame. Certainly, challenging tasks do require more time for in-depth exploration and understanding, compared to simpler, routine tasks. Additionally, the classroom organization into homogeneous groups and the constant need for students to wait for her assistance whenever they faced difficulties (as explained earlier) made it harder to effectively manage time. This challenge was somehow addressed by the end of the PLD program:

> At times, I encountered time constraints, limiting the amount of attention I could give to each student. Enablers proved valuable, as they allowed me to step back while students remained engaged, fostering independence. Also, using extenders benefited more advanced students, granting them autonomy, and helping me to dedicate more time to those who require more assistance. (VCS9, lines 226-231)

The utilization of enablers and extenders proved effective in helping Michelle manage time while catering to diverse student needs. With these tools, student engagement was not always dependent on her availability.

Unresolved Challenges

Using enablers. In L1, the students heavily relied on Michelle's scaffolding to start working on the task. For instance, while working autonomously, three different pairs of students called out "Ma'am, we need help" or were consistently raising their hands, awaiting her assistance. Similar interactions were evident in L2 and L3. The tool of enablers was embraced by Michelle and supported her in maintaining cognitive activation with less advanced students; yet, it did not assist them in making significant progress on the task.

Michelle used enablers for the first time in L4 (focusing on ratios). She designed *multiple enablers* that catered to varying levels of student proficiency. Both enablers featured *representations* of empty juice containers. The first enabler had containers that were pre-divided into quarters, aimed at aiding students with greater difficulties, while the second required students to decide how to divide the containers, targeting those with fewer challenges (see Figure 41b). Despite designing two enablers, only Enabler 1 was used by five students who were not making any progress on the task. Michelle did not immediately distribute the enablers to the students; instead, she engaged them in initial discussions to gauge their understanding of the task. Below is an illustrative example of such an interaction between Michelle and a student:

Michelle: What do you think we should do to solve the task?
Student 1: (After a couple of seconds:) I don't know.
Michelle: Read what it says here [showing Recipe A].
Student 1: "One-fourth of the juice is apple juice, and the rest is orange."
Student 2: (Interjects:) [The recipes] A and D [have the same flavor]!
Michelle: (Addressing Student 2 and stopping him from sharing his answer:) Shh! Please, write down your thoughts, and don't chime in. Later, I want you to tell me how you arrived at it. (Addressing Student 1 again:) So, one-fourth is apple, and the rest is orange. Great. Read the other recipes and tell me if there's another recipe that tells you the same thing, maybe in a different way. I'll let you think about it, and if you can't figure it out, raise your hand. (L5, min 3:28-4:19)

In this excerpt, Michelle made efforts to extract the student's initial ideas, allowing him time to respond (lines 1-2) and work on the task (lines 11-12). His main difficulty was his inability to comprehend the task instructions (line 2). This struggle could stem from his limited exposure to the context of crafting and tasting mixed juices of various strengths. Considering that Michelle mentioned multiple times over the sessions that most of her students had limited real-world experiences, this student may not have been able to make sense of the task context and mathematize it. Notably, Michelle refrained from providing ready-made answers and prevented another student from sharing his ideas (lines 6-8). Instead, she attempted to clarify the task requirements, although not entirely successfully, since the student continued to face difficulties:

Student 1: Ma'am, they're all the same. They all talk about orange and apple juice.
Michelle: Do they all refer to the same ratio? [Leaves and goes to get the enablers] Here's what you can do: Use this to display the mixed juice of each recipe. Use red color for apple juice and orange color for orange juice. (L5, min 6:23-7:03)

The student was encountering unaddressed difficulties in understanding the wording or the mathematics of the task (lines 1-2). Her decision to offer the enabler could support the student in visually representing the recipes (lines 3-6). After allowing the student to work on the enabler, Michelle revisited him to assess his progress:

Michelle: You haven't written anything... What did you find?
Student 1: (Desperately:) But I didn't understand what to do...

Michelle: When it says that one-fourth is apple juice, how many boxes will you color?
Student 1: This one. (Points to a quarter of the container.)
Michelle: Good! Use colors to differentiate the juice parts.
Student 1: I'll write an 'A' for Apple. Here it [i.e., Recipe B] says "one liter of apple juice and three liters of orange." I don't know what a liter is.
Michelle: Think of one box. A liter means 'one.' Do it. (L5, min 11:29-13:21)

Despite the use of the enabler, his progress in the task continued to lag (lines 1-2), confirming that his actual difficulties were not addressed and were becoming increasingly evident (lines 7-8). It is important to underscore that this was the first and only time that Michelle used enablers in her teaching, and she faced challenges that had not been resolved by the end of the program. Encountering challenges is a common experience when trying out new ideas for the first time, as it often involves a learning curve, and it might take time to fully understand how to use them effectively. With additional opportunities to employ this practice, there was potential for her to refine her approach and achieve noticeable progress.

Nurturing a classroom culture for enhanced cognitive activation and differentiation. Michelle consistently emphasized the pivotal role of classroom culture in either facilitating or impeding cognitive activation and differentiation, as well as their interplay. She maintained the belief that culture was primarily *shaped by students*. Particularly, when discussing barriers to promoting cognitive activation, Michelle acknowledged that "*student culture*" could act as a hindrance (VCS1, line 32). Expanding on this challenge, she noted that students in higher grades are *used to* gravitating toward routine tasks rather than engaging in complex thinking and seeking assistance from peers or teachers when faced with difficulties (PRI4, lines 101-102).

She aspired to develop practices that foster a learning culture encouraging cognitive activation and differentiation (RC4, question 4). Despite issues of classroom culture being touched upon in VCS9, Michelle remained preoccupied with this issue, aiming to acquire methods for transforming the attitudes of higher-grade students without a strong mathematics learning culture *from the outset of the school year* (EPI, lines 216-218). Perhaps, the ideas for promoting such a culture were provided somewhat late. This challenge persisted even after the PLD program, with Michelle emphasizing the school's role rather than the teacher's in shaping the classroom culture ("The mentality acquired over the years in the school requires significant effort to change." EPI, lines 150-151).

Addressing student indifference. Michelle faced challenges in motivating her students to participate in intellectually challenging tasks ("It is very difficult for them to get engaged with the task; if I don't work closely with them, they won't do anything.", POI1, lines 26-28). She believed that their disengagement and disinterest in tasks were primarily due to their indifference. In response to whether she would implement the practices discussed in VCS1 to engage students with varying ability levels in such tasks, she stated, "Yes, I will use some of the presented ideas that *'force'* all students to participate" (RC1, question 2, emphasis added). This expressed the persisting concern of student indifference that she carried into and throughout the PLD program. The following vignette vividly portrays her frustration as her students remained indifferent:

Interviewer: Do you think the less advanced students were cognitively activated?

Michelle: [Nodding negatively]

Interviewer: Why?

Michelle: Because they show complete indifference to everything... and I don't know how to handle it. It's not a matter of difficulty or level. It's that they're not interested in learning.

[...]

Michelle: What concerns me greatly are those three students [points to students' desks] who show complete indifference. These four students [points to other desks] do nothing when left alone... How can I keep them engaged in the lesson?

Interviewer: Do you think it's a matter of comprehension or indifference?

Michelle: It's not a matter of comprehension. For example, they were working on a simple calculation like three plus five plus seven, so I believe it's no longer a comprehension issue.

Interviewer: So, you are telling me that you need to learn some practices—

Michelle: [interrupting:] Yes, [I want to learn] practices on how to grab their interest. (POI2, lines 20-35; 52-60)

This challenge troubled her deeply, especially regarding certain students who consistently displayed this behavior (lines 5-7, 9-12), regardless of the complexity of the content (lines 14-16). Therefore, she would like to learn teaching practices capable of capturing their attention and inspiring active participation in lessons (lines 18-19).

She consistently revisited this issue (e.g., "There are certain students that show little interest not only in this particular lesson but also in all subjects." VCS2, lines 40-45). Student indifference was a significant challenge for her in addressing both axes:

TE1: Any other challenges [faced in your videotaped lesson]?

Michelle: *The students' indifference!*

TE1: Is it the same as handling different student levels?

Georgia: No, it's not the same.

Michelle: [agreeing with Georgia:] Yes, it's related to the classroom culture.

[…]

Michelle: My students don't care about anything! I have three students whom I tell, "Listen to what your classmate says and tell me what they said." *No response!*

TE2: Did you identify the reason [behind this unresponsiveness]?

Michelle: They're simply bored. These are kids who are neglected.

[…]

TE2: Could the indifference arise from asking them to do something too challenging?

Michelle: *No!* The other day, I gave them an easy addition with single-digit numbers. So, it wasn't something that required them to think.

[...]

TE2: If you use scenarios relevant to their daily lives, could they engage with the task?

Michelle: Do you know what they're into? *Doing graffiti and getting bored!* The principal utilized graffiti on a school wall! [laughter] It's a *very* unique class. [...] There are times that I let those students be in their own world and focus on the others, but I really feel remorse. (VCS3, lines 336-348; 426-434; 506-650, emphasis in the original)

This exchange highlighted that Michelle attributed student indifference to student-related factors like boredom or parental neglect (line 11). However, this "blaming game" appear to have prevented her from fully exploring how her teaching choices might contribute to their lack of interest and motivation.

This challenge remained unalleviated until the end of the PLD program, presenting an ongoing obstacle to her exploration of alternative methods aimed at sparking student motivation ("No matter how I approach these students, they don't even try." VCS9, lines 562; "If they lack the interest to learn, no matter what you do, they won't learn!" EPI, line 235). This persistent challenge left her grappling with a pivotal question: whether to redirect her efforts towards those students who exhibited enthusiasm for learning and potentially let go of those who displayed disinterest ("I am wondering whether it's wiser to let go of those who lack interest and not invest so much effort in trying to engage them. Should I direct more of my attention towards students who are motivated to learn?" POI4, lines 111-123). This highlights the complexity of handling this challenge.

7.5 Learning Level: Evolution of Michelle's Conceptualizations Around the Two Axes and Their Interplay

Following Kirkpatrick's model, this section explores how Michelle's understanding and perspectives evolved. The qualitative analysis illustrates her evolution from unfamiliarity with cognitive activation and a narrow view of differentiation towards a nuanced understanding of challenging tasks and the interplay between cognitive activation and differentiation. The next section further explores her learning development.

7.5.1 Michelle's initial conceptualizations

Unfamiliarity with the concept of cognitive activation. Michelle's initial conceptualizations around cognitive activation were underdeveloped; she could not identify any factors that could influence the set-up and implementation of challenging tasks. For example, she *refrained from using some highly challenging tasks in her teaching*, concerned about the students' difficulties and reactions, believing that they could not handle these tasks. She admitted that before participating in the PLD program, whenever she encountered a doing-mathematics task, she left it behind and moved on to tasks that were easier and more standardized ("There were times when I would say, 'Oh Lord, they won't be able to handle this investigation. Let's leave it behind", EPI, lines 65-66).[47]

Furthermore, Michelle joined the PLD program without recognizing that students' thinking, creativity, and flexibility can be constrained when the teacher establishes a highly structured environment and provides excessive guidance to students. For instance, during VCS1, the group observed a video clip of a lesson where the teacher excessively intervened to guide students during their autonomous work on a challenging task—ending up doing the thinking for his students. After watching it, a teacher educator raised concerns regarding the effectiveness of such teaching practices for students. Michelle began by asserting that the teacher's intervention is a form of support when students require assistance in developing their mathematical thinking (as in the sense of "caring"). She acknowledged that while some students can grasp the underlying mathematical concepts immediately, others may need guidance right from the start of autonomous work to effectively engage with the task. Notably, she failed to realize that the teacher in the video clip was overly guiding the students.

When pressed to provide further clarification on identifying the specific students who benefit from this kind of intervention, she rephrased her initial argument: she

[47] During an informal meeting with the teacher participants one year after the program's completion, Michelle without any prompting reaffirmed that *before* her participation in the program she was not using the demanding tasks or simply delegating them to more advanced students, who could voluntarily work on them at home.

asserted that it is helpful for all students, as they are not yet ready to *independently* tackle such advanced tasks.

Teacher Educator (TE): Which students benefit and which students do not benefit from this kind of support?
Michelle: I think it helps all students.
TE: Why?
Michelle: Because they're immature [laughing]. They haven't developed all the brain structures to be able to do everything on their own. For example, even a more capable student can sometimes work on something more advanced, or solve a task completely independently, while other times they may need a hint, and so on. (VCS1, lines 10-15).

Despite acknowledging that even the more advanced students sometimes need support, this incident highlights Michelle's difficulty in recognizing when a teacher's intervention might reduce the cognitive demand of a task.

In addition, during a subsequent discussion, she expressed the belief that lower elementary-grade students need to work in a highly guided environment; she referred to "working in a box" to imply working within defined boundaries or constraints (VCS1, lines 122-125). She could not acknowledge that providing explicit procedures and steps may lead to the routinization of problematic aspects of the task. When a teacher educator attempted to challenge her ideas, questioning whether this approach traps students and teaches them to expect everything to be readily provided, Michelle was problematized about these issues probably for the first time and responded hesitantly by saying "Perhaps..." (VCS1, lines 126-127).

Moreover, she demonstrated a lack of familiarity with cognitive activation practices. This became evident in VCS1, which focused on negative numbers. She expressed her intention to engage the students intellectually by utilizing various tasks from the student textbook that she hoped would pique their interest (PRI1, lines 8-10). However, by the end of L1, she realized that her attempts to promote cognitive engagement had failed, as most of the students did not actively participate in the task and instead played with the cubes she had provided (PRI1, lines 5-10).

In sum, Michelle had a limited understanding of cognitive activation, as evidenced by her avoidance of challenging tasks, difficulty in recognizing incidents of overly guiding students, and lack of knowledge regarding practices that promote cognitive activation. Also, she was not familiar with the concept of cognitive activation and how it contributes to student learning, as well as the factors that promote or hinder it. Her understanding of differentiation was also limited, as explained further below.

Differentiating mainly for students who face challenges. At the outset of the PLD program, Michelle mainly viewed differentiation as a means to support struggling students; her conceptualizations of it primarily revolved around differentiating downwards to meet their needs. An illustrative example of this tendency was evident in her rationale behind her approach to grouping students. In VCS2, she provided a detailed account of how she grouped students into *permanent* homogeneous groups based on their levels throughout the previous school year:

> Last year, I did this thing, but I didn't tell anyone: I divided the class into four groups, according to their level. First, I grouped my top students [i.e., in terms of academic performance] and told them, "I'll give you the task, I don't know what you'll do, find a way to solve it!" Then, I formed a slightly lower group [in performance], so I would go there occasionally or send one of the more advanced students to help. The third group included average to below average [in performance], and the last group had struggling students. This year, my colleague who teaches this class told me they took a math test, and the lowest-scoring student got 80 [out of 100]. Why? Because the previous year I sat with them, and they didn't move on until they completed the task and understood what we were doing. (VCS2, lines 288-297)

From this excerpt, two key ideas emerge regarding Michelle's approach to differentiating support among different groups of students. Firstly, she believed that more advanced students could solve a task without any guidance, and in fact, early finishers "enjoy" helping their struggling classmates (VCS2, line 302). Secondly, less advanced students require *constant* support from the teacher to complete a task, understand the mathematical concepts, and succeed in the subject.

Her initial pre- and post-lesson interviews corroborated that she was more concerned with how to differentiate her teaching to help less advanced students. In L1, focusing on negative numbers, she relied solely on the tasks of the textbook because her class was "generally at a low level, and students have many gaps in their knowledge" (VCS2, lines 36-37). The challenging task involved utilizing a vertical number line that encompassed both positive and negative numbers, resembling an elevator chute of an apartment building (see Figure 39). She provided tangible materials to certain students whom she considered to be struggling (i.e., two different-colored cubes were used, with each color representing floors above or below ground level, PRI1, lines 16-19). In contrast, she did not differentiate for more advanced students, arguing that such students were unlikely to "exist" in her classroom, and if they did "exist", they could work on extra tasks in the student textbook (PRI1, lines 23-24).

In conclusion, Michelle's initial conceptualization around differentiation was restricted. She believed that differentiation was promoted largely by grouping students homogeneously and by providing materials and resources to less advanced ones. In fact, she directed her attention towards groups of struggling students, assuming that more advanced ones do not require any form of differentiation and can solve the task independently.[48]

7.5.2 Evolution of Michelle's Conceptualizations

Progressing on cognitive activation: issues resolved and uncertainties. From VCS2 onward, Michelle started identifying practices implemented by other teacher-participants in their videotaped lessons that would assist her in maintaining the cognitive level of the task, such as *asking students to explain the instructions in their own words* (VCS2, line 98); and *asking students to individually write down their solutions*, rather than relying on listening to a classmate and copying their solution (VCS2, lines 183-185).

Also, in VCS2, for the first time, she identified an instance in which she was over-guiding the students during L1. She shared an excerpt from L1 (on negative numbers) where she had become overly directive with a student working on the challenging task of asking to find the distance between different floors of an apartment building. During the discussion around this excerpt, she admitted:

> I felt that I was *providing too much guidance to the students*. *I didn't let the student answer it on her own*. She [i.e., the student she was scaffolding in the shared clip] is the best student in the class, so I could have been less directive... She had said, "Miss, I'm having difficulties," and *I started telling her what to do*. (VCS2, lines 160-165, emphasis added)

In this excerpt, she discerned her moves and questions that decreased the level of challenge offered to students while circulating in the class during autonomous work. Notably, even with the most accomplished student, Michelle felt compelled to provide

[48] Michelle's conceptualization of each axis individually was not well-developed, which subsequently hindered her ability to conceptualize the interplay between them. Hence, the data included no evidence of her initial conceptualization regarding the interplay between cognitive activation and differentiation.

support by telling her what to do rather than affording her opportunities to explain her thinking, through appropriate questioning.

In VCS3, the PLD group watched a clip of a whole-class discussion where students shared their solutions to a fifth-grade algebraic task and the teacher's questions stole the students' thinking. Unlike the previous episode, during the discussion around this clip, Michelle was unable to identify the decrease in demands:

Teacher Educator (TE): What was the level of mathematical challenge the students worked on [in this clip]? Were they engaged with the task? Why?
Michelle: Yes! [laughs] They explained how they had worked on it themselves, then they found the eleven thousands [by implementing the rule], and eventually understood the pattern rule.
Nancy: [The teacher] asked how many rows of cubes each pattern term has, and how many constants there are.
Georgia: Also, they arrived at a mathematical formula.
Nancy: Moreover, they substituted the formula with numbers.
TE: Did you feel at any point that their thinking was stolen?
Pina: [Yes, when] he asked how many rows term two has, term three has, etc.
Michelle: But since it had already been discussed individually in the groups, shouldn't the whole-class discussion lead somewhere?
Georgia: He could avoid this kind of guidance, and if some children wanted clarification, he could have told them.
Nancy: Of course, we don't know what happened before [this clip].
TE: [There was] an individual discussion as the teacher circulated.
Nancy: So, they were given this opportunity [to think], and they reached a conclusion.
Suzanna: So, it wasn't stealing [of the thinking].
Michelle: Yes, that's what we're saying!
Georgia: At this point, more time was needed because they could have solved it on their own. (VCS3, lines 273-294)

Apart from Pina and to some extent Georgia, the other teacher participants failed to recognize the '*stealing of students' thinking'* during whole-class discussions. Michelle considered that if the cognitive level is maintained during autonomous work, it is not necessary for the whole-class discussion to promote high cognitive activation:

> If the teacher *has already had individual discussions* with the students and explained it to those who were struggling during autonomous work, then only the correct solution could be presented. [During the phase of whole-class discussion] we are simply *concluding* the lesson. (VCS3, lines 389-391, emphasis added)

Although there was a slight shift in her understanding of cognitive activation compared to her initial beliefs, she held the misconception that whole-class discussions were primarily aimed at summarizing key points, given that individual interactions with students had already taken place during student autonomous work. Furthermore, while she could recognize instances where there was a reduction in demands, she struggled to identify such instances when they occurred during whole-class discussions. Perhaps, for someone with limited experience and understanding of cognitive activation, like Michelle, the decrease in demands may have gone unnoticed, particularly during this lesson phase.

In VCS5 and VCS6, the PLD group revisited the topic of classifying tasks according to levels of cognitive demand based on the TAG (see Table 1). Michelle was now able to recognize various demanding task characteristics, such as allowing for multiple solutions (VCS5, line 406), leading to generalizations, providing explanations (VCS5, lines 464-466), or involving a complex representation (VCS6, line 320). However, she faced some challenges in distinguishing the cognitive level of tasks that involved both visual and symbolic representations. For example, she had doubts about the level of demands of a task asking students to describe the algorithm for the subtraction of three-digit whole numbers with base-ten blocks' representations (VCS5, lines 258-259). Even after the PLD group discussed and classified the task as level 3, Michelle was still unsure about its cognitive level ("If it involved only an algorithm, I would classify it as a memorization task." VCS5, lines 267). She had associated the presence of an algorithm with tasks of low level—as memorization of a set of steps, regardless of other task characteristics.

Furthermore, as time passed, she became adept at recognizing when her students were cognitively activated in her lessons. For instance, in POI3, when asked to provide examples from L3 (focusing on the commutative and associative properties of addition) in which her students were mathematically challenged, she explained:

> This group solved the problem without any calculations. They thought that since the tablet and the mobile phone have the same price, the camera should also have the same price as the laptop. Another instance was when a student observed [the commutative property of addition and asked], "Miss, does the same thing apply to multiplication?" This means that he moved a step farther. (POI3, lines 9-15)

She was now more attentive to students' reasoning compared to earlier videotaped lessons, in which she typically responded to pertinent questions with a general answer (e.g., "I believe the students understood the main objective", POI1, lines 4-5).

In summary, Michelle's conceptualizations around cognitive activation evolved, demonstrating growth in her understanding. She began identifying practices that upheld the cognitive level of tasks, such as contextualizing the task or asking students for explanations, or practices that reduce the demand, such as providing ready-made answers and hints to students before assessing their actual challenges and needs. However, she still faced challenges in recognizing instances of decreased cognitive demand during whole-class discussions. Also, while she improved in recognizing the characteristics of challenging tasks, she had difficulty distinguishing the cognitive level in tasks involving both visual and symbolic representations. Her evolving understanding and remaining uncertainties highlight the complexity of conceptualizing cognitive activation and the need for continued PLD in this area.

Persistent differentiation beliefs and practices. Michelle delved deeper into the axis of cognitive activation and the interplay of the two axes, setting aside issues exclusively related to differentiation. She was consistent in how she conceptualized differentiation throughout the PLD program. She considered that differentiation involved grouping students based on their performance levels (PRI2, lines 36-37; VCS3, lines 169-174; PRI3, line 16; POI4, lines 72-74), to allocate more time to assist students facing difficulties: "Less advanced students sit together so that it is easier for me to assist them during the lesson. This [grouping] makes it easier for me to be in the same location consistently" (PRI4, lines 31-33). Also, she thought that providing materials to less advanced students (VCS3, lines 178-180; PRI4, line 43) facilitates differentiating downward, which was her main concern. Michelle’s view of differentiation is summarized in the following statement made before teaching L3:

> I will provide materials [to less advanced students] and I hope that they will engage with them. I will spend considerable time with this group of students! The goal is for the more advanced students to work independently so that I can sit with the students who are facing challenges and assist them. (PRI3, lines 40-42)

In this excerpt, a misconception about differentiation is identified. She assumes that by providing materials solely to less advanced students and spending considerable time with them while expecting more advanced students to work independently, students’ learning will automatically be enhanced. However, this approach oversimplifies the concept of differentiation, as it overlooks that differentiation involves providing appropriate support to *all* groups of students. Thus, Michelle did not seem to have made much progress in her understanding of differentiation.

Undergoing a significant shift regarding her conceptualizations of the interplay. Her awareness and understanding of the interplay of cognitive activation and differentiation emerged and evolved over time. She demonstrated increased attentiveness to analyzing the dynamics of teacher-student interactions. Her progress was evident across all three phases of teaching. In particular, she was actively involved in the discussion that emerged around some narratives included in the EDUCATE materials, providing insights into the teacher-student interactions. In task launching, she became more adept at recognizing when the teacher's questions elicit all students' comprehension of the task's instructions or when the teacher tends to over-explain the instructions, thereby reducing the demands (VCS4, lines 130-162).

Regarding student autonomous work, Michelle suggested asking for explanations when the students have completed the task ("Can you explain what you did?" VCS4, line 219; "How did you work on the task?" line 238; "How does this idea help you work on the task?" line 247). She also proposed asking students to rephrase the task instructions when they are stuck (VS4, line 250). She argued that these questions help the teacher to ensure that more advanced students truly understand what they have been doing (VCS4, line 276). In addition, she doubted her ability to help these students by posing different types of questions, insisting that they would continue to struggle regardless of her efforts (VCS4, lines 176-197).

In the whole-class discussion phase, Michelle supported that an incorrect student solution should be presented and discussed in the plenary to problematize all students, especially less advanced ones, and help them reflect on its incorrectness. She justified the need to engage those students in a whole-class discussion, pointing out that if not done "The students who have made the mistake will not go through the process of reflecting on why it is wrong. They will simply think, 'I copied the correct one, let's move on."(VCS3, lines 234-237). Furthermore, she suggested asking students for explanations and paraphrasing their peers' answers (VCS4, lines 112-113).

Notably, there seems to be a contrast between the phases of autonomous work and whole-class interactions. In fact, Michelle had many issues with struggling students for the phase of student autonomous work but not for whole-class discussion. It is speculated that this difference may be attributed to the fact that the difficulties and diverse needs of different student groups are amplified during autonomous work, which consequently poses greater challenges for a teacher. Yet, in her EPI, Michelle also admitted experiencing challenges in the phase of whole-class discussion (e.g., less advanced students struggle to follow the discussion, EPI, lines 89-95).

The most significant shift in Michelle's conceptualization regarding the interplay of the two axes was observed from VCS5 onward. In VCS5, she emphasized the importance of practices that can engage all students, *including* those who face difficulties, in challenging tasks. This marked a departure from her previous belief that there was little a teacher could do to support the learning of less advanced students. In her EPI, Michelle highlighted multiple times the use of enablers and extenders, along with the incorporation of effective and diverse questioning, as the most invaluable and tangible tools provided by the program to enhance her teaching practice (e.g., EPI, lines 123-125).

Specifically, during VCS5, Michelle emphasized the significance of using open-ended questions during the launch of tasks. Rather than providing direct hints, she suggested asking students to explain their understanding of the task instructions (VCS5, lines 39-41). She also recognized the importance of dedicating time to clarify the key terms of the task instructions, ensuring students' comprehension of the task (VCS5, lines 48-52). Additionally, she highlighted the value of posing questions that assess students' understanding and help them identify their errors, instead of immediately resolving their doubts during autonomous work (VCS5, lines 134-135). This shift in perception showcased her growing awareness of the potential of questioning in supporting all students throughout their learning process.

Regarding the enablers and extenders, during the PLD group's discussion on a challenging task from Pina's lesson, Michelle actively contributed by offering several suggestions related to enablers and extenders in VCS5. The task involved students discovering which numbers would activate each bulb of a computer, with the blue bulb representing multiples of five and the green bulb representing multiples of two (see Figure 25):

Michelle: [As an extender] the students can determine which numbers will not activate either bulb.
TE1: How would that engage them further?
Michelle: The extender will stimulate their thinking by asking them to employ reverse thinking. [Additional extenders are shared by other teacher participants.]
TE1: [Turning the discussion to enablers:] How can we support students who are facing difficulties?
Michelle: The teacher can provide them with a table with all the numbers up to a hundred.

Georgia: Indeed, offering a table with all the numbers can be a helpful resource.

[…]

TE1: Is there a risk of the teacher taking over the students' thinking with these ideas? Remember, the goal is to assist the students without depriving them of their thinking process.

Michelle: [No, because] for a struggling student, this table can be a key tool [for solving the task]. (VCS5, lines 552-555; 610-619)

In this excerpt, Michelle exhibits a solid understanding of the concept of enablers and extenders. She proposed providing a task that involves reverse thinking as an extender for early finishers (lines 3-4) and additional resources as an enabler for students facing difficulties (lines 11-12). Contrary to her earlier conceptualizations about supporting less advanced students, what stands out in her suggestions is her acknowledgment of the importance of supporting those students without undermining their thinking process by providing appropriate enablers (lines 17-18).

In the subsequent sessions, she delved into additional aspects of developing and using enablers, specifically focusing on *how* and *when* they should be utilized. For instance, in VCS6, she proposed incorporating an *intermediate step* within a multistep task "to address abrupt transitions between different parts of the task" as an enabler (VCS6, lines 472-476). This insightful suggestion shows that she developed a better understanding of how enablers can facilitate smoother transitions between different parts of the task, particularly when there is a gap in difficulty levels.

In VCS7, Michelle demonstrated her understanding of when to provide an enabler, as evidenced by her response to the questions of Ariana, another teacher participant:

Ariana: I didn't give them the other enabler right from the beginning... Should I have done that?

TE: Just because you prepared it, it doesn't mean you have to provide it; they [i.e., the students] might not need it... It should arise as a necessity.

Ariana: I don't know who will need it.

Michelle: Didn't you wait for some time, let the students work on the task, and then give the enabler to them?

Ariana: I was waiting…

Michelle: You should have given them a chance to think first. (VCS7, lines 279-286)

This vignette shows that Michelle realized that an enabler should be provided to students when they face challenges that they are unable to overcome independently.

Rather than offering the enabler right from the beginning, she recognized the value of allowing students to initially engage with the task individually (lines 6-7).

In conclusion, Michelle's understanding, and awareness of the interplay between both axes evolved noticeably over time. She analyzed more teacher-student interactions, recognizing the role of questioning. Interestingly, her conceptualization shifted notably from VCS5 onward, after the tools of enablers and extenders were introduced. The introduction of these tools marked an “aha moment” for Michelle, as she considered them a viable tool for engaging all students—including those facing difficulties—in challenging tasks without overshadowing their thinking process, offering astute suggestions and examples during the PLD discussions.

7.5.3 Michelle’s Final Conceptualizations

From unfamiliarity to the embracement of challenging tasks. One of Michelle’s most significant gains from EDUCATE was her engagement in using challenging tasks in her lessons. When asked about how the PLD program helped her regarding cognitive activation, she shared the following account:

Michelle: I felt like I gained [cognitive activation] practices to make all the students interested in the lesson. Now, there won't be any students saying, "I won't do it," nor will I say, "Oh, this is too difficult for my students, let's leave it behind."

Interviewer: So, there were times when you used to do something differently?

Michelle: Yes. [Before participating in EDUCATE] I would say, "Oh Lord! Students won't manage this investigation task, let's leave it behind." But now, with the things we discussed and the practices we learned, I felt safer even to say, "No, I won't leave these tasks behind, we will work on them." (EPI, lines 57-68)

In this exchange, Michelle vividly discussed her confidence in embracing highly challenging tasks instead of avoiding them (lines 1-4). She no longer shied away from challenging tasks and *felt safer* working on them with her students (lines 8-10). As a result of her persistent use of such tasks, she noticed that “the percentage of students who understand the new mathematical concepts surpasses the typical percentage previously observed” (EPI, lines 250-252). This observation highlighted that Michelle focused more on students’ *conceptual understanding* rather than on their procedural fluency.

Beyond implementing challenging tasks, she explicitly discussed the ideas and practices she acquired through the PLD program. These included avoiding the use of hints or leading questions that provide ready-made answers (VCS9, lines 77-78); and presenting problems within scenarios that closely resemble real-life situations and align with her students' interests (EPI, lines 335-336). Thus, the PLD program appears to have equipped Michelle with certain practices to promote cognitive activation.

All in all, before she participated in EDUCATE, Michelle was not familiar with the idea of cognitive activation and its potential impact on student learning. After the completion of the PLD program, she experienced a notable shift in her perspectives, recognizing the value of cognitive activation in enhancing student conceptual understanding. Conversely, she did not experience similar levels of conceptual change regarding differentiation, as explained next.

Persistent beliefs: Differentiation mainly for less advanced students. When it comes to differentiation, Michelle's conceptualizations remained relatively stable. In particular, she continued believing in grouping students based on their performance levels and providing constant scaffolds to less advanced groups. When discussing her approach to handling different student groups in L4, she explained:

> High-performing students had already completed some tasks and found it much easier to grasp the concepts. So, they progressed much faster compared to less advanced students who needed me to sit with them and have one-on-one sessions, like a private lesson, to reach a certain level of understanding. (POI4, lines 13-16)

Simply relying on more advanced students to work independently and dedicating all attention to less advanced students did not fully speak to the idea of differentiation. While it is important to provide additional support to students who may be struggling, true differentiation involves tailoring teaching to meet the diverse needs of *diverse* students.

Classroom organization could have been one of her barriers to differentiation:

Michelle: If I had more space [in the classroom], I would set up additional learning centers.

TE1: How would that help?

Michelle: Early finishers would have their learning center, while the students who struggle would have learning aids. They wouldn't have to *constantly wait for their turn* to work with the teacher. (VCS9, lines 422-430, emphasis added)

Once more, Michelle admitted that her students were heavily reliant on her intervention. She focused on creating multiple learning centers in the classroom, which was hindered by the lack of physical space in the classroom (lines 1-2). The intention behind this approach was to promote independent learning (lines 4-6). The assumption was that early finishers would have their own designated space for extenders, while struggling students would have access to learning aids, reducing their dependence on constant support. However, it is important to note that while the concept of learning centers could be beneficial, Michelle's implicit assumption that they could replace other important elements of differentiation, such as ongoing monitoring of student progress and flexible grouping, was suggestive of her limited understanding of differentiation.

In conclusion, Michelle's unwavering beliefs regarding attending to less advanced students and leaving more advanced students on their own limited her understanding of differentiation. At the end of the PLD program, she began to recognize the importance of attending to both groups. Although she still prioritized the former group, her shift showed a deeper comprehension of the interplay of both axes, as elaborated upon below.

Shifting perspectives on the interplay of both axes: making use of challenging tasks and expanding her teaching toolkit. Towards the end of the program, Michelle critically examined her prior ideas on how to maintain the challenge for *all* students and debunked many of them. Reflecting on her PLD experience, she shared a comprehensive account, drawing comparisons between her previous and newfound insights:

> I have gained the courage to try new practices and have developed greater self-assurance in my abilities. I have come to realize that it is acceptable to deviate from the task sequence outlined in the student textbook. Additionally, I have learned that not all students need to work on the same part of a task simultaneously. Throughout the program, I've delved into various practices to provide the right level of challenge for struggling students and to give that extra push to students who seek more engagement, saying "Miss, I'm done, I want something else." It is not always necessary to prepare separate tasks for these students. Instead, I've discovered that by asking early finishers to solve the task with additional methods, I can effectively maintain the level of challenge that suits their capabilities. (EPI, lines 29-38)

This excerpt is very telling of Michelle's conceptual gains from the PLD program for promoting the interplay for all students. She gained a sense of confidence in experimenting with innovative ideas. She also recognized that occasionally straying from the rigid structure of the student textbook and allowing students to asynchronously work on tiered tasks is a valuable approach (EPI, line 80). Moreover, she learned that more advanced students do not always need extra tasks; instead, she differentiated the main task with extended components for them. This shift in perspective showcases her growth in promoting an inclusive learning environment.

Moreover, she eloquently described how the implementation of newly acquired practices had a positive impact not only on her self-confidence but also on her *students' self-assurance*. She noticed a significant increase in their willingness to take risks and try new things. They became more proactive in expressing their ideas, participating in class discussions, and seeking assistance when needed (VCS9, lines 224-226). This transformation was facilitated by the *use of enablers and extenders*, which played a crucial role in supporting multiple student needs (VCS9, lines 228-230). By employing enablers, she argued that she no longer had to be physically present everywhere in the classroom. However, she persistently prioritized her focus on supporting less advanced students over more advanced students, as already explained in the previous section.

Apart from enablers and extenders, Michelle learned about how *questioning* can enhance the interplay of cognitive activation and differentiation. Particularly, she emphasized the importance of not doing the thinking for the students through the questions posed. As a result, she had become highly attentive to and cautious of the formulation of her questions (EPI, lines 85-87):

> I differentiate my questions based on the needs of each student. I understand that a less advanced student may require additional support in foundational concepts. For instance, if multiplication is a prerequisite knowledge for a task, I will help with my questions accordingly before guiding them through the task. Yet, I would not do that for a more advanced student. (EPI, lines 134-138)

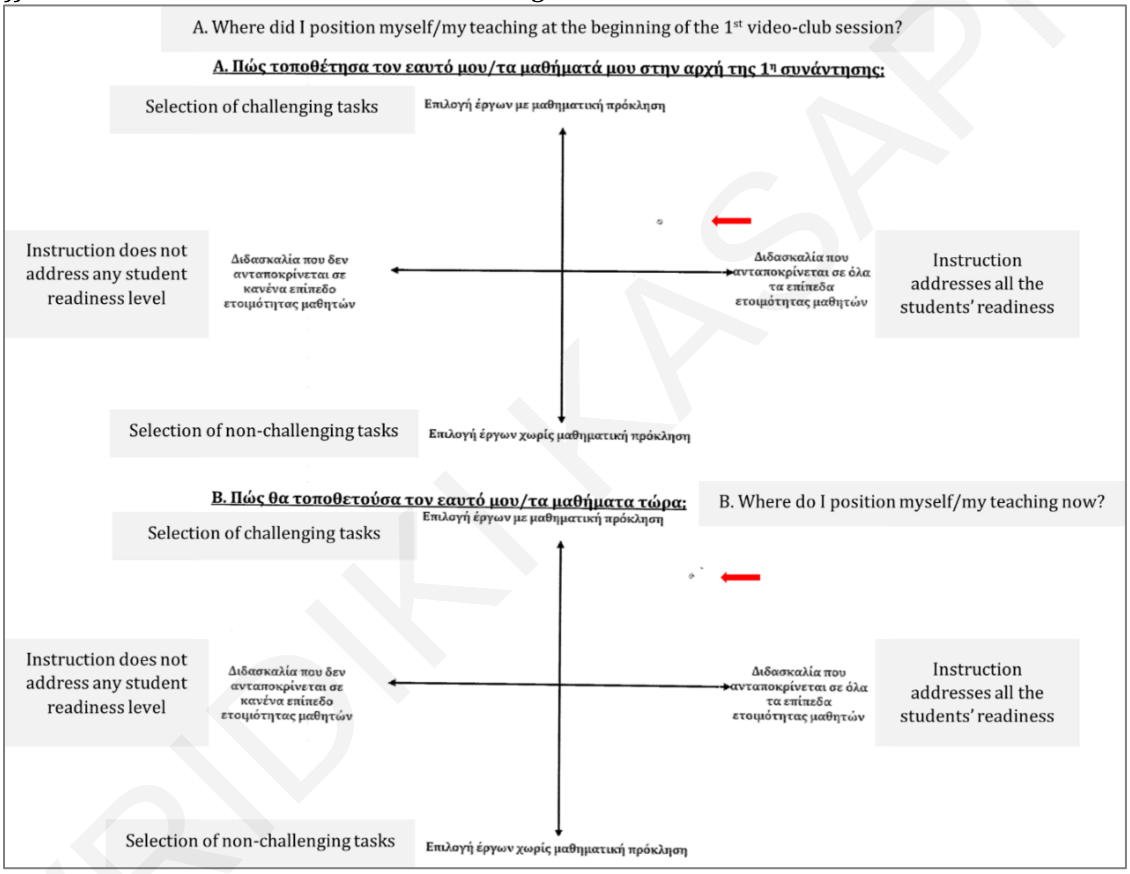
Remarkably, in this snippet, Michelle supports that tailoring her questioning to student needs supports students who require additional support. However, with her over-emphasis on using questioning to scaffold less advanced students, there is a risk of over-assisting them, potentially hindering their development of independent thinking.

After all her experiences in the EDUCATE project, Michelle supported that maintaining a high cognitive level for all students in the classroom is feasible ("For example, helping struggling students solve a task promotes their cognitive activation. Without this support, they may give up. This does not hinder the progress of more advanced students since they do not require the same level of scaffolding", EPI, lines

177-181). This excerpt highlighted a consistent pattern in her conceptualizations—she was much more concerned with less advanced students rather than with more advanced ones. The program made her more aware of the needs of more advanced students, but still, her main concern was with the former group of students.

Figure 43

Four-Quadrant Diagram Depicting Michelle's Positioning of Her Teaching (A) Before and (B) After Participation in the Program Concerning Cognitive Activation and Differentiation, With Its Translation in English



Finally, during the first and final sessions (VCS1 and VCS9, respectively), teachers were asked to indicate where they positioned themselves and their teaching in the two-dimensional space shown in Figure 43, with the vertical axis representing the extent to which they selected challenging tasks for use in their teaching and the horizontal axis corresponding to the degree to which their teaching addressed all the students' different readiness levels. In VCS9, Michelle situated her teaching in a different spot compared to VCS1, demonstrating primarily an upward shift in selecting challenging tasks. Asked to jot down things that helped her make this shift, she provided the following account:

> The difference exists because I felt more confident in selecting tasks that include mathematical challenges. This was further supported by suggestions on how to assist less advanced students in acquiring knowledge while simultaneously enabling more advanced students to fully utilize their abilities.

One notable aspect in this excerpt was her transformation in the conceptualization of cognitive activation because she now found a way to use more challenging tasks even with less advanced students. Also, although she prioritized less advanced students, she progressed by recognizing the importance of catering to the needs of both student groups. However, remarkedly, her positioning did not demonstrate much progress in terms of differentiation, as only a minimal shift to the right was observed.

7.6 Learning and Behavior Levels: A Comparison Between Michelle's Conceptualization and Practice

The trajectory of Michelle's PLD, as elucidated through her participation in EDUCATE, presented a multifaceted panorama of the complex relationship between her conceptual understanding of cognitive activation, differentiation, and their interplay on the one hand, and their tangible manifestation in her classroom, on the other hand.

Concerning her practice, although she utilized certain tools and showed improvement in their use during the program, her overall progression did not manifest a definitive trajectory of change. On a positive note, she was a caring teacher, earnestly dedicated to facilitating her students' learning journey. However, her approach to caregiving revealed an imbalance between cognitive activation and differentiation and her low tolerance for student struggle and errors. She overly directed students, aiming to remove learning barriers; despite being well-intentioned, this behavior inadvertently limited opportunities for productive struggle.

Regarding her conceptualizations, Michelle demonstrated growth in identifying practices that influence cognitive demand and developed a better understanding of teacher-student interactions. However, her conceptualizations around differentiation seem to have remained consistent, with her primary focus on supporting less advanced students. The introduction of enablers and extenders marked a shift in her understanding of the interplay between both axes yet achieving a balanced approach to meet all students' diverse needs represented an area that required more work on her part.

A comparative analysis of the insights derived from her evolving conceptualizations and the critical examination of her teaching practice revealed a lack of alignment between the two, with "consistent" inconsistencies surfacing. A palpable dissonance was revealed between her conceptual change and her practical application on the axis of cognitive activation. On the one hand, her journey in understanding cognitive activation reflected a transformative process, from ignorance to understanding. Initially, her conceptualization was characterized by misconceptions and a reluctance to incorporate challenging tasks, stemming from a fear of student struggles. However, as her understanding deepened, she began to embrace these tasks, recognizing their potential to enhance student engagement and foster conceptual understanding. In practice, this transformation was not linear. There were moments of progress, where she attempted to implement challenging tasks and restrained excessive intervention. Yet, inconsistency was a prevalent theme, with instances of reverting to more directive teaching methods because of her low tolerance to student struggle. This inconsistency emphasized the complex nature of translating conceptual understanding into classroom practice, particularly when navigating the delicate balance of providing support without diminishing the demands. While her understanding of cognitive activation had matured, translating this understanding into consistent classroom practice remained a challenge. This highlights a common phenomenon in teacher professional development where mindset shifts do not necessarily or directly result in changes in practice. This is because altering old teaching habits or practices that have been deeply entrenched over the years might take a lot of effort as well as a lot of time to materialize.

Her conceptualization of differentiation remained relatively stable throughout her development, primarily focusing on supporting less advanced students. Her understanding of differentiation was rooted in this supportive role, viewing it as a means to assist struggling learners. In practice, this translated to a predominant use of homogeneous grouping and a provision of tangible materials to less advanced students, even at the culmination of the program. While this approach might have stemmed from a tendency to care for her students, it inadvertently marginalized more advanced learners, as they were expected to work independently or assist their peers, with limited differentiation provided to cater to their needs. In this aspect, there was an alignment between her conceptualization and practice. However, a closer examination of her teaching suggested that she demonstrated a path of change leading to high-level implementation of specific differentiation practices across her lessons, such as asynchronous work and encouraging diverse forms of expression—something that was not observed in the other two axes of her practice. This suggested that despite the stability in her conceptualization of differentiation, she was exhibiting an improved implementation of differentiation in her teaching practice. The question that then arises is: To what extent did Michelle implement certain tools and practices acquired during the program, even though they were not theoretically conceptualized or underpinned, and—more critically, would she continue utilizing them in her practice moving forward?

An incongruence was also evident between her conceptualization and practice regarding the interplay of cognitive activation and differentiation, which represented a critical juncture in her PLD. The introduction of tangible tools, namely enablers and extenders, in VCS5, marked a significant shift in her understanding of the interplay between cognitive activation and differentiation, which evolved significantly by the end of the PLD program. She began to recognize the importance of balancing support and challenge, being mindful of her questioning techniques, and utilizing enablers and extenders to cater to diverse student needs. In practice, although she showed moments of progress, by improving in using extenders and posing questions, her overall practice did not demonstrate a clear direction of change. Her well-intentioned "caring" approach sometimes led to over-guiding, diminishing student autonomy and cognitive activation. Balancing support across different student groups remained a complex endeavor, and there were moments when her support for less advanced students overshadowed the needs of more advanced learners.

In synthesizing the comparison between Michelle's conceptualization and practice, it becomes evident that her journey through the PLD program has been marked by growth, challenges, and moments of inconsistency. The discrepancy between her conceptual growth and practice suggested that while PLD programs can catalyze shifts in teachers' mindset, ensuring that these shifts translate into consistent classroom practice requires ongoing support, tangible tools, reflection, and perhaps a re-evaluation and transformation of the teachers' ingrained teaching habits and beliefs about teaching and learning mathematics. It is quite possible that she had deeper, ingrained beliefs that posed challenges to her ability to progress and effectively balance cognitive activation with differentiation. These beliefs could stem from her previous experiences, educational background, or long-held views on teaching and learning.

Such deep-seated beliefs can significantly influence a teacher's approach to teaching, potentially creating barriers to adopting new strategies or shifting perspectives. For instance, if she believed that her primary role as a teacher was to provide direct instruction and prevent student struggle at all costs, this could lead her to intervene prematurely during lessons, hindering cognitive activation. Similarly, if she held the belief that differentiation primarily means providing additional support to less advanced students, this could result in a skewed application of differentiation practices, neglecting the needs of more advanced learners. Addressing these deeper beliefs would require a reflective and transformative PLD experience, where Michelle could critically examine her own beliefs, understand how they impact her teaching practices, and work towards aligning her beliefs with effective teaching strategies.

Additionally, it is crucial to acknowledge that the relationship between conceptualization and practice is bidirectional; Michelle's day-to-day experiences and practices in the classroom might also have played a pivotal role in shaping or refining her conceptual understanding. For instance, a successful implementation of a challenging task may have reinforced her belief in the value of cognitive activation, while observing her students' diverse responses to differentiation strategies and the persistent challenges of less advanced students might have led her to remain primarily focused on supporting less advanced students and lacked major adjustment in her conceptualization of differentiation. In this way, her practice might have become a dynamic component of her PDL, creating a feedback loop that enhanced her conceptual understanding and often guided her toward more informed and effective teaching strategies. Acknowledging the limited scope of this exploration and that changes in teaching practice and learning outcomes do not manifest instantaneously, it is essential to consider the long-term commitment required for educational innovation. Perhaps with a commitment to ongoing experimentation, Michelle could introduce more substantial changes to her teaching practice in the ensuing years. This consideration highlights the ongoing nature of educational transformation and the potential for continued development beyond the current observation period.

In sum, Michelle emerged as a caring teacher, who was trying to do the best for her students and help them. Her learning path through the PLD program highlighted improvements in her understanding of cognitive activation, yet her approach to differentiation required further refinement to adequately address the diverse needs of her entire classroom. Her teaching journey revealed a duality: on one hand, she demonstrated commendable growth in certain practices. On the other, her genuine desire to support her students occasionally overshadowed the need for cognitive challenge, leading to over-guidance even by spoon-feeding them with answers, thus blurring the lines between providing supportive guidance and diminishing the mathematical challenge. Hence, her overall performance did not show a certain improvement path. Drawing inspiration from Nel Noddings' relational approach to education, Michelle's portrayal underscores the complex balance between offering support and allowing productive struggle. While she grappled with challenges in integrating cognitive activation with differentiation, her commitment to continuous improvement and genuine care for her students remained evident. All in all, Michelle's portrait painted a picture of a teacher in evolution, navigating the complexities of ambitious teaching.

CHAPTER 8. DISCUSSION

> The final chapter consolidates key findings from an evaluation of a PLD program, focusing on cognitive activation, differentiation, and their interplay. It advocates for a multifaceted approach to assessing PLD effectiveness, guided by Kirkpatrick's model, and outlines the program's strengths and weaknesses to inform future PLD designs. Additionally, the study reveals diverse pathways of teacher change and learning, highlighting the differential impact of PLD on teacher development. Theoretically, it emphasizes the complexity of teacher change and underscores the significance of integrating cognitive activation with differentiation. Methodologically, it calls for thorough evaluations of PLD programs, it validates the use of video clubs for fostering ambitious teaching practices among elementary teachers, and it points to the value of employing both mean and maximum performance metrics for a more nuanced assessment of teacher development. Practically, the study demonstrates the viability of blending cognitive activation and differentiation in teaching, identifies critical PLD features facilitating this blend, and stresses the need for precise assessments of teachers' needs for more effective PLD program customization. The chapter concludes by acknowledging the study's limitations and suggesting directions for future research to further enrich the field of teacher PLD.

8.1 Introduction

Drawing on the richness of the existing literature on cognitive activation (e.g., Hsu & Yao, 2023) and differentiation (e.g., Tomlinson & Imbeau, 2023), this dissertation endeavored to work at the nexus of these two critical strands within the context of a PLD program aimed at fostering ambitious teaching (Charalambous et al., 2023b). While each axis—cognitive activation (e.g., Stein et al., 2007) and differentiation (e.g., Tomlinson, 2014)—has been explored extensively in isolation, their synergistic potential remained largely at a nascent level. Recent investigations (e.g., Charalambous et al. 2022, 2023a; Delaney & Gurhy, 2019; Mellroth et al. 2021; Psycharis et al., 2019) have begun to illuminate this composite area, yet discourse on their integration within PLD programs, particularly in the realm of mathematics education—a field that has increasingly recognized the value of ambitious teaching practices (e.g., Blazar, 2015)—is only emerging. This study sought to contribute to this evolving dialogue by holistically evaluating the effectiveness of a PLD program focusing on cognitive activation, differentiation, and their interplay.

Synthesizing the insights gleaned on the characteristics of PLD programs in Chapter 2, it was concluded that such programs represent a promising avenue for developing teachers' competencies in *either* cognitive activation (e.g., Boston & Wilhelm, 2017; Smith & Stein 2023) *or* differentiation (e.g., Gheyssens et al., 2020; Valiandes & Neophytou, 2018). With respect to cognitive activation, it was concluded that despite the successes of earlier programs, research still underscores the need for dynamic PLD aimed at supporting teachers in attending to student thinking (Tekkumru-Kisa et al., 2020), as well as identifying and addressing challenges teachers face in implementing challenging tasks (Superfine & Superfine, 2023). With regards to differentiation, it was pointed out that, while teachers may benefit from participating in PLD focused on differentiation, such programs also exhibit certain shortcomings. These include limited direct observation of teaching (Prast et al., 2018), insufficient hands-on practice (Slade et al., 2006), restricted opportunities for teachers to develop a conceptual understanding of differentiation (Santamaria, 2009), and inadequate monitoring of teacher learning and teaching practice over time (Santana, 2020). These challenges, coupled with the variability in differentiation quality across educational contexts, signal a pressing need for more PLD programs focusing on understanding teachers' challenges and supporting them in implementing differentiation (Maulana et al., 2023). Therefore, it was concluded that there remains fertile ground for further research and teacher development in *either* axis through PLD.

More critically, the review of the existing literature suggested that the effectiveness of PLD programs across educational levels in nurturing the *interplay* of cognitive activation and differentiation—or ambitious mathematics teaching—remains a topic of debate. Studies on PLD effectiveness in ambitious teaching have predominantly focused on outcomes related to teachers (such as their learning, professional noticing, mathematical knowledge for teaching, and teaching practice). Yet, the body of evidence presents a spectrum of methodological approaches and outcomes, from transformative shifts to minimal or null changes in teacher learning and teaching practices. While qualitative studies shed light on collective progress in ambitious teaching practices among PLD participants (e.g., Fauskanger & Bjuland, 2019; Gibbons & Okun, 2023; Jakopovic, 2021), they also reveal a critical oversight in capturing the individualized experiences and challenges faced by teachers (Wæge & Fauskanger, 2021, 2023). This gap highlights a broader issue within existing research: a tendency to view teacher learning and improvement in teaching practice as a uniform process. This scenario is compounded by a notable absence of direct observation and analysis of ambitious teaching practices (Gibbons et al., 2017). Indeed, both qualitative and mixed methods studies revealed notable variations and challenges in how different teachers implement new ambitious teaching practices in their classrooms (Anthony et al., 2018; Charalambous et al., 2023a; Leong et al., 2021; Witherspoon et al., 2021), underscoring the complexities involved in enacting ambitious teaching. Furthermore, quantitative research revealed significant variability in PLD effectiveness in teaching practice and learning (Dash et al., 2012; Hill et al., 2018; Jacob et al., 2017; Kraft & Hill, 2020; Lindvall et al., 2022, 2023), suggesting that PLD may not impact all teachers uniformly (Shumway et al., 2020).

The recent studies by Lindvall's group (2022; 2023) scrutinized the effectiveness of a yearlong PLD program on ambitious teaching demonstrating the importance of considering not only the *end* results in teaching practice of the program but also the timing and progression of changes in teaching practice *throughout*. In addition to teaching practice, the work of Horn and Garner (2022) underscored the necessity to broaden PLD research focus on *teacher learning* beyond merely theorizing it as an outcome of PLD. Instead, they advocate for a nuanced understanding of teacher learning as dynamic processes intertwined with teachers' prior experiences, teaching practice, and social histories, as well as the context of PLD and the socio-cultural milieu.

The juxtaposition of these findings suggests a shift towards mixed-methods research that not only investigates the *collective* effectiveness of PLD programs but also delves into a more granular exploration of teachers' *individual* experiences both throughout and at the end of the PLD program, examining the evolution and the dynamic interplay between their conceptualizations, practices, and challenges they face. This exploration is critical for designing PLD programs that effectively integrate cognitive activation and differentiation, responsive to the diversity of teacher needs.

In this context, four research questions were posed herein, aligning with the Kirkpatrick four-level framework (cf. Kirkpatrick & Kirkpatrick, 2007)—used as a lens to probe into the effectiveness of the PLD program implemented in this study. In particular, the study aimed to explore what teachers gain from participating in a PLD program focusing on cognitive activation, differentiation, and their interplay. Firstly, it analyzed the end *Results*, questioning the frequency and the mean and maximum performance of teachers' teaching practices in their concluding lessons. Secondly, the investigation turned to teachers' *Behavior*, examining how their teaching behavior evolved, the practices with which they most frequently experimented, the changes introduced in their practice, and the challenges encountered as they navigated new teaching terrains throughout the PLD program. Thirdly, regarding their *Learning*, the study sought to understand how teachers' learning progressed over time, how they (re)conceptualized cognitive activation, differentiation, and their interplay, and how these (re)conceptualizations compare with their practices, across the PLD. Lastly, the study examined teachers' *Reactions*, aiming to capture their perceptions and feedback on the program. The mapping of the research questions onto the four-tiered structure of Kirkpatrick's model lays the groundwork for an exploration that underscores the importance of delving into all four levels. This multi-level perspective catalyzes the discussion of the main findings in the next section.

Instead of discussing the results of each research question in isolation, this chapter identified and focused on four topics, collectively addressing the previously outlined research questions. The first topic pertains to studying the phenomenon of PLD effectiveness through multiple lenses, thus discussing the power of collectively examining all four levels of Kirkpartick's model. Following this, the second topic illuminates the strengths and limitations of the PLD program under consideration, enriching the literature on the features of effective PLD, which are crucial for informing future designs and implementations of PLD. The third topic ventures into the differential effectiveness of PLD across different teacher learning and teaching pathways, underscoring the imperative for more adaptive interventions. The concluding topic illuminates Kirkpatrick's model affordances and limitations in evaluating teacher PLD. After presenting and discussing these topics, the chapter outlines the theoretical, methodological, and practical contributions and implications of the study. It concludes by outlining the methodological limitations of the study and proposing directions for further research.

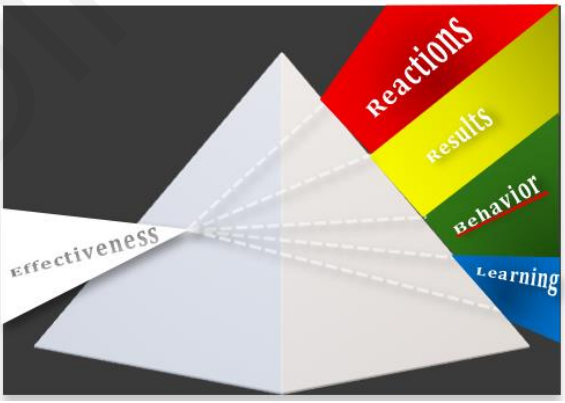
8.2 Unpacking the Concept of Effectiveness in Teacher PLD Through a Multi-Focal Lens

Kennedy (2019) underscored the complexity of evaluating the effectiveness of teacher PLD programs, suggesting that traditional evaluation metrics and methods may not fully capture the nuanced changes for teachers. She wondered how researchers would know whether the teachers had learned or improved enough. Several scholars have questioned the methodological limitations through which the research community evaluates the effectiveness of a PLD program, acknowledging the possibility that teachers may learn and improve in ways that are not visible or accounted for in PLD evaluations (e.g., Goldsmith et al., 2014; Lindvall et al., 2022). This perspective compelled a *reevaluation* of effectiveness in teacher PLD, advocating for a more holistic approach that recognizes the multifaceted nature of teacher development.

In response to this research call, the current evaluation of the focal program's effectiveness for teachers illustrated how a multi-dimensional analysis can provide a comprehensive understanding of the program's effectiveness, evocative of a "prism." Analogous to the way a prism disperses light to unveil its full-color spectrum, the multifaceted analytical approach of this study dissects the variegated dimensions of program effectiveness (see Figure 44). Specifically, through lenses that individually assess teachers' *reactions* to the program, their teaching *results* at the program's culmination, their *behavioral* evolution as well as their *learning* evolution across the program's timeline, the study decomposes the complex notion of "effectiveness" into its various aspects. This approach reveals the diverse facets of effectiveness that might otherwise remain concealed, emphasizing that a singular focus on any one aspect skews the comprehension of effectiveness (e.g., Goldsmith et al., 2014; Kennedy, 2019).

Figure 44

A Prism Approach to Evaluating Teacher PLD Effectiveness



For instance, analyzing the PLD effectiveness through its *Reaction* level offers a window into its *perceived* effectiveness. This aspect revealed teachers' overwhelmingly positive experiences with various hallmarks of the program, which is suggestive of an effective PLD initiative. Such positive findings are in line with PLD effectiveness evaluation studies that qualitatively analyzed teacher self-reported data (e.g., Fauskanger & Bjuland, 2019; Gibbons & Okun, 2023; Jakopovic, 2021). Although earlier research has demonstrated a satisfactory level of agreement between teachers' self-reported data and classroom observations—particularly in terms of the frequency, rather than the quality, of teaching practices (e.g., Desimone, 2009)—relying solely on emotional and cognitive responses may provide an incomplete picture of PLD effectiveness. This approach potentially overlooks or fails to directly correlate with the actual impact of PLD on teachers' teaching practices and learning (Lindvall et al., 2022). Therefore, there is a need to investigate the teacher participants' practice.

Shifting the focus to the *Results* level, the end-of-program quantitative evaluation returned a mixed picture of its effectiveness, commonly seen in quantitative PLD evaluation research (e.g., Dash et al., 2012; Hill et al., 2018; Jacob et al., 2017; Kraft & Hill, 2020). In their culminating lessons, teachers as-a-group extensively experimented with cognitive activation practices, applied differentiation practices to varying extents, and less frequently integrated both axes. Notably, attempts to focus on the interplay of both axes often yielded high-quality teaching, despite observable challenges in orchestrating whole-class interactions across all axes. The mixed results at the program's conclusion invite a deeper inquiry into its "success" (Jacob et al., 2017).

This question is particularly compelling considering that understanding and integrating new teaching concepts and practices as advocated by the focal PLD program typically takes one to two years (e.g., Kennedy, 2019; Timperley, 2008). Yet, Lindvall et al. (2023) propose that the duration of PLD *itself* does not guarantee positive results. This is attributed to the considerable time teachers need to master new teaching practices, along with the necessity for *multiple* opportunities for experimentation and reflection, as emphasized by Darling-Hammond et al. (2017). This level of effectiveness suggests that success in PLD initiatives is inherently context-sensitive and demands comprehensive evaluation beyond singular assessments, such as baseline comparisons or longitudinal analysis, which can reveal the underlying processes concerned with teachers' teaching behavior that led to the final outcomes or products (cf. Creemers et al., 2013).

Although valuable, focusing solely on the end results in teachers' practice (*product perspective*) undervalues the rich processes that may have contributed to changes or the lack thereof (*process perspective*), potentially leading to misinterpretations of what constitutes PLD effectiveness (Borko, 2004; Borko et al., 2010). To grasp the PLD program's full effectiveness, the study focused on analyzing teachers' teaching *Behavior* over the course of the PLD, confirming that the relation between PLD input and final results is more complex than assumed (Creemers et al., 2013). A longitudinal view, spanning across the program's timepoints, showed the *dynamic* nature of the development of teachers' practice, contrasting with the *static* snapshot provided at the program's endpoint. Quantitative analysis of teachers' collective teaching behavior throughout the program showed a steadily increasing emphasis on cognitive activation, fluctuating engagement with differentiation, incremental and high-quality efforts to use teaching practices that promote the interplay of both axes, as well as low-quality in managing whole-class interactions.

Remarkably, Wilcoxon signed-rank tests pinpointed instances of statistically significant improvements in some practices, spotlighting the program's effectiveness in refining certain aspects of teaching quality. Most statistically significant differences were observed in the *maximum* performance—especially in student autonomous work—of teachers rather than in their *mean* performance. This finding offers insightful implications into the dynamics of the program's effectiveness and teacher development. First, it implies that teachers experimented with the program ideas and achieved significant high-performance levels at least once in their lessons, which is suggestive of the program's effectiveness. Second, it could also indicate that improvements were noted in areas in which teachers had the potential to excel, leading to notable improvements in their maximum levels of performance. This underscores the importance of considering a broader range of teacher performance metrics, rather than solely focusing on mean scores.

Third, it denotes that while the general cohort of teachers may have exhibited minimal significant shifts in mean performance across timepoints, the outliers or high performers within that group may have shown greater changes. This could imply that PLD interventions might be particularly effective for those at the higher end of the performance spectrum, or conversely, that only highly accomplished teachers can leverage these interventions to their full advantage. Case studies could explore whether these teachers "experience the program as a starting point, not a finished package" (Carpendale et al., 2021, p. 1). Research suggests that transitioning to ambitious mathematics teaching is challenging, with many teachers failing or not attempting it due to its complexity (Kennedy, 2016). Thus, this finding indicates the need for focusing on individual growth paths which may call for targeted and customized PLD programs to elevate the performance of all teachers, not just those already excelling (Creemers et al., 2013; Hill et al., 2018). This perspective could lead to a reevaluation of how PLD effectiveness is measured, recognizing the diverse potential within teacher populations.

In this sense, the use of both mean and maximum score metrics offers methodological insights at the *macro level*, which help explain the mixed findings observed in the quantitative component (Hill et al., 2018). Goldsmith and associates (2014) called for adopting a *micro-level* perspective, by focusing on an in-depth examination of the learning processes of teachers during the PLD, which would likely unveil different learning pathways. The learning processes often go unexplored, relying instead on implicit theories suggesting that PLD leads to shifts in learning, which subsequently influence teaching practice (Clarke & Hollingsworth, 2002). What is lacking in these accounts is a clear understanding of the processes driving these changes. Therefore, by focusing on the *Learning* level, teachers' conceptual changes can be revealed, encompassing shifts in their understanding of cognitive activation, differentiation, and their interplay, thus addressing the missing aspect of "how these changes [shifts in teachers' practice] came about" (Horn & Garner, 2022, p.2).

In this context, the case-study component of the current study delved deeper into the process layer with a focus on teachers' learning and behavioral processes, as well as their alignment as they evolve (e.g., Helsing et al., 2008; Kazemi & Hubbard, 2008). This analysis illuminated three distinct paths of teacher development, revealing themes related to teaching and learning and bringing to light the challenges faced by teachers as they navigate the terrain of ambitious mathematics teaching (Antoniou et al., 2015). These cases are further explored in the third topic (see Section 8.4).

In summary, the metaphor of the 'prism' evaluation problematizes the tendency to focus solely on the products of a PLD program—*what* teachers are doing in their culminating lessons— suggesting that a different delineation of its success might emerge when attention is given to *how* and *why* certain outcomes were achieved, namely the processes leading to such products (Creemers et al., 2013; Hill et al., 2018; Kennedy, 2016). This approach can help unravel the complexity of PLD effectiveness (akin to white light), allowing us to see beyond surface-level results and understand the multiple levels of PLD effectiveness (akin to the spectrum of colors), thereby enabling a deeper understanding of the program's influence on teacher development. Each color of the spectrum provides a richer, more complete picture of PLD effectiveness, emphasizing that a holistic view—integrating both the processes and its products—provides a clearer insight into the program's true impact. The four levels of

Kirkpatrick's model provide evidence that contributes to Kennedy's urge (2016; 2019) to better unpack the PLD effectiveness in order to better understand and appraise it.

8.3 Enhancing Teacher PLD: The Case of the EDUCATE PLD Program

Acknowledging the significance of integrating established effective features into the design and implementation of effective PLD programs (Borko et al., 2011), the focal PLD program which was based on the EDUCATE project made conscientious efforts in this realm (see Table 8 in Chapter 3). Whereas in the preceding section, we emphasized the importance of analyzing the effectiveness of a PLD program from multiple perspectives, in this section we focus on examining the particular characteristics of the EDUCATE PLD program that appear to render it effective. Echoing Kennedy's admonition (2019) to identify effective aspects of PLD programs so that we can then "package" them "so that others can adopt them" (p. 149), in this section, we aim at pinpointing to such features, aspiring that doing so could contribute to ongoing discussions about improving PLD efforts.

In Chapter 2, after examining empirically validated conceptual frameworks (cf. Darling-Hammond et al., 2017; Desimone, 2009; Merchie et al., 2016; Timperley, 2008), seven key features of effective PLD were synthesized into a model with three layers (i.e., the individual teacher-learner, the PLD group, and the PLD program itself), providing a helpful *heuristic device* for PLD designers (see Figure 11). However, the reactions of the teacher participants in his study to the program revealed that although these seven features are often discussed as distinct elements in the literature, they appear to have a more dynamic relationship and may contribute to the effectiveness of PLD as a *collective*. This suggests that they should be considered holistically as a cohesive set or "package" of features, each playing a significant role in enhancing PLD outcomes.

Notwithstanding the interactive nature of these features, *all* participating teachers agreed that the following four features were particularly effective: (a) Opportunities for participating in professional learning communities; (b) Developing teacher professional inquiry and self-regulatory learning skills to increase ownership; (c) Extended duration with multiple opportunities to revisit and experiment with new practices; and (d) Active learning by reflecting on experience . It should be noted that for teachers, these features are not considered separately; instead, they were discussed by the teachers in a blended manner, with one being integrated within the others.

Specifically, all four features were enhanced by the video-club component, which was the link between the *collaborative* nature of teacher learning and the *contextualized* nature of teaching (Xia et al., 2022). By incorporating clips from real classrooms in which teachers were implementing ambitious teaching practices, teachers were not only aided in developing images of this unfamiliar teaching approach (e.g., Barnhart & van Es, 2020) but also in envisioning themselves implementing it and receiving feedback from peers on their implementations. Such collective action and critical reflection on teaching among teachers enabled what literature refers to as "a sense of collective responsibility" (Creemers et al., 2013, p. 57), which stood in contrast to the isolation typically associated with teaching (Xia et al., 2022).

Teachers also emphasized the significance of video-clubs in cultivating a supportive *professional learning community* (Alles et al., 2019). The small group size, the collective video analysis of their lessons and those of their peers (Van Es, 2012), the trustful and open exchange of ideas and advice, the discussions addressing common content, concerns, and challenges, and attempts to find solutions (Merchie et al., 2016) contributed to this environment.

Furthermore, the inclusion of the video-club component in the PLD enriched teachers' self-awareness during their teaching endeavors (Timperley, 2008). This aided teachers in developing an increased awareness and understanding of their teaching practice. By reviewing and deliberating on videoclips of their lessons alongside their peers, teachers engaged in reflective analysis of their practices, generated new ideas, established connections to their own work, and developed diverse perspectives. Consequently, teachers obtained valuable insights into their strengths, weaknesses, and areas requiring improvement (i.e., self-regulatory learning skills), actively seeking feedback to enhance their teaching practices.

Also, teachers benefited from actively co-constructing knowledge with teacher educators rather than having ideas imposed upon them (Desimone, 2009; Merchie et al., 2016), during stimulating, thought-provoking peer discussions, indicative of their involvement in intellectually challenging work (Kennedy, 2016). Due to the video-club component and the extended duration of the program, they had multiple opportunities to experiment with practices in their lessons and to reflect upon them (Sims et al., 2021). The analysis of the three teacher cases revealed that their reflection on videotaped lessons during video-club discussions often facilitated changes in their learning and teaching practices. These findings corroborate the theoretical model proposed by Clarke and Hollingsworth (2002), which suggests that experimentation and reflection are factors that mediate teacher learning and teaching practice.

Notably, the multifaceted analysis of the effectiveness of the EDUCATE program revealed two practices not systematically covered in the literature on PLD effectiveness but deemed necessary for consideration. These practices can further enhance and expand the synthesis with the key features of effective PLD. The first pertained to equipping teachers with *praxis tools* (cf. Windschitl et al., 2011). These tools bridge theory and practice, offering actionable ideas for teachers to implement directly in their planning, teaching, and reflection. Given the inherent ambiguity of ambitious teaching, teachers often struggle to operationalize ambitious teaching, especially when juxtaposed with the abundance of tools available for traditional methods, such as teacher-led discussions, textbook-centered lessons, and extensive curriculum coverage (Sykes et al., 2010).

Teachers highlighted the importance of *naming* these praxis tools, enabling them to observe theoretical ideas in action (Van Driel & Berry, 2012), such as enablers and extenders, diverse mathematically challenging questions, and task analysis. Introduced in video-club sessions from Timepoint 3, these tools marked a pivotal juncture in EDUCATE. Several significant differences between initial (Timepoints 1-2) and later (Timepoints 3-5) sessions underscored the exploration of these praxis tools in teachers' lessons. The several statistically significant differences in the axis of cognitive activation and the interplay of both axes contribute to prior research, showing how praxis tools, such as enablers and extenders, can provide concrete examples of ambitious mathematics teaching for practicing teachers (cf. Charalambous et al., 2022). Along the same lines, it could be argued that the inferior improvement observed in teachers' practice in the axis of differentiation may be attributed to the program's comparatively limited provision of praxis tools for this axis in comparison to the other axes (i.e., cognitive activations and the interplay of the axes)—which signifies a potential weakness of EDUCATE. The complexity of implementing differentiation, requiring additional experimentation, time, and practice for teachers to effectively integrate it into their teaching, may further exacerbate this issue (Kyriakides et al., 2009).

The second effective practice identified by and for teachers was meta-cognitive in nature, referred to as *organizing the praxis tools into meaningful toolkits or packets*. Specifically, within the focal PLD, various praxis tools were organized into three packets tailored to each phase of the lesson: task launching, student autonomous work, and whole-class interactions (during VCS3-VCS5). Moreover, these three packets were integrated into an overarching organization for lesson planning (during VCS6-VCS8, see Table 7 in Chapter 3). The positive changes identified in the analysis of teachers' teaching may demonstrate the possible influence of organizing the praxis tools into meaningful packets. The organization was also evident in how teachers discussed these practices during their interviews and video-club sessions. This practice potentially provided insights into how cognitive activation and differentiation could be seamlessly integrated into a lesson, given the challenge teachers face in maintaining a consistent focus on both axes throughout the teaching process (Charalambous et al., 2023b, p. 181).

Despite the contributions of the program, findings from the four research questions also revealed certain limitations concerning the PLD feature of involvement of knowledgeable and high-quality experts in planning and facilitating PLD (see Figure 11). While teacher educators utilized all of the key practices for conducting productive video-club discussions (Van Es et al., 2014), they neither *systematically* addressed the individual differences among teachers nor provided personalized feedback to teachers.

In particular, recognizing the diversity among participants in the video club, as evidenced by prior research (e.g., Creemers et al., 2013; van Es et al., 2017a), efforts were made to elicit and accommodate teachers' needs through their reflection cards at the end of every video-club session. However, the case studies suggest that this measure may not have adequately supported the monitoring and customization of the PLD to teachers' individual needs and progress. Therefore, participants' individual developmental stages may not have been systematically considered, resulting in the oversight of diverse needs, and treating teachers as a homogeneous entity throughout the program. The case of the EDUCATE reveals a complex "*balancing act*" for teacher educators (cf. Arizona Group et al., 1996) when facilitating PLD initiatives on ambitious mathematics teaching. Balancing the support provided to teachers to harmonize cognitive activation with differentiation, accommodating diverse teacher needs, and juggling multiple program features within time constraints may pose intricate challenges for teacher educators, limiting accurate tracking of teacher development over time.

Furthermore, while teacher educators offered teachers the opportunity to receive feedback at their convenience (e.g., before each video-club session), this clashed with teachers' available free time, such as in the case of Michelle, resulting in missed opportunities for further improvement. As some teachers themselves acknowledged, they could have benefited more from receiving *systematic* feedback and suggestions on their practice related to their priorities for improvement (Creemers et al., 2013).

In summary, the EDUCATE program integrated several effective features of PLD into its design and implementation. The findings highlighted the significance of several key practices for teacher improvement, including fostering professional learning communities; developing teacher self-awareness, inquiry skills, and ownership; providing multiple and extended opportunities for experimentation; and promoting active learning through reflection. The video-club PLD model played a pivotal role in facilitating these practices. Additionally, the analyses suggested the effectiveness of two less commonly emphasized practices in the literature: providing *praxis tools* for the direct implementation of ambitious mathematics teaching, and *organization into meaningful tool packets* for maintaining the focus on this kind of teaching throughout the teaching process. Despite these strengths, the program faced challenges in *systematically* tailoring PLD to individual teacher needs, and *systematically* providing feedback to teachers, suggesting a roadmap for refining future PLD initiatives.

8.4 Identifying Multiple Learning and Behavioral Change Paths Within a PLD Program

The individual trajectories of the three cases—Pina, Kate, and Michelle—detailed in Chapters 5-7 within the EDUCATE PLD program vividly illustrate that not all participants experience the same level of gain or follow identical paths of development. This finding aligns with both theoretical assumptions (e.g., Goldsmith et al., 2014; Witherspoon et al., 2021) and earlier empirical studies (e.g., Anthony et al., 2018; Charalambous et al., 2018, 2022, 2023; van Es & Sherin, 2017a) supporting the notion that the same PLD opportunity yields varied outcomes on different individuals. The idea encapsulated by Creemers et al. (2013, p. 53) that "No single strategy will always work in every school, for every teacher, all of the time" underpins these concepts.

In particular, the three profiles highlight the diverse pathways of teachers participating in PLD programs, showcasing a spectrum from transformative growth to the nuanced challenges of balancing cognitive activation and differentiation. The first profile, Pina, serves as a testament to the potential of PLD programs to foster significant and integrated teacher development when the program focus aligns with the individual's prior experiences, growth mindset, and readiness for change. The second profile, Kate, captured as a series of fluctuating progress and regressions, mirrors the non-linear and often unpredictable nature of professional growth. This profile reflects the complexities of integrating new practices, emphasizing the need for perseverance and adjustive PLD support mechanisms, particularly when exploring unfamiliar teaching territories. The third profile illustrates *caring* (cf. Noddings, 2001) but sometimes counterproductive efforts to support students, reflecting the challenges faced by teachers who must learn to moderate their support to enhance, rather than inhibit, student autonomy and thinking. Collectively, these cases echo an increasingly discussed theme in the literature highlighting the importance of assessing and monitoring teacher needs, stages, and priorities, to adjust and align PLD initiatives with them (e.g., Hill et al., 2018). Next, the profile of each teacher, the factors that may have influenced their developmental trajectory, as well as the kind of support they may have needed to grow further are discussed.

Pina's progressive and coherent development in learning and teaching behavior, along with her shift towards focusing on the synergy between cognitive activation and differentiation—instead of viewing these elements separately—exemplifies the success of extended video-club PLD programs in supporting practicing elementary teachers with ambitious mathematics practice. This adds to the existing body of research on the effectiveness of the video-clubs as a PLD model for practicing teachers (e.g., Santagata et al., 2021; Van Es et al., 2017b).

Additionally, her case postulates factors that contribute to making such progress, including the beneficial impact of aligning PLD content with a teacher’s background, prior learning experiences, mindset, and readiness. Specifically, her robust background in mathematics teaching, enhanced by a master’s in mathematics education—identified in the literature as supportive of enhancing cognitive activation and ambitious mathematics teaching (e.g., Chapman, 2013; Hill et al., 2005; Kelcey et al., 2019), together with her previous beneficial encounters with PLD centered on cognitive activation and mathematically challenging tasks, may have set a solid foundation for her effective involvement with the program's focus (e.g., Hill et al., 2018). Furthermore, Pina might have been "ready to make the transformation" towards ambitious mathematics teaching, with the EDUCATE program aligning closely and meaningfully with her starting point at that time—her existing practices and conceptual understanding (Otten et al., 2022, p. 1445). Her growth mindset—believing that experimentation and addressing challenges help her students and herself grow as a teacher—may also have supported her growth during PLD (e.g., Tomlinson & Imbeau, 2023). Despite her improvements, her teaching in the final lesson showed potential for further improvement. This may be linked to certain challenges she encountered, which are discussed later for all three cases. Maybe cases similar to Pina's could be better supported by offering PLD tailored to her specific challenges, such as refining the use of certain praxis tools, such as enablers, and expanding her repertoire of praxis tools.

Kate's teaching performance across the program ebbed-and-flowed, while she consolidated certain practices. Trying to simultaneously juggle multiple practices, she often enacted a new practice in one of her videotaped lessons and then abandoned it in the next; in other instances, she implemented a new practice in a decrescendo way across her lessons. This pattern may suggest that what the teacher educators believed to be a progressive introduction of ideas rather might have been at odds for her; focusing on a new idea or tool each time might have detracted from productively solidifying her experimentation with the older ones.

Reflecting on Kate's case, her situation can be seen as a reverse illustration of *"the problem of enactment"* (cf. Kennedy, 1999, p. 70). The problem of enactment occurs when teachers agree with new ideas in theory but stick to old habits in practice, struggling to adopt different teaching practices due to entrenched behaviors (see also Cohen, 1990, the case of Mrs. Oublier). Instead, Kate willingly and actively experimented with new practices from the PLD program but struggled to *sustain* them amidst the gravitational pull of her established habits and the overwhelming nature of handling multiple practices. Her case might more accurately be termed *"the problem of sustainability,"* highlighting her initial relatively successful adoption of new practices but difficulty in consistently implementing them, pointing to the need for PLD programs to not only focus on the introduction of new ideas but also support their enduring integration into daily teaching.

In addition, Kate's case underscores the significance of aligning PLD programs with teachers' pedagogical interests and their previous PLD experiences, suggesting that relevance to participant's interests may be a key factor in the effectiveness of PLD programs. Her keen knowledge and interest in differentiation, influenced by prior PLD engagements, primarily drove her participation in the EDUCATE program. According to Horn and Garner (2022), successful PLD designs must consider teachers' preconceptions, experiences, and social histories, emphasizing the importance of acknowledging and building upon their existing understandings. For Kate, starting with the familiar concept of differentiation and gradually introducing new concepts like cognitive activation could serve as a strategic approach to facilitate her growth during PLD.

Michelle's case exemplifies a common scenario in teacher PLD, where conceptual learning and practice do not always move in tandem. Despite improving the use of certain practices throughout the program, such as incorporating challenging tasks in her teaching, her overall teaching trajectory did not exhibit a clear pattern of change. Described as a "caring teacher"—to use Noddings' term (2001)—she was dedicated to her students' learning. However, her caregiving approach was characterized by a low tolerance for student mistakes and a tendency to excessive guidance (e.g., Smith & Stein, 1998). While her understanding of cognitive activation and the interaction between various teaching dimensions saw significant improvement, her approach to differentiation remained somehow static.

Michelle's (and Kate's) pattern, identified in the literature as an improvement in learning/conceptualization without a corresponding pattern of change in practice, stands in contrast to Pina's experience, which saw advancements in both areas (e.g., Tam, 2015). This underscores the "idiosyncratic and individual nature of teacher professional growth" as described by Clarke and Hollingsworth (2002, p. 965), highlighting the diverse and personalized paths of PLD. Yet, Michelle's pattern might suggest that changes in conceptualizations do not necessarily lead to corresponding, immediate, changes in teaching practices. Cases like Michelle's highlight the complicated nature of teaching and the necessity for PLD programs to not only focus on enhancing teachers' conceptual understanding but also to provide tailored and robust mechanisms that facilitate the translation of these conceptual gains into tangible classroom practices.

It is therefore necessary to explore potential factors that may have prevented the corresponding change in her practice. It is hypothesized that her conceptualizations could have been shaped by her theories, beliefs, and dispositions about teaching and learning, which oftentimes are deeply ingrained and resistant to change (e.g., Boston, 2013; Creemers et al., 2013; Erotocritou-Stavrou & Koutselini, 2016; Thompson et al., 2013). Furthermore, her teaching pattern may have been influenced by her "ethic of care" towards less advanced students, echoing the findings from another case study by

Anthony and colleagues (2018, p. 655) involving Tina. Similar to Tina's approach, Michelle's practice of consistently grouping students into fixed ability levels dominated her strategies for support while navigating between ensuring accessibility and offering mathematical challenges. Driven by genuine care for her students, she wanted to facilitate her student learning by truly trying to engage less advanced students with the tasks, potentially at the expense of more advanced students.

This suggests that enabling Michelle to evolve her teaching practice may require a focused examination of her deeply ingrained beliefs about teaching and learning. Specifically, it involves resolving the conflict/contradiction between her ethic of care for less advanced students and the imperative to offer an appropriate level of mathematical challenge, without sidelining any student group (Anthony et al., 2018). Additionally, it calls for navigating the tensions and discrepancies between her existing and developing practices (Anagnostopoulos et al., 2020); a reassessment of her roles and responsibilities, alongside those of her students, within classroom dynamics (Hunter, 2008); and an alteration of her misconceptions about differentiation, to understand that it extends beyond simply creating homogeneous groups and increasing or reducing students' work (Tomlinson, 2017).

Moreover, the slight improvements observed in her final lesson, particularly in sustaining the mathematical challenge over a longer period, reiterate the importance of recognizing the extended time needed for transformative change (e.g., Kennedy, 2019; Timperley, 2008). With a dedication to continuous experimentation through PLD, Michelle could be able to implement more profound changes to her learning and teaching practice in the following years.

In addition, the experiences of Pina, Kate, and Michelle collectively during their participation in the PLD program illustrate a nuanced landscape of both progress and persistent challenges at the nexus of cognitive activation and differentiation. Beyond their patterns of change, their heterogeneity was further underscored by the diverse challenges they each faced. The examination of these challenges offers a comprehensive insight into the effectiveness of the EDUCATE program.

Although certain challenges were effectively addressed and mitigated during the PLD, others remained unresolved, indicating areas requiring further intervention. Notably, all three teachers demonstrated improvements in minimizing directive interactions and enhancing student participation in mathematical reasoning. The adoption of extenders and enablers along with asynchronous work as praxis tools to accommodate diverse student needs and readiness levels further underlined the program's impact on differentiating the mathematical challenges presented to students. Documenting the teachers' progress in these challenges contributes valuable evidence to the effectiveness of PLD programs in promoting student-centered learning environments.

However, unresolved challenges such as balancing the depth of the mathematical challenge for all students within time constraints, appropriately developing and enacting enablers, supporting both less advanced without marginalizing more advanced students, cultivating a classroom culture conducive to cognitive activation and differentiation, and addressing student-related challenges (e.g., student indifference, motivational barriers for students from immigrant backgrounds, and bridging the knowledge gaps of higher-grade students) highlight the complexities of enacting ambitious teaching practices. The identification and articulation of these enduring issues emphasize the need for ongoing support and can guide future refinement of PLD efforts aimed at addressing ambitious mathematics teaching.

Finally, the current study makes a significant contribution to the research on the challenges of ambitious mathematics teaching by articulating two unaddressed challenges not discussed in the literature. Specifically, Pina grappled with the challenge of *judiciously interrupting autonomous work*, characterized by uncertainty about the optimal timing for transitioning from individual student work to whole-class discussions. This challenge delves into the complex decision-making process regarding when to encourage autonomous exploration of concepts by students and when to convene whole-class discussions that aim to reinforce or expand upon these individual learnings. The second challenge refers to *maintaining the mathematical challenge during whole-class interactions, despite bringing the class to the plenary at appropriate checkpoints*, faced by Kate. Organizing the lesson into three phases likely aided in identifying these challenges, due to their association with transitioning between student autonomous work and whole-class interactions. These challenges highlight the novelty of this study, shedding light on previously underexplored aspects of teaching practice. Furthermore, by drawing attention to these challenges, the study encourages a more detailed exploration of the scaffolds teachers require to address them.

In sum, this study delves into the differential effectiveness of the EDUCATE PLD program on three teachers—Pina, Kate, and Michelle—highlighting the individualized, non-linear nature of teacher growth and the challenges encountered in ambitious mathematics teaching. It underscores the pivotal role of aligning PLD content with teachers' unique backgrounds, learning experiences, and readiness for change to foster meaningful but not uniform changes across all participants. Furthermore, the research identifies two previously unexplored challenges—judiciously interrupting autonomous work, and maintaining the mathematical challenge during whole-class interactions—contributing new insights into the complexities of ambitious mathematics teaching. These insights call for further research into PLD strategies that effectively support diverse teacher profiles and facilitate sustainable professional growth. Moving forward, it is crucial to explore the scaffolds and support mechanisms that can address the specific challenges identified, ensuring that PLD programs are both adaptable and targeted to meet the evolving needs of teachers.

8.5 Kirkpatrick's Model Revisited: Complexities and Challenges in Evaluating Professional Learning and Development

Although not its primary objective, this thesis inadvertently provided empirical validation of Kirkpatrick's model (2007), showcasing its affordances and limitations. The research findings demonstrated that the model served as both a *theoretical* and an *analytical* tool. Initially, it was used to assess and measure the outcomes of the PLD program (analytical use). The study showed that the effectiveness of a PLD program should be measured *both* quantitatively and qualitatively, during and at the conclusion of the program.

Specifically, the qualitative component was crucial in providing deeper insights into the diverse teaching behaviors and the mixed final results observed, as well as in uncovering the context within which learning and practice occur, explaining why certain outcomes emerged. For example, the quantitative analysis portrayed Kate as not particularly successful in the PLD program, whereas the qualitative analysis depicted her as a teacher who actively took up and experimented with different ideas in her practice, despite facing challenges that might have hindered her performance beyond what her lesson analysis suggested. Perhaps, she needed more time to refine her practice, and the program inadvertently overloaded her. Conversely, Michelle's qualitative analysis highlighted a conflict or dilemma between the ethics of care and ambitious teaching explaining her inconsistent teaching performance as captured by the quantitative component. Cases like Kate and Michelle, who were eager participants in EDUCATE, raise a red flag on how they can better be supported and how the PLD can be differentiated to do so. Even for Pina, who might be considered to meet the program's expectations due to observed improvements in quantitative metrics, the qualitative analysis provided insights into how she progressed and identified unresolved issues that could potentially hinder her further development. Therefore, the qualitative component is critical in identifying variables that quantitative metrics might overlook, including emotional responses, levels of engagement, or even *subtle* shifts in conceptualizations and practice that are significant yet less apparent in numerical data.

Furthermore, the application of the model also shed light on the *understanding and conceptualization of PLD effectiveness* (theoretical use). As discussed in Section 8.2, the Kirkpatrick's model (2007) promoted a holistic multi-faceted view of PLD effectiveness by evaluating multiple levels, thus responding to calls for more comprehensive PLD evaluation (McChesney & Aldridge, 2019). This approach not only aids in measuring effectiveness but also challenges the theoretical understanding of what makes PLD successful. For example, the final results showed mixed effectiveness, while teachers' reactions suggested a more successful story. This discrepancy highlights the need to consider multiple dimensions of evaluation to gain a comprehensive understanding of PLD effectiveness. Traditional single-level PLD evaluations leave blind spots, failing to capture the full scope of a program's impact.

However, the model is not without its limitations. Contrary to theoretical assumptions that each level of Kirkpatrick's model sequentially impacts the next (e.g., positive reactions lead to improved learning, which then leads to behavioral change, and ultimately to enhanced results, e.g., Guskey, 2024), the findings showed that success at one level does not necessarily lay the foundation for success at a higher level, nor does it predict or support success at subsequent levels. Specifically, the predominantly positive evaluation of the program by the participants (reaction level) could suggest potential for high teaching performance during (behavior) and at the conclusion of the program (results). Yet, this was not observed, as teachers followed diverse learning and teaching paths. Furthermore, the cases of Michelle and Kate illustrated that learning and teaching progress do not necessarily coincide; although they enhanced their conceptualization of cognitive activation/differentiation and their interplay, their teaching practices did not exhibit corresponding improvements. Thus, based on these observations, it is clear that we cannot speak of lower and upper levels of PLD evaluation, as the relationship between the levels remains unclear, particularly between the levels of learning and teaching behavior, which sometimes coincide and at other times are distinct, indicating a non-linear relationship between evaluation levels. These diverse patterns challenge the straightforward, predictive utility of the model across different cases.

The inability to clearly define and differentiate between 'lower' and 'upper' levels of evaluation suggests that these levels are not distinct stages but perhaps part of a more *interconnected* and *overlapping* process. In reality, they often interact in more complex ways, such as in feedback loops or simultaneous development, influenced by the impact of internal and external factors. Learning and behavioral changes can influence each other in cycles of enactment and reflection (cf. Clarke & Hollingsworth, 2002). For instance, as teachers implement new practices (behavior), they reflect on the effectiveness of these changes and adapt their understanding (learning), which further influences subsequent behavior, making a feedback loop. Alternatively, changes in behavior and learning can occur *simultaneously* rather than sequentially. For example, a teacher might simultaneously learn a new concept and apply it in practice, with the learning and application phases blending into each other and informing each other in real-time. Of course, the interplay between different levels of effectiveness can be affected by factors such as organizational culture, support, or pressures (cf. Guskey, 2024). These factors can modify how one moves through the levels of evaluation or how these levels influence one another. Kirkpatrick's model does not explicitly define these relationships, indicating that further research is necessary to explore them.

This nuanced critique of Kirkpatrick's model invites a rethinking of how PLD *effectiveness* is defined, measured, and interpreted, advocating for an approach that acknowledges the complexity and individuality of teacher development. The findings of this thesis propose a fresher conceptualization of PLD effectiveness as a holistic process across multiple levels. Focusing on any single level of evaluation or relying solely on single metrics is neither ethically sound nor sufficient to provide an accurate depiction of PLD effectiveness. In this respect, the following question arises: how *ethical* or *equitable* is it to measure the impact of PLD on *all* teacher participants based solely on their final results or the quantitative assessment of their teaching behavior? Teachers' personal experiences, the specific contexts of their educational environments, and their individual differences result in varied teaching outcomes. More relevant questions emerge: Should the evaluations be summative or formative?

Should the focus be on informing the research community about the effectiveness of PLD, or should it be on providing teacher educators and teachers with data for continuous improvement? In essence, this thesis contributes to the discourse on the main challenges of measuring PLD effectiveness, which include the content—the what, the methods or properties—the how, as well as the timing of measurement—the when (Alicea et al., 2023; Desimone, 2009). As a research community, we have yet to reach a consensus on this issue. Recognizing the affordances and limitations of Kirkpatrick's model, the current thesis highlights the necessity for further elaboration and research to fully understand and optimize its application in varied PLD settings.

8.6 Implications

The study presents implications that are theoretical, methodological, and practical:

Theoretical implications. Theoretically, it contributes insights into (a) the nature of teacher change and learning, and (b) the importance of intertwining cognitive activation with differentiation, treating teaching practices as a complex net rather than as separate entities.

Regarding the first theoretical contribution, the study provides crucial insights into the developmental trajectories of teachers, encompassing both their learning/conceptualizations and their teaching practice, particularly in the realm of PLD (e.g., Schoenfeld, 2023). Specifically, it sheds light on the diverse ways teachers learn and implement new teaching practices (e.g., Witherspoon et al., 2021). The comparison between teachers' patterns of learning and their practice patterns illustrated "multiple change sequences and a variety of possible teacher growth networks" (Clarke & Hollingsworth, 2002, p.965), with some aligning and others misaligning.

Hence the study enriches theoretical perspectives on teachers' change showing that it is personal, non-linear, situated, and more complex than previously assumed (e.g., Guskey, 2002), encompassing an ongoing interplay between teacher learning and practice. By focusing on teachers' learning and behavior from a *process* perspective, we can trace the journey to the endpoint of their teaching practice—a *product* perspective, which are further discussed in the methodological implications. This approach illuminates "how teachers interpret and use the available understandings and skills", thus offering a more nuanced view of teacher development—the black box linking the

PLD program to their ultimate final teaching practice (Timperley & Alton-Lee, 2008, p. 340).

With respect to the second theoretical contribution, the study contributes to theoretical understandings of the intertwined potential of cognitive activation and differentiation within the domain of ambitious mathematics teaching. Delving into the dynamic interaction between these two axes underscores the necessity for conceptualizing the teaching process as a holistic system wherein components interact synergistically in multiple ways, rather than as a mere aggregation of discrete, interchangeable elements (cf. Hiebert & Grows, 2007).

Notably, the three teacher case studies gradually focused on the synergistic potential of the two axes, rather than viewing them as separate entities (as shown in their learning processes). This shift underscores a convergence between teachers' conceptual frameworks and the theoretical frameworks of researchers, challenging traditional teaching theories that treat practices as isolated or mutually exclusive (cf. Charalambous & Praetorius, 2020).

Methodological Implications. From a methodological standpoint, the study illustrates three key implications: (a) the value of holistically evaluating PLD programs in multiple levels, by assessing both products and processes, (b) the effectiveness of video clubs in facilitating ambitious teaching practices among practicing elementary teachers, and (c) the affordances of utilizing both mean and maximum performance metrics in the evaluation of teachers' work and their experimentation with different ideas.

Concerning the first implication, the study calls for a re-evaluation of PLD success criteria, advocating for assessments that consider both the *product* and the *process* of teacher development. Specifically, the first topic contributes to the concerns regarding the ways we learn about the effectiveness of teacher PLD programs (e.g., Timperley & Anton-Lee, 2008). By looking at the processes, researchers can identify *what* works, *how* it works, and *under what conditions* it works best. By valuing both the journey (process) and the outcome (product), this comprehensive evaluation approach offers richer insights into the effectiveness of the PLD program.

The use of the Kirkpatrick's model suggests a shift towards a holistic evaluation of teacher PLD effectiveness, advocating for a nuanced approach that goes beyond traditional metrics, such as focusing solely to teachers' *reactions* or before-and-after program *results*. Emphasizing the multi-faceted nature of teacher development, this perspective draws on the metaphor of a prism to illustrate the need for examining PLD impacts across various dimensions—teachers' reactions, teaching results, behavioral and learning evolutions. Such frameworks consider the temporal aspect of teacher development that can track the progressive effectiveness of PLD over time, through both qualitative and quantitative methods.

Furthermore, the employment of video clubs to facilitate ambitious teaching practices among elementary *practicing* teachers stands out as a significant methodological contribution in this study (e.g., van Es et al., 2017a). Uniquely, it ventures into uncharted territory by examining how these teachers conceptualize and experiment with practices of cognitive activation and differentiation, as well as their interplay in their classrooms. Through this PLD model, the study illuminates which praxis tools teachers find most beneficial, along with the presence of consistency and inconsistency in the implementation of these tools.

As regards to the third methodological implication, using both mean and maximum scores to evaluate teacher performance in PLD programs carries significant implications for understanding and enhancing teacher change and the complexities involved in the PLD process. This dual metric approach enables a more nuanced analysis of teacher performance over time. Mean scores shed light on the general or average level of performance, offering insights into the consistency of teaching practices. Conversely, maximum scores illuminate the moments of highest potential, particularly when teachers experiment with new ideas from the PLD program. Recognizing the importance of peak performances alongside consistent teaching quality helps balance the pursuit of continuous improvement with realistic teaching practice expectations. It acknowledges that not every lesson will achieve the same high level and that maintaining high-quality practice in every teaching practice or lesson phase is impractical. Yet, it affirms the existence and potential for nurturing such high performance. By employing both metrics, this approach challenges the traditional single-metric evaluation models, advocating for a more comprehensive assessment strategy. This leads to the development of more sophisticated metrics that accurately reflect the complexity and dynamism of teaching performance.

Practical Implications. The insights gleaned from the EDUCATE PLD program offer significant implications for PLD designers/providers and teacher educators, especially in designing, delivering, and facilitating programs that focus on cognitive activation, differentiation, and their interplay. The practical implications of the study unfold in three key areas: (a) it demonstrates the feasibility of teachers working at the nexus of cognitive activation and differentiation; (b) it illuminates which PLD features effectively support teachers in striving toward this goal; and (c) it highlights the importance of teacher needs-assessment analysis during the PLD.

Regarding the first implication, it addresses the ongoing debate regarding the feasibility of teachers working at the nexus of these two critical axes (e.g., Charalambous et al., 2023a). The study presents images of teachers experimenting with various ambitious teaching practices (Lampert et al., 2011), while also addresses a notable research gap by highlighting the challenges they face—challenges that are not always conducive to the adoption of these ambitious methods (Horn & Garner, 2022). Given the identified challenges in ambitious mathematics teaching, there is a clear need for further research to develop, evaluate, and refine specific support mechanisms, tools, and practices that can aid teachers in overcoming them.

In regard to the second implication, understanding the dynamic nature of the effective PLD features mentioned in Section 2.5.2 (see Figure 11) is paramount. A PLD program focused on cognitive activation and differentiation should not just list these as separate features but integrate them into a holistic system that supports teachers' development comprehensively. This integration necessitates a deliberate design that encourages teachers to engage deeply with the content, critically reflect on their practice, and actively experiment with new strategies in their classrooms.

Also, the use of *video clubs* has emerged as a particularly effective tool in this regard. By incorporating real classroom footage, teacher educators can provide vivid, concrete examples of cognitive activation and differentiation in action. This component aids teachers in visualizing these abstract concepts in practice, facilitating a deeper understanding and more effective implementation. Therefore, PLD designers should consider incorporating the video-club as a core component of their programs.

Moreover, equipping teachers with *praxis tools* specifically designed to enhance cognitive activation and differentiation is crucial. These tools ought to be specific, concrete, and directly applicable in the classroom; they should also be easy to design and implement, as well as time-efficient, thereby enabling teachers to effectively bridge the gap between theory and practice—a finding that aligns with Haug and Mork's (2021) research in ambitious science teaching. *Organizing these tools into meaningful packets or toolkits* can further support teachers by providing a structured approach to implementing these practices. This organizational strategy not only aids in clarity but also in ensuring that teachers can maintain focus on cognitive activation and differentiation throughout different phases of their lessons.

Finally, addressing the individual needs and challenges of teachers within the PLD program is another critical practical implication of the study. To do so, the teacher educators must collect assessment data at the *beginning* of and *throughout* a PLD (*process perspective*). Initially, diagnostic surveys or interviews can be employed to gather detailed information on teachers' previous experiences, current practices, perceived strengths, areas for development, and their expectations from the PLD. Incorporating baseline observations of teachers' classroom practices can provide practical insights into their teaching practices and observed challenges. Throughout the program, continuous assessment can be facilitated through reviewing the videotaped video-club discussions, in which teachers document their experiences, challenges, and insights as they try new practices, as well as their videotaped lessons. Regular feedback sessions with facilitators offer another layer of ongoing assessment, enabling adjustments to the program based on evolving needs. This way teacher educators can design their programs to be flexible enough to cater to this diversity. This could involve providing options for differentiated learning paths within the PLD, leading to more meaningful and sustained changes in teaching practice and student learning.

8.7 Study Limitations and Directions for Future Research

This study has presented and discussed its findings within the context of its inherent limitations—beyond those already discussed in the third chapter (see Section 3.10), necessitating a cautious interpretation of the results. Firstly, while this study systematically documents the multi-level effectiveness of the EDUCATE program, its exploratory nature did not aim to establish definitive relationships between the PLD program and teachers' change. Instead, it proposed a range of hypotheses—some potentially alternative—regarding factors contributing to the observed (differential) effectiveness of the PLD program and its potential influences on teachers' practice and learning. Given the tentative character of these hypotheses/postulates, further empirical research is necessary to explore and refine these hypotheses.

Moreover, the initial needs, previous experiences, and knowledge of the teachers were not assessed and there was no baseline measurement of their teaching practice, due to concerns about exerting undue pressure, which may have led to teacher attrition from the study. Future studies should aim to gather such data to tailor PLD interventions more effectively to the specific backgrounds and initial needs of participating teachers.

The voluntary nature of participating teachers in the EDUCATE program should also be considered when drawing conclusions. The self-selection bias (e.g., Bullman, 2021) means that the perspectives gathered may predominantly represent individuals already inclined towards positive engagement with and a more favorable view of the program due to their motivation and interest in PLD and its focal axes. While valuable, one should acknowledge the study's potential limitations and future research could explore PLD effectiveness among non-volunteer teacher populations, which may yield different outcomes.

The methodological approach, utilizing a single PLD model—the video-club format—may yield different results under alternative PLD models, such as coaching (e.g., Fernandez & Yoshida, 2004) or lesson study (e.g., Desimone & Pak, 2017). Additionally, the study employed the Kirkpatrick (2007) theoretical model to assess the PLD program's effectiveness. Alternative theoretical models (e.g., Clarke & Hollingsworth, 2002; Desimone, 2009; Evan, 2004; Opfer & Pedder, 2011), could provide different insights into the PLD's impact.

Conducted within the specific educational context of Cyprus—a centralized educational system with a single textbook published for each grade level—and engaging a small cohort of eight volunteer teachers, the study's findings are not intended to generalize across different contexts or educational levels, but rather to present a case of a PLD program engaging a small group of teachers. Replication studies in diverse contexts, in other subjects, and with different teacher demographics are necessary to expand upon these initial findings.

In addition, the study did not track the participating teachers following the conclusion of the intervention, limiting the understanding of the PLD program's longitudinal impact. Future research should adopt a longitudinal approach to assess the sustainability and evolution of the pedagogical changes initiated by PLD over a more extended period, ideally following the teachers for at least one full school year post-intervention (Kennedy, 2019). This approach could help understand whether changes due to interventions persist over time or teachers revert back to their baseline level after the intervention is over (Creemers et al., 2013).

Lastly, a significant limitation is the inability of the study to directly link the documented changes in teachers' teaching practice and learning to changes in student learning outcomes, a question beyond the scope of this study, but one that definitely merits consideration in future research. In fact, future studies could aim to categorize different teacher profiles or pathways and investigate how these variations in teacher learning and practice might contribute to student learning outcomes. Additionally, incorporating student learning as a critical aspect of the prism approach could significantly enhance the evaluation of PLD effectiveness, which can be a critique of Kirkpatrick's model (2007) and a potential way in which this model could be enhanced.

Appendices

Appendix 1 – List of Practices of the Coding Protocol

PRACTICES
COGNITIVE ACTIVATION PRACTICES

Task Launching

- **TL1-CA**-Selecting mathematically challenging tasks
- **TL2**-CA-Maintaining the cognitive demands of the task as presented to students during task launching
- **TL3**-CA-Discussing mathematical ideas as presented to students
- **TL-CA-overall**-Holistic estimate of the challenging work during this phase

Student Autonomous Work

- **SAW1**-CA-Providing mathematical prompts to students without trivializing their thinking
- **SAW2**-CA-Asking students to engage in mathematical reasoning and/or meaning-making activities;
- **SAW3**-CA-Engaging students in mathematical reasoning and/or meaning-making activities
- **SAW-CA-overall**-Holistic estimate of the challenging work during this phase

Whole-Class Interactions

- **WCI1**-CA-Eliciting instances of student mathematical reasoning and/or meaning/making
- **WCI2**-CA-Synthesizing and extending important mathematical ideas
- **WCI3**-CA-Asking students to compare or evaluate different solution approaches
- **WCI4**-CA-Engaging students in mathematical reasoning and meaning-making activities
- **WCI-CA-overall**-Holistic estimate of the challenging work during this phase

DIFFERENTIATION PRACTICES

Task Launching

- **TL1**-DIF-Selecting tasks which are potentially accessible to all students
- **TL2**-DIF-Making clear the organizational decisions or management procedures for working autonomously on the task
- **TL-DIF-overall**-Holistic estimate of the differentiation during this phase

Student Autonomous Work

- **SAW1**-DIF-Using student asynchronous work to accommodate different learning readiness levels and needs

- **SAW2**-DIF-Encouraging multiple expressions of content, process, and/or product
- **SAW-DIF-overall**-Holistic estimate of the differentiation during this phase

Whole-Class Interactions

- **WCI1**-DIF-Sequencing student solutions in a reasonable progression to support student access to the ideas shared
- **WCI2**-DIF-Highlighting important mathematical ideas during the sharing to ensure that these are made clear to as many students as possible
- **WCI3**-DIF-Students express mathematical ideas that are visible and/or audible to all students (as well as the teacher)
- **WCI-DIF-overall**-Holistic estimate of the differentiation during this phase

PRACTICES THAT PROMOTE THE INTERPLAY OF COGNITIVE ACTIVATION AND DIFFERENTIATION

Task Launching

- **TL1**-IN-Explaining potentially unfamiliar non-mathematical aspects of the wording of the task
- **TL2**-IN-Clarifying mathematical aspects of the task
- **TL3**-IN-Posing questions that indicate the level of support that students need in order to engage in the task
- **TL4**-IN-Activating relevant existing mathematical knowledge and strategies
- **TL-IN-overall**-Holistic estimate of the interplay between the two axes during this phase

Student Autonomous Work

- **SAW1**-IN-Directing different types of questions to different students for engaging them in meaning-making, conceptual connections, or mathematical reasoning
- **SAW2**-IN-Providing enablers to facilitate access to the task at hand without reducing the challenge;
- **SAW3**-IN-Providing extenders to advanced learners or early finishers
- **SAW-IN-overall**-Holistic estimate of the interplay between the two axes during this phase

Whole-Class Interactions

- **WCI1**-IN-Holding students accountable for attending to and understanding their classmates' sharing
- **WCI2**-IN-Using incorrect or incomplete student solutions as resources for all student learning
- **WCI-IN-overall-**Holistic estimate of the interplay between the two axes during this phase.

Appendix 2 – An Example of the Coding Process

A practice promoting Cognitive Activation in the phase of Student Autonomous Work	Scoring
Teacher provides mathematical prompts to students without trivializing the thinking for the students (e.g., by giving directive hints or ready-made answers)	**0:** teacher does not really engage with student work (e.g., teacher circulates around and provides generic comments about student work, such as “Good job!”)
	1: during teacher-student interactions, the teacher provides mathematical prompts that reduce the demands
	2: during teacher-student interactions the teacher provides mathematical prompts without reducing the demands, but the students do not seem to make progress on the task and take up the challenge
	3: during teacher-student interactions the teacher provides mathematical prompts without reducing the demands; the prompts help students make some progress on the task and take up the challenge just

Reference List

Abell, M. M., Jung, E., & Taylor, M. (2011). Students' perceptions of classroom instructional environments in the context of 'Universal Design for Learning'. *Learning Environments Research*, *14*(2), 171-185. https://doi.org/10.1007/s10984-011-9090-2

Adami, A. F. (2004). Enhancing students' learning through differentiated approaches to teaching and learning: A Maltese perspective. *Journal of Research in Special Educational Needs*, *4*(2), 91-97. https://doi.org/10.1111/j.1471-3802.2004.00023.x

Adnyani, L. P. A. P. (2020). Applying cognitive conflict strategy to develop mathematical critical thinking ability and character of students. *Journal of Mathematics Education*, *5*(1), 30-38. http://doi.org/10. 31327/jme.v5i1.1174

Aftab, J. (2015). Teachers' beliefs about differentiated instructions in mixed ability classrooms: A case of time limitations. *Journal of Education and Educational Development, 2*(2), 94 - 114. https://doi.org/10.22555/joeed.v2i2.441

Agathangelou, S. A., Hill, H. C., & Charalambous, C. Y. (2024). Customizing professional development opportunities to teachers' needs: Results from a latent profile analysis. *The Elementary School Journal, 124*(3), 000-000. https://doi.org/10.1086/728590

Akcil-Okan, O., & Tekkumru-Kisa, M. (2021, April 7-10). *Enacting rigorous lessons: Leveraging students' ideas for enhancing demand on student thinking* [Paper presentation]. 94th National Association for Research in Science Teaching Annual Meeting, Virtual conference. https://par.nsf.gov/biblio/10252033

Alicea, S., Buckley, K., Cordova-Cobp, D., Husain, A., Meili, L., Merrill, L., Morales, K., Schmitt, L., Schwatz, N., Tasker, T., Thames, V., & Worthman, S. (2023). *Measuring teacher professional learning: Why it's hard and what we can do about it.* Research Partnership for Professional Learning. https://annenberg.brown.edu/sites/default/files/Measuring%20Teacher%20Professional%20Learning.pdf

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education*, *66*(2), 136-149. https://doi.org/10.1177/0022487114560646

Allen, L., & Turville, J. (2010). *Differentiating by readiness: Strategies and lesson plans for tiered instruction, Grades K-8*. Taylor & Francis.

Alles, M., Seidel, T., & Gröschner, A. (2019). Establishing a positive learning atmosphere and conversation culture in the context of a video-based teacher learning community. *Professional Development in Education, 45*(2), 250-263, https://doi.org/10.1080/19415257.2018.1430049

Alsalamah A, & Callinan C. (2021). Adaptation of Kirkpatrick's four-level model of training criteria to evaluate training programmes for head teachers. *Education Sciences, 11*(3), 1-25. https://doi.org/10.3390/educsci11030116

Altintaş, E., and Özdemir, A. S. (2015). Evaluating a newly developed differentiation approach in terms of student achievement and teachers' opinions. *Educational Sciences: Theory and Practice, 15*(4), 1103–1118. https://doi.org/10.12738/estp.2015.4.2540

Amador, J. M., Wallin, A., Keehr, J., & Chilton, C. (2021). Collective noticing: teachers' experiences and reflection on a mathematics video club. *Mathematics Education Research Journal, 35,* 557-582. https://doi.org/10.1007/s13394-021-00403-9

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261
Anagnostopoulos, D., Cavanna, J., & Charles-Harris, S. (2020). Managing to teach ambitiously in the first year?. *The Elementary School Journal*, *120*(4), 665-691. https://doi.org/10.1086/708660
Anthony, G., Hunter, R., & Hunter, J. (2018). Challenging teachers' perceptions of student capability through professional development: a telling case. *Professional Development in Education, 44*(5), 650-662. https://doi.org/10.1080/19415257.2017.1387868
Anthony, G., Hunter, R., Hunter, J., & Duncan, S. (2015). How ambitious is "ambitious mathematics teaching". *Set: Research Information for Teachers*, *2*, 45-52.
Antoniou, P., & Kyriakides, L. (2013). A dynamic integrated approach to teacher professional development: Impact and sustainability of the effects on improving teacher behaviour and student outcomes. *Teaching and Teacher Education*, *29*, 1-12. https://doi.org/10.1016/j.tate.2012.08.001
Antoniou, P., Kyriakides, L., & Creemers, B. P. M. (2015). The Dynamic Integrated Approach to teacher professional development: rationale and main characteristics. *Teacher Development, 19*(4), 535–552. https://doi.org/10.1080/13664530.2015.1079550
Applebaum, M., & Leikin, R. (2014). Mathematical challenge in the eyes of the beholder: Mathematics teachers' views. *Canadian Journal of Science, Mathematics and Technology Education, 14,* 388-403. https://doi.org/10.1080/14926156.2014.958624
Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40*(3), 685-730. https://doi.org/10.3102/000283120400036
Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: a catalyst for change?. *Journal of Mathematics Teacher Education, 8* , 499-536. https://doi.org/10.1007/s10857-006-6585-3
Arbaugh, F., Lannin, J., Jones, D. L., & Park-Rogers, M. (2006). Examining instructional practices in Core-Plus lessons: Implications for professional development. *Journal of Mathematics Teacher Education, 9* , 517-550. https://doi.org/10.1007/s10857-006-9019-3
Arizona Group, Guilfoyle, K., Hamilton, M. L., Pinnegar, S., & Placier, M. (1996). Negotiating balance between reforming teacher education and forming self as teacher educator. *Teacher Education Quarterly, 23*(3), 153–168. http://www.jstor.org/stable/23477793
Askey, R. (2001). Good intentions are not enough. In. T. Loveless (Ed.), *The great curriculum debate* (pp. 163–183). Brookings Institution.
Avidov-Ungar, O. (2016). A model of professional development: Teachers' perceptions of their professional development. *Teachers and Teaching*, *22*(6), 653-669. https://doi.org/10.1080/13540602.2016.1158955
Bal, A. P. (2016). The Effect of the differentiated teaching approach in the algebraic learning field on students' academic achievements. *Eurasian Journal of Educational Research*, *63*, 185-204. http://dx.doi.org/ 10.14689/ejer.2016.63.11
Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal, 90*(4), 449–466. https://doi.org/10.1086/461626

Ball, D. L. (2003). Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education. RAND Corporation. https://www.rand.org/pubs/monograph_reports/MR1643.html. Also available in print form.
Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond and G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3–31). Jossey-Bass.
Ball, D. L., Hill, H. H., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? A*merican Educator*, *29*(1), 14-46. http://hdl.handle.net/2027.42/65072
Barnhart, T., & van Es, E. A. (2020). Developing a critical discourse about teaching and learning: The case of a secondary science video club. *Journal of Science Teacher Education, 31*(5), 491-514. https://doi.org/10.1080/1046560X.2020.1725724
Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133-180. https://doi.org/10.3102/0002831209345157
Baumgartner, T., Lipowski, M. B., & Rush, C. (2003). *Increasing reading achievement of primary and middle school students through differentiate instruction* [Master dissertation]. Saint Xavier University.
Bayar, A. (2014). The components of effective professional development activities in terms of teachers' perspective. *International Online Journal of Educational Sciences, 6*(2), 319-327. http://dx.doi.org/10.15345/iojes.2014.02.006Beecher, M., & Sweeny, S. M. (2008). Closing the achievement gap with curriculum enrichment and differentiation: one school's story. *Journal of Advanced Academics*, *19*(3), 502-530. https://doi.org/10.4219/jaa-2008-815
Beisiegel, M., Mitchell, R., & Hill, H. C. (2018). The design of video-based professional development: an exploratory experiment intended to identify effective features. *Journal of Teacher Education, 69*(1), 69-89. https://doi.org/10.1177/0022487117705096
Beltramo, J. (2017). Developing adaptive teaching practices through participation in cogenerative dialogues. *Teaching and Teacher Education, 63*, 326-337. https://doi.org/10.1016/j.tate.2017.01.007
Birnie, B. F. (2015). Making the case for differentiation. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 88*(2), 62-65. https://doi.org/10.1080/00098655.2014.998601
Blazar, D. (2015). Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review*, *48*, 16-29. https://doi.org/10.1016/j.econedurev.2015.05.005
Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational evaluation and policy analysis, 37*(4), 542-566. https://doi.org/10.3102/0162373715579487
Blomberg, G., Sherin, M. G., Renkl, A., Glogger, I., & Seidel, T. (2014). Understanding video as a tool for teacher education: investigating instructional strategies to promote reflection. *Instructional Science, 42* , 443-463. https://doi.org/10.1007/s11251-013-9281-6
Boaler, J. (2002). Learning from teaching: Exploring the relationship between reform curriculum and equity. *Journal for research in mathematics education, 33*(4), 239-258. https://doi.org/10.2307/749740

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record, 110*(3), 608–645. https://doi.org/10.1177/016146810811000030
Bobis, J.M., Way, J., Anderson, J., & Martin, A.J. (2016). Challenging teacher beliefs about student engagement in mathematics. *Journal of Mathematics Teacher Education, 19*, 33-55. https://doi.org/10.1007/s10857-015-9300-4
Bolhuis, S. (2003). Towards process-oriented teaching for self-directed lifelong learning: a multidimensional perspective. *Learning and Instruction, 13*(3), 327-347. https://doi.org/10.1016/S0959-4752(02)00008-7
Bondie, R. S., Dahnke, C., & Zusho, A. (2019). How does changing “one-size-fits-all” to differentiated instruction affect teaching?. *Review of Research in Education, 43*(1), 336-362. https://doi.org/10.3102/0091732X18821130
Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, *33*(8), 3-15. https://doi.org/10.3102/0013189X033008003
Borko, H., Eisenhart, M., Brown, C. A., Underhill, R. G., Jones, D., & Agard, P. (1992). Learning to teach hard mathematics: Do novice teachers and their instructors give up too easily? *Journal for Research in Mathematics Education, 23*(3), 194–222. https://doi.org/10.5951/jresematheduc.23.3.0194
Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. *International Encyclopedia of Education, 7*(2), 548-556. https://doi.org/10.1016/B978-0-08-044894-7.00654-0
Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and teacher education*, *24*(2), 417-436. https://doi.org/10.1016/j.tate.2006.11.012
Borko, H., Jacobs, J., Seago, N., & Mangram, C. (2014). Facilitating video-based professional development: planning and orchestrating productive discussions. In Y. Li, E. A. Silver, & S. Li (Eds.), *Transforming mathematics instruction: Multiple approaches and practices* (pp. 259–281). Springer. https://doi.org/10.1007/978-3-319-04993-9_16
Borko, H., Koellner, K., Jacobs, J., & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM Mathematics Education, 43*(1), 175-187. https://doi.org/10.1007/s11858-010-0302-5
Boston, M. (2012). Assessing instructional quality in mathematics. *The Elementary School Journal, 113*(1), 76-104. https://doi.org/10.1086/666387
Boston, M. (2013). Connecting changes in secondary mathematics teachers’ knowledge to their experiences in a professional development workshop. *Journal of Mathematics Teacher Education, 16*(1), 7–31. https://doi.org/10.1007/s10857-012-9211-6
Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119-156. https://doi.org/10.2307/40539329
Boston, M. D., & Smith, M. S. (2011). A ‘task-centric approach’ to professional development: Enhancing and sustaining mathematics teachers’ ability to implement cognitively challenging mathematical tasks. *ZDM Mathematics Education, 43*(6-7), 965-977. https://doi.org/10.1007/s11858-011-0353-2
Boston, M. D., & Wilhelm, A. G. (2017). Middle school mathematics instruction in instructionally focused urban districts. *Urban Education*, *52*(7), 829-861. https://doi.org/10.1177/0042085915574528

Boylan, M., Coldwell, M., Maxwell, B., & Jordan, J. (2018). Rethinking models of professional learning as tools: a conceptual analysis to inform research and practice. *Professional Development in Education, 44*(1), 120–139. https://doi.org/10.1080/19415257.2017.1306789

Brändström, A. (2005). *Differentiated tasks in mathematics textbooks. An analysis of the levels of difficulty.* [Licentiate dissertation]. Luleå Tekniska Universitet. https://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-18110

Brantlinger, A., Sherin, M. G., & Linsenmeier, K. A. (2011). Discussing discussion: a video club in the service of math teachers' National Board preparation. *Teachers and Teaching: theory and practice*, *17*(1), 5-33. https://doi.org/10.1080/13540602.2011.538494

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative *Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brighton, C., Hertberg, H., Moon, T., Tomlinson, C., & Callahan, C. (2005). *Feasibility of high-end learning in the academically diverse middle school*. The National Research Center on the Gifted and Talented, University of Virginia.Brimijoin, K. (2005). Differentiation and high-stakes testing: An oxymoron? *Theory into Practice, 44*(3), 254-261. https://doi.org/10.1207/s15430421tip4403_10

Brophy, J. (Ed.). (2004). *Using video in teacher education*. Elsevier.

Bruner, J. S. (1960). *The process of education*. Harvard University Press.

Bullman, A. (2021). *Teacher autonomy in professional development selection* [Doctoral Dissertation]. Baker University. http://www.bakeru.edu/images/pdf/SOE/EdD_Theses/Bullman_April.pdf

Büscher, C. (2019, February). Conceptual learning opportunities in teachers' differentiated task designs for inclusive mathematics education. In U. T. Jankvist, M. van den Heuvel-Panhuizen & M. Veldhuis (Eds.), *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education* (pp. 3604-3611). Freudenthal Group und ERME.

Cambridge Dictionary. (n.d.). Citation. In *dictionary.cambridge.org dictionary*. https://dictionary.cambridge.org/dictionary/english

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal, 111*(3), 430-454. https://doi.org/10.1086/657654

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2019). Malleability, plasticity, and individuality: How children learn and develop in context. *Applied Developmental Science, 23*(4), 307-337. https://doi.org/10.1080/10888691.2017.1398649

Carolan, J., & Guinn, A. (2007). Differentiation: lessons from master teachers. *Educational Leadership*, *64*(5), 44-47. https://www.ascd.org/el/articles/differentiation-lessons-from-master-teachers

Carpendale, J., Berry, A., Cooper, R., & Mitchell, I. (2021). Balancing fidelity with agency: understanding the professional development of highly accomplished teachers. *Professional Development in Education*, 1–19. https://doi.org/10.1080/19415257.2021.1972436

Carpenter, T. P., & Fennema, E. (1992). Cognitively guided instruction: Building on the knowledge of students and teachers. *International Journal of Educational Research*, *17*(5), 457-470. https://doi.org/10.1016/S0883-0355(05)80005-9

Carson, R. (2010). *High school mathematics teacher' thinking regarding exploratory learning activities* [Master Dissertation]. University of Calgary. https://doi.org/10.11575/PRISM/3877

Cazden, C.B. (2001). The language of teaching and learning (2nd ed.). Heinemann.

Chapman, O. (2013). Mathematical-task knowledge for teaching. *Journal of Mathematics Teacher Education*, *16*(1), 1-6. https://doi.org/10.1007/s10857-013-9234-7

Charalambous, C. Y. (2008). Mathematical knowledge for teaching and the unfolding of tasks in mathematics lessons: Integrating two lines of research. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano, & A. Sepulveda (Eds.), *Proceedings of the 32nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 281-288). PME. https://doi.org/10.1080/07370000802177235

Charalambous, C. Y. (2010). Mathematical knowledge for teaching and task unfolding: An exploratory study. *The Elementary School Journal, 110*(3), 247-278. https://doi.org/10.1086/648978

Charalambous, C. Y., & Hill, H. C. (2012). Teacher knowledge, curriculum materials, and quality of instruction: Unpacking a complex relationship. *Journal of Curriculum Studies*, *44*(4), 443-466. https://doi.org/10.1080/00220272.2011.650215

Charalambous, C. Y., & Praetorius, A. K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation, 67*(1), 1-8. https://doi.org/10.1016/j.stueduc.2020.100894

Charalambous, C. Y., Agathangelou, S. A., Kasapi, E., & Christofidou, E. (2023a). Learning to teach ambitiously: a multiple case study of practicing teachers' experimentation with enablers and extenders. *Journal of Mathematics Teacher Education, 26*, 363-394. https://doi.org/10.1007/s10857-022-09532-9

Charalambous, C. Y., Hill, H. C., & Mitchell, R. N. (2012). Two negatives don't always make a positive: Exploring how limitations in teacher knowledge and the curriculum contribute to instructional quality. *Journal of Curriculum Studies*, *44*(4), 489-513. https://doi.org/10.1080/00220272.2012.716974

Charalambous, C. Y., Philippou, S., & Olympiou, G. (2018). Reconsidering the use of video clubs for student-teachers' learning during field placement: Lessons drawn from a longitudinal multiple case study. *Teaching and Teacher Education, 74*, 49-61. https://doi.org/10.1016/j.tate.2018.04.002

Charalambous, C. Y., Philippou, S., Olympiou, G., & Georgiou, K. (2022). Experimenting with enablers and extenders to support ambitious teaching in mathematics: A video-club case study of student teachers during their field placement. *Teaching and Teacher Education, 119*, 1-19. https://doi.org/10.1016/j.tate.2022.103874

Charalambous, C.Y., Agathangelou, S., Delaney, S., Papadouris, N. (2023b). Engaging all students in challenging mathematical work: Working at the intersection of cognitively challenging tasks and differentiation during lesson planning and enactment. In J. Cai, G. J. Stylianides, & P. A. Kenney. (Eds.), *Research studies on learning and teaching of mathematics: Dedicated to Edward A. Silver* (pp. 179-218). Springer. https://doi.org/10.1007/978-3-031-35459-5_9

Cheeseman, J., Clarke, D., Roche, A., & Walker, N. (2016). Introducing challenging tasks: Inviting and clarifying without explaining and demonstrating. *Australian Primary Mathematics Classroom, 21*(3), 3-7. https://search.informit.org/doi/10.3316/informit.350316936851015

Cheeseman, J., Clarke, D.M., Roche, A., & Wilson, K. (2013). Teachers' views of the challenging elements of a task. In V. Steinle, L. Ball & C. Bardini (Eds.), *Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia* (pp. 154–161). MERGA. http://www.merga.net.au/documents/Cheeseman_&_Clarke_MERGA36-2013.pdf

Choppin, J. (2011). The impact of professional noticing on teachers' adaptations of challenging tasks. *Mathematical Thinking and Learning, 13*(3), 175-197. https://doi.org/10.1080/10986065.2010.495049

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and teacher education*, *18*(8), 947-967. https://doi.org/10.1016/S0742-051X(02)00053-7
Clarke, D., Cheeseman, J., Roche, A., & van der Schans, S. (2014a). Teaching strategies for building student persistence on challenging tasks: insights emerging from two approaches to teacher professional learning. *Mathematics Teacher Education and Development, 16*(2), 46-70. https://files.eric.ed.gov/fulltext/EJ1052607.pdf
Clarke, D., Roche, A., Cheeseman, J., & Sullivan, P. (2014b). Encouraging students to persist when working on challenging tasks: Some insights from teachers. *Australian Mathematics Teacher, 70*(1), 3-11. https://search.informit.org/doi/10.3316/informit.231468585575226
Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational evaluation and policy analysis, 12*(3), 311-329. https://doi.org/10.3102/01623737012003311
Cohen, D. K. (2011). *Teaching: Practice and its predicaments*. Harvard University Press.
Cohen, J., Hutt, E., Berlin, R., & Wiseman, E. (2022). The change we cannot see: Instructional quality and classroom observation in the era of common core. *Educational Policy, 36*(6), 1261-1287. https://doi.org/10.1177/0895904820951114
Cordingley, P. (2015). The contribution of research to teachers' professional learning and development. *Oxford Review of Education, 41*(2), 234-252. https://doi.org/10.1080/03054985.2015.1020105
Cordingley, P., Bell, M., Isham, C., Evans, D., & Firth, A. (2007). *Continuing professional development (CPD): What do specialists do in CPD programmes for which there is evidence of positive outcomes for pupils and teachers?* EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/CPD4%20Report%20-%20SCREEN.pdf?ver=2007-09-28-142054-167
Creemers, B. P. M., Kyriakides, L., & Antoniou, P. (2013). *Teacher professional development for improving quality in teaching*. Springer. https://doi.org/10.1007/978-94-007-5207-8
Creemers, B. P., & Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement*, *17*(3), 347-366. https://doi.org/10.1080/09243450600697242
Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage Publications Inc. https://toc.library.ethz.ch/objects/pdf/z01_978-1-4129-7517-9_01.pdf
Croninger, R., Lareson, J. C., & VonSecker, C. E. (2006, April). *Effects of teacher qualifications, practices, and content on 4th and 5th grade mathematics achievement in high- and low-poverty classes*. Paper presented at the Annual Meeting of the American Educational Research Association. https://www.researchgate.net/profile/Robert-Croninger/publication/240620333_Effects_of_Teacher_Qualifications_Practices_and_Content_on_4th_and_5th_Grade_Mathematics_Achievement_in_High_and_Low-Poverty_Classes1/links/552524c20cf22e181e73e127/Effects-of-Teacher-Qualifications-Practices-and-Content-on-4th-and-5th-Grade-Mathematics-Achievement-in-High-and-Low-Poverty-Classes1.pdf
Dack, H. (2018). Structuring teacher candidate learning about differentiated instruction through coursework. *Teaching and Teacher Education, 69*, 62–74. https://doi.org/10.1016/j.tate.2017.09.017

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24*(2), 97–140. https://doi.org/10.1080/10888691.2018.1537791

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development.* Learning Policy Institute. https://bibliotecadigital.mineduc.cl/bitstream/handle/20.500.12365/17357/46%20Effective_Teacher_Professional_Development_REPORT.pdf?sequence=1

Dash, S., Magidin de Kramer, R., O'Dwyer, L. M., Masters, J., & Russell, M. (2012). Impact of online professional development or teacher quality and student achievement in fifth grade mathematics. *Journal of research on technology in education*, *45*(1), 1-26. https://doi.org/10.1080/15391523.2012.10782595

de Jager, T. (2017). Perspectives of teachers on differentiated teaching in multi-cultural South African secondary schools. *Studies in Educational Evaluation, 53*, 115–121. https://doi.org/10.1016/j.stueduc.2016.08.004

Delaney, S. (2017). *Become the primary teacher everyone wants to have: A guide to career success.* Routledge.

Delaney, S., & Gurhy, A. M. (2019, February). Combining differentiation and challenge in mathematics instruction: A case from practice. In U. T. Jankvist, M. Van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Proceedings of the eleventh congress of the European Society for Research in mathematics education*. Freudenthal Group & Freudenthal Institute, Utrecht University and ERME. https://hal.science/hal-02423365/

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199. https://doi.org/10.3102/0013189X08331140

Desimone, L. M. (2011). A primer on effective professional development. *Phi Delta Kappan*, *92*(6), 68-71. https://doi.org/10.1177/003172171109200616

Desimone, L. M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory Into Practice*, *56*(1), 3-12. https://doi.org/10.1080/00405841.2016.1241947

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81-112. https://doi.org/10.3102/01623737024002081

Desimone, L., Smith, T. A., & Phillips, K. (2013). Linking student achievement growth to professional development participation and changes in instruction: A longitudinal study of elementary students and teachers in Title I schools. *Teachers College Record, 111*(5), 1–46. https://doi.org/10.1177/016146811311500508

Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review, 24*, 31-54. https://doi.org/10.1016/j.edurev.2018.02.002

Dijkstra, E. M., Walraven, A., Mooij, T., & Kirschner, P. A. (2017). Factors affecting intervention fidelity of differentiated instruction in kindergarten. *Research Papers in Education*, *32*(2), 151-169. https://doi.org/10.1080/02671522.2016.1158856

Dixon, F. A., Yssel, N., McConnell, J. M., & Hardin, T. (2014). Differentiated instruction, professional development, and teacher efficacy. *Journal for the Education of the Gifted, 37*(2), 111-127. https://doi.org/10.1177/0162353214529042

Dobie, T. E., & Anderson, E. R. (2015). Interaction in teacher communities: Three forms teachers use to express contrasting ideas in video clubs. *Teaching and Teacher Education, 47*, 230-240. https://doi.org/10.1016/j.tate.2015.01.003

Dooley, K. (2009). Intercultural conversation: Building understanding together. *Journal of Adolescent and Adult Literacy*, *52*(6), 497-506. https://eprints.qut.edu.au/28342/1/c28342.pdf

Dotger, S., & Causton-Theoharis, J. (2010). Differentiation through choice using a think-tac-toe for science content. *Science Scope, 33(*6), 18. https://osu-wams-blogs-uploads.s3.amazonaws.com/blogs.dir/548/files/2010/11/Dotger-2010-Think-Tac-Toe.pdf

Doyle, W. (1983). Academic work. *Review of educational research, 53*(2), 159-199. https://doi.org/10.1080/0022027860180402

Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*(2), 167-180. https://doi.org/10.1207/s15326985ep2302_6

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, *20*(4), 399-483. https://doi.org/10.1207/S1532690XCI2004_1

Erotocritou-Stavrou, T. E., & Koutselini, M. (2016). Differentiation of teaching and learning: The Teachers' perspective. *Universal Journal of Educational Research, 4*(11), 2581-2588. https://doi.org/10.13189/ujer.2016.041111

European Commission. (2013). Supporting teacher competence development for better learning outcomes. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2012:0374:FIN:EN:PDF

Evans, L. (2014). Leadership for professional development and learning: enhancing our understanding of how teachers develop. *Cambridge Journal of Education, 44*(2), 179-198. https://doi.org/10.1080/0305764X.2013.860083

Evers, W. J., Brouwers, A., & Tomic, W. (2002). Burnout and self-efficacy: A study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of educational psychology*, *72*(2), 227-243. https://doi.org/10.1348/000709902158865

Fauskanger, J., & Bjuland, R. (2019). Learning ambitious teaching of multiplicative properties through a cycle of enactment and investigation. *Mathematics Teacher Education and Development*, *21*(1), 125-144. https://files.eric.ed.gov/fulltext/EJ1216014.pdf

Fernandez, C., & Yoshida, M. (2004). *Lesson study: A case of a Japanese approach to improving instruction through school-based teacher development.* Lawrence Erlbaum. https://doi.org/10.4324/9781410610867

Firmender, J. M., Reis, S. M., & Sweeny, S. M. (2013). Reading comprehension and fluency levels ranges across diverse classrooms: The need for differentiated reading instruction and content. *Gifted Child Quarterly*, *57*(1), 3-14. https://doi.org/10.1177/0016986212460084

Foley, D. J., Khoshaim, H. B., Alsaeed, M., & Er, N. S. (2012). Professional development in statistics, technology, and cognitively demanding tasks: Classroom implementation and obstacles. *International Journal of Mathematical Education in Science and Technology*, *43*(2), 177-196. https://doi.org/10.1080/0020739X.2011.592616

Förtsch, C., Werner, S., von Kotzebue, L., & Neuhaus, B. J. (2016). Effects of biology teachers' professional knowledge and cognitive activation on students' achievement. *International Journal of Science Education*, *38*(17), 2642-2666. https://doi.org/10.1080/09500693.2016.1257170

Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American educational research journal, 38*(3), 653-689. https://doi.org/10.3102/00028312038003653

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F.K. Lester, Jr., (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225-256). Information Age Publishing.

Fullan, M. (2007). Change the terms for teacher learning. *Journal of Staff Development, 28*(3), 35-36. https://www.proquest.com/openview/bd8bc2338e4b16812a90e9516585faf2/1?pq-origsite=gscholar&cbl=47961

Gaitas, S., Carêto, C., Peixoto, F., & Castro Silva, J. (2022). Differentiated instruction:'to be, or not to be, that is the question'. *International Journal of Inclusive Education,* 1-17. https://doi.org/10.1080/13603116.2022.2119290

Gallagher, H. A., Arshan, N., & Woodworth, K. (2017). Impact of the National Writing Project's college-ready writers' program in high-need rural districts. *Journal of Research on Educational Effectiveness*, *10*(3), 570-595. https://doi.org/10.1080/19345747.2017.1300361

Gallagher, M. A., Parsons, S. A., & Vaughn, M. (2020). Adaptive teaching in mathematics: a review of the literature. *Educational Review*, *74*(2), 1-23. https://doi.org/10.1080/00131911.2020.1722065

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945. https://doi.org/10.3102/00028312038004915

Garrett, R., Citkowicz, M., & Williams, R. (2019). How responsive is a teacher's classroom practice to intervention? A meta-analysis of randomized field studies. *Review of research in education, 43*(1), 106-137. https://doi.org/10.3102/0091732X19830634

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review, 16*, 41-67. https://doi.org/10.1016/j.edurev.2015.06.001

Geisler, J. L., Hessler, T., Gardner, R., & Lovelace, T. S. (2009). Differentiated writing interventions for high-achieving urban African American elementary students. *Journal of Advanced Academics*, *20*(2), 214-247. https://doi.org/10.1177/1932202X0902000202

Gheyssens, E, Consuegra, E, & Struyven, K (2020). Good things come to those who wait: The importance of professional development for the implementation of differentiated instruction. *Frontiers in Education, 5*(96). https://doi.org/10.3389/feduc.2020.00096

Ghousseini, H., & Kazemi, E. (2023). Professional Learning Tasks Through Job-Embedded Teacher Professional Development. In Cai, J., Stylianides, G.J., Kenney, P.A. (eds) *Research Studies on Learning and Teaching of Mathematics: Dedicated to Edward A. Silver* (pp. 27-48). Springer International Publishing. https://doi.org/10.1007/978-3-031-35459-5_2

Gibbons, L. & Okun, A. (2023). Examining a coaching routine to support teacher learning. *Investigations in Mathematics Learning, 15*(1), 11-28. https://doi.org/10.1080/19477503.2022.2139094

Gibbons, L., Kazemi, E., Hintz, A., & Hartmann, E. (2017). Teacher time out: Educators learning together in and through practice. *NCSM Journal, 18*(2), 28-46. https://www.mathedleadership.org/docs/resources/journals/NCSMJournalVol18Num2.pdf#page=32

Gibbs, K., & McKay, L. (2021). Differentiated teaching practices of Australian mainstream classroom teachers: A systematic review and thematic analysis. International *Journal of Educational Research, 109*, 1-9. https://doi.org/10.1016/j.ijer.2021.101799

Gibson, V., & Hasbrouck, J. (2008). *Differentiated instruction: Grouping for success*. McMillan-McGraw.

Gilbert, M. C. (2016). Relating aspects of motivation to facets of mathematical competence varying in cognitive demand. *The Journal of Educational Research, 109*(6), 647-657. https://doi.org/10.1080/00220671.2015.1020912

Goddard, Y., Goddard, R., & Kim, M. (2015). School instructional climate and student achievement: An examination of group norms for differentiated instruction. *American Journal of Education*, *122*(1), 111-131. https://doi.org/10.1086/683293

Goldsmith, L. T., Doerr, H. M., & Lewis, C. C. (2014). Mathematics teachers' learning: A conceptual framework and synthesis of research. *Journal of Mathematics Teacher Education, 17*(1), 5-36. https://doi.org/10.1007/s10857-013-9245-4

Grant, S. G., & Gradwell, J. (2009). The road to ambitious teaching: Creating big idea units in history classes. *Journal of Inquiry and Action in Education*, *2*(1), 1-26. https://digitalcommons.buffalostate.edu/jiae/vol2/iss1/1/

Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational psychology review, 15*(2), 147-179. https://doi.org/10.1023/A:1023477131081

Guay, F., Roy, A., & Valois, P. (2017). Teacher structure as a predictor of students' perceived competence and autonomous motivation: The moderating role of differentiated instruction. *British Journal of Educational Psychology, 87*(2), 224-240. https://doi.org/10.1111/bjep.12146

Guild, P.B., and Garger, S (1998). *What is differentiated instruction? Marching to different drummers* (2nd Ed.). Association for Supervision & Curriculum. https://files.eric.ed.gov/fulltext/ED269839.pdf

Guskey, T. R. (2000). *Evaluating professional development.* Corwin Press.

Guskey, T. R. (2003). What makes professional development effective?. *Phi Delta Kappan*, *84*(10), 748-750. https://doi.org/10.1177/003172170308401007

Guskey, T. R. (2024). Look beyond the satisfaction survey: A framework to evaluate results of professional learning. *The Learning Professional, 45*(1), 28-33. https://tguskey.com/wp-content/uploads/LP-24-Prof-Lrng-Eval-PDF.pdf

Guskey, T. R., & Yoon, K. S. (2009). What works in professional development?. *Phi Delta Kappan*, *90*(7), 495-500. https://doi.org/10.1177/003172170909000709

Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal*, *28*(4), 567–590. https://doi.org/10.1080/0141192022000005832

Hall, T. (2002). Differentiated instruction. National Center on Accessing the General Curriculum.

Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 U.S. early childhood and elementary classrooms*. Foundation for Child Development. https://www.fcdus.org/wp-content/uploads/2016/04/BuildingAScienceOfClassroomsPiantaHamre.pdf

Hardré, P. L., & Sullivan, D. W. (2008). Teacher perceptions and individual differences: How they influence rural teachers' motivating strategies. *Journal of Teaching and Teacher Education, 4*(7), 1-17. https://doi.org/10.1016/j.tate.2008.04.007

Harris, D. M. (2012). Varying teacher expectations and standards: Curriculum differentiation in the age of standards-based reform. *Education and Urban Society, 44*(2), 128-150. https://doi.org/10.1177/0013124511431568

Hathcock, S. J. Dickerson, D., Eckhoff, A., & Katsioloudis, P. (2015). Scaffolding for creative product possibilities in a design-based STEM activity. *Research in Science Education, 45*(5), 727–748. https://doi.org/10.1007/s11165-014-9437-7

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. https://helpingchildrenwithdyslexia.com/gintest/wp-content/uploads/2019/10/PDF-5.pdf

Haug, B. S., & Mork, S. M. (2021). Taking 21st century skills from vision to classroom: What teachers highlight as supportive professional development in the light of new demands from educational reforms. *Teaching and teacher education, 100*, 1-12. https://doi.org/10.1016/j.tate.2021.103286

Hawley, W.D., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp.127-150). Jossey-Bass.

Hedrick, K. A. (2012). Differentiation: A strategic response to student needs. *The Education Digest, 78*(4), 31-36. https://www.proquest.com/openview/f2a7f4bbfe2c4296afa34e42d6578c12/1.pdf?pq-origsite=gscholar&cbl=25066

Helsing, D., Howell, A., Kegan, R., & Lahey, L. (2008). Putting the" development" in professional development: Understanding and overturning educational leaders' immunities to change. *Harvard educational review, 78*(3), 437-465. https://doi.org/10.17763/haer.78.3.888l759g1qm54660

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for research in mathematics education, 28*(5), 524-549. https://doi.org/10.5951/jresematheduc.28.5.0524

Herbst, P. G. (2006). Teaching geometry with problems: Negotiating instructional situations and mathematical tasks. *Journal for Research in Mathematics Education, 37*(4), 313-347. https://doi.org/10.2307/30034853

Herbst, P., Fujita, T., Halverscheid, S., & Weiss, M. (2017). *The learning and teaching of geometry in secondary schools: A modeling perspective*. Routledge. https://www.routledge.com/The-Learning-and-Teaching-of-Geometry-in-Secondary-Schools-A-Modeling-Perspective/Herbst-Fujita-Halverscheid-Weiss/p/book/9780415856911

Hertberg-Davis, H. (2009). Myth 7: Differentiation in the regular classroom is equivalent to gifted programs and is sufficient: Classroom teachers have the time, the skill, and the will to differentiate adequately. *Gifted Child Quarterly*, *53*(4), 251-253. https://doi.org/10.1177/0016986209346927

Hertberg-Davis, H. L., & Brighton, C. M. (2006). Support and sabotage principals' influence on middle school teachers' responses to differentiation. *Journal of Secondary Gifted Education, 17*, 90–102. https://doi.org/10.4219/jsge-2006-685

Heyd-Metzuyanim, E., Nachlieli, T., Weingarden, M., & Baor, R. (2020). Adapting a professional development program for cognitively demanding instruction across

shifting contexts. *Educational Studies in Mathematics, 104*, 385-403. https://doi.org/10.1007/s10649-020-09967-y
Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: a project of the National Council of Teachers of Mathematics* (pp. 371-404). Information Age Publishing.
Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, *30*(2), 393-425. https://doi.org/10.3102/00028312030002393
Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H., Smith, M., ... & Gallimore, R. (2005). Mathematics teaching in the United States today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis*, *27*(2), 111-132. https://doi.org/10.3102/01623737027002111
Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher, 42*(9), 476–487. https://doi.org/10.3102/0013189X13512674
Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and instruction, 26*(4), 430-511. https://doi.org/10.1080/07370000802177235
Hill, H. C., Corey, D. L., & Jacobs, R. (2018). Dividing by zero: Exploring null results in a mathematics professional development program. *Teachers College Record, 120*(6), 1–42. https://doi.org/10.1177/016146811812000602
Hill, H., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406. https://doi.org/10.3102/00028312042002371
Hill, J., Hawk, K., & Taylor, K. (2002). Professional development: What makes it work. *SET: Research Information for Teachers*, (2), 12-15. https://doi.org/10.18296/set.0739
Hlas, A. C., & Hlas, C. S. (2012). A review of high-leverage teaching practices: Making connections between mathematics and foreign languages. *Foreign Language Annals, 45*(1), 76-97. https://doi.org/10.1111/j.1944-9720.2012.01180.x
Hollingsworth, H., Lokan, J., & McCrae, B. (2003). *Teaching mathematics in Australia: Results from the TIMSS video study* (TIMSS Australia Monograph No. 5). Australian Council for Educational Research. https://research.acer.edu.au/cgi/viewcontent.cgi?article=1003&context=timss_video
Horn, I., & Garner, B. (2022). *Teacher learning of ambitious and equitable mathematics instruction: A sociocultural approach*. Routledge. https://doi.org/10.4324/9781003182214
Hsu, H. Y., & Yao, C. Y. (2023). A Review of the mathematical tasks framework and levels of cognitive demand. In J. Cai, G. J. Stylianides, & P. A. Kenney. (Eds.), *research studies on learning and teaching of mathematics: Dedicated to Edward A. Silver*, (pp. 219-252). https://doi.org/10.1007/978-3-031-35459-5_10
http://dx.doi.org/10.18296/set.0017
Hugener, I., Pauli, C., Reusser, K., Lipowsky, F., Rakoczy, K., & Klieme, E. (2009). Teaching patterns and learning quality in Swiss and German mathematics lessons. *Learning and Instruction*, *19*(1), 66-78. https://doi.org/10.1016/j.learninstruc.2008.02.001
Hunter, R. (2008). Facilitating communities of mathematical inquiry. In M. Goos, R. Brown, & K. Makar (Eds.). *Navigating currents and charting directions: Proceedings*

of the 31st annual conference of the Mathematics Education Research Group of Australasia (Vol. 1, pp. 31–39). MERGA. https://files.eric.ed.gov/fulltext/ED503747.pdf

Ingram, N., Holmes, M., Linsell, C., Livy, S., & Sullivan, P. (2016). Teacher actions that encourage students to persist in solving challenging mathematical tasks. In B. White, M. Chinnappan & S. Trenholm (Eds.), *Opening up mathematics education research- Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia* (pp. 657-660).MERGA. https://files.eric.ed.gov/fulltext/ED572389.pdf

Ingvarson, L., Meiers, M. & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives, 13*(10), 1-28. https://doi.org/10.14507/epaa.v13n10.2005

Jackson, K. J., Shahan, E. C., Gibbons, L. K., & Cobb, P. A. (2012). Launching complex tasks. *Mathematics Teaching in the Middle School, 18*(1), 24-29. https://doi.org/10.5951/mathteacmiddscho.18.1.0024

Jackson, K., & Cobb, P. (2010, May). *Refining a vision of ambitious mathematics instruction to address issues of equity* [Paper presentation]. Presented at the Annual Meeting of the American Educational Research Association. https://cadrek12.org/sites/default/files/Jackson%20%20Cobb%20Ambitious%20and%20equitable_vision_100702.pdf

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education*, *44*(4), 646-682. https://doi.org/10.5951/jresematheduc.44.4.0646

Jackson, K., Gibbons, L., & Dunlap, C. (2014). Teachers' views of students' mathematical capabilities: A challenge for accomplishing ambitious reform. *Teachers College Record, 119*(7), 1-43. https://doi.org/10.1177/016146811711900708

Jacob, R., Hill, H., & Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness, 10*(2), 379-407. https://doi.org/10.1080/19345747.2016.1273411

Jakopovic, P.M. (2021). Coaching to develop teacher professional noticing: planning with students and mathematics in mind. *International Journal of Mentoring and Coaching in Education, 10*(3), 339-354. https://doi.org/10.1108/IJMCE-10-2020-0064

Johnsen, S. (2003). Adapting instruction with heterogeneous groups. *Gifted Child Today, 26*(3), 5–6. https://doi.org/10.1177/107621750302600302

Jones, R. E., Yssel, N., & Grant, C. (2012). Reading instruction in tier 1: Bridging the gaps by nesting evidence-based interventions within differentiated instruction. *Psychology in the Schools*, *49*(3), 210-218. https://doi.org/10.1002/pits.21591

Jordan, A., Schwartz, E., & McGhie-Richmond, D. (2009). Preparing teachers for inclusive classrooms. *Teaching and Teacher Education*, *25*(4), 535-542. https://doi.org/10.1016/j.tate.2009.02.010

Jürimäe, M., Kärner, A. and Tiisvelt, L. (2014). Teachers Taking Ownership of Educational Change via Participation in Professional Learning Communities. In: Nyhamn, F. & Hopfenbeck, T. N. (Eds.). *From political decisions to change in the classroom: successful implementation of classroom policy* (pp. 58-77). The Norwegian Directorate for Education and Training.

https://www.researchgate.net/profile/Anita-Kaerner/publication/281064266_Teachers_Taking_Ownership_of_Educational_Change_via_Participation_in_Professional_Learning_Communities/links/55d322b108ae0a3417225da4/Teachers-Taking-Ownership-of-Educational-Change-via-Participation-in-Professional-Learning-Communities.pdf

Kane, T. J., and D. O. Staiger. 2012. *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Measures of Effective Teaching Project, Bill & Melinda Gates Foundation. https://files.eric.ed.gov/fulltext/ED540960.pdf

Kanellopoulou, E. M., & Darra, M. (2022). Pedagogical differentiation in primary education: conceptual determinants and definitions. *International Education Studies, 15*(2), 138-148. https://doi.org/10.5539/ies.v15n2p138

Kang, N. H. (2017). Korean teachers' conceptions of models and modeling in science and science teaching. *Journal of the Korean Association for Science Education*, *37*(1), 143-154. https://doi.org/10.14697/jkase.2017.37.1.0143

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: attending to the coevolution of teachers' participation across contexts. *Journal of Teacher Education, 59*(5), 428-441. https://doi.org/10.1177/0022487108324330

Kazemi, E., Franke, M., & Lampert, M. (2009, July). Developing pedagogies in teacher education to support novice teachers' ability to enact ambitious instruction. In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 12-30). https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3491651526c45d41a06d38d83848f0ee0b57da2d

Kelcey, B., Hill, H. C., & Chin, M. J. (2019). Teacher mathematical knowledge, instructional quality, and student outcomes: a multilevel quantile mediation analysis. *School Effectiveness and School Improvement, 30*(4), 398-431. https://doi.org/10.1080/09243453.2019.1570944

Kennedy, M. M. (1999). The role of preservice teacher education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of teaching and policy* (pp. 54-86). Jossey Bass. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b632b7bf1039784bec47c63f4a8a502f5dbed74c

Kennedy, M. M. (2016). How does professional development improve teaching?. *Review of Educational Research*, *86*(4), 945-980. https://doi.org/10.3102/0034654315626800Kennedy, M. M. (2019). How we learn about teacher learning. Review of Research in Education,43(1), 138-162. https://doi.org/10.3102/0091732X19838970Kim, J. H., & Choi, I. (2021). Choosing the level of significance: A decision-theoretic approach. *Abacus, 57*(1), 27-71. https://doi.org/10.1111/abac.12172

Kim, J., Frank, K., Youngs, P., Salloum, S., & Bieda, K. (2020). Teacher evaluation, ambitious mathematics instruction, and mathematical knowledge for teaching: Evidence from early career teachers, *Journal for Research in Mathematics Education, 53*(3), 181-203. https://doi.org/10.5951/jresematheduc-2020-0093

Kinser-Traut, J. Y., & Turner, E. E. (2020). Shared authority in the mathematics classroom: Successes and challenges throughout one teacher's trajectory implementing ambitious practices. *Journal of Mathematics Teacher Education*, *23*(1), 5-34. https://doi.org/10.1007/s10857-018-9410-x

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2007). Implementing the Four Levels: A Practical Guide for Effective Evaluation of Training Programs. Berrett-Koehler Publishers.
Knapp, N. F., & Peterson, P. L. (1995). Teachers' interpretations of "CGI" after four years: Meanings and practices. *Journal for Research in Mathematics Education, 26*(1), 40–65. https://doi.org/10.5951/jresematheduc.26.1.0040
Koellner, K., & Jacobs, J. (2015). Distinguishing models of professional development: The case of an adaptive model's impact on teachers' knowledge, instruction, and student achievement. *Journal of Teacher Education*, *66*(1), 51-67. https://doi.org/10.1177/0022487114549599
Kokkinos, T., & Gakis, P. (2021). Student teachers' differentiated teaching practices for high-achieving students. *Journal of Further and Higher Education, 45*(7), 916-931. https://doi.org/10.1080/0309877X.2020.1827374
Konstantinou-Katzi, P., Tsolaki, E., Meletiou-Mavrotheris, M., & Koutselini, M. (2013). Differentiation of teaching and learning mathematics: An action research study in tertiary education. *International Journal of Mathematical Education in Science and Technology, 44*(3), 332-349. https://doi.org/10.1080/0020739X.2012.714491
Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643-680. https://doi.org/10.3102/0034654315626799
Kotob, M. M., & Jbaili, F. (2020). Implementing differentiation in early education: The impact on student's academic achievement. *Journal of Applied Linguistics and Language Research*, *7*(2), 110-133. https://jallr.com/index.php/JALLR/article/view/1104
Koutselini, M. (2001). *Curriculum development – theory – research – praxis*. Lythrodondas Press. [In Greek]
Koutselini, M., & Agathangelou, S. (2009). Human rights and equity in teaching. In A. Ross (ed.), *Human rights and citizenship education (pp.237-243)*. CiCe.
Koutselini, M. (2006). *Differentiating teaching – learning in mixed ability classes: Philosophy and application.* Book A. Lythrodontas Press. [In Greek]
Kraft, M. A., & Hill, H. C. (2020). Developing ambitious mathematics instruction through web-based coaching: A randomized field trial. *American Educational Research Journal, 57*(6), 2378-2414. https://doi.org/10.3102/0002831220916840
Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of educational research, 88*(4), 547-588. https://doi.org/10.3102/0034654318759268
Kronberg, R., York-Barr, J., Arnold, K., Gombos, S., Truex, S., Vallejo B., & Stevenson, J. (1997). *Differentiated teaching and learning in heterogeneous classrooms: Strategies for meeting the needs of all students*. Minneapolis. Institute on Community Integration, Minnesota University. https://files.eric.ed.gov/fulltext/ED418538.pdf
Kronborg, L., Plunkett, M., Kelly, L., & Urquhart, F. (2008). Student attitudes towards learning in differentiated settings. *Australasian Journal of Gifted Education*, *17*(2), 23-32. https://search.informit.org/doi/10.3316/informit.720791580709578Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, *9* , 231-251. https://doi.org/10.1007/s10984-006-9015-7
Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive activation in the mathematics classroom and professional*

competence of teachers: Results from the COACTIV project. Springer. https://doi.org/10.1007/978-1-4614-5149-5
Kunter, M., Tsai, Y. M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction*, *18*(5), 468-482. https://doi.org/10.1016/j.learninstruc.2008.06.008
Kyriakides, L., Creemers, B. P., & Antoniou, P. (2009). Teacher behaviour and student outcomes: Suggestions for research on teacher training and professional development. *Teaching and teacher education, 25*(1), 12-23. https://doi.org/10.1016/j.tate.2008.06.001
Lalvani, P. (2013). Privilege, compromise, or social justice: Teachers' conceptualizations of inclusive education. *Disability and Society*, *28*(1), 14-27. https://doi.org/10.1080/09687599.2012.692028
Lampert, M. (1992). The practice and problems of teaching and learning authentic mathematics in school. In F. Oser, A. Dick, & J. Patry (Eds.), *Effective and responsible teaching: The new synthesis* (pp. 295-314). Jossey-Bass Publishers.
Lampert, M. (2001). *Teaching problems and the problems of teaching*. Yale University Press.
Lampert, M. (2010). Learning Teaching in, from, and for Practice: What Do We Mean? *Journal of Teacher Education, 61*(1-2), 21-34. https://doi.org/10.1177/0022487109347321
Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *The Elementary School Journal*, *109*(5), 491-509. https://doi.org/10.1086/596998
Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In M.K. Stein and L. Kucan (Eds.) *Instructional explanations in the disciplines* (pp. 129-141). Springer. https://doi.org/10.1007/978-1-4419-0594-9_9
Lampert, M., Boerst, T., & Graziani, F. (2011). Organizational resources in the service of school-wide ambitious teaching practice. *Teachers College Record, 113*(7), 1361–1400. https://doi.org/10.1177/016146811111300706
Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*(3), 226–243. https://doi.org/10.1177/0022487112473837
Lawrence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards-based learning that benefit the whole class. *American Secondary Education*, *32*(3)*,* 34-62. https://www.jstor.org/stable/41064522
Lee, B., Cawthon, S., & Dawson, K. (2013). Elementary and secondary teacher self-efficacy for teaching and pedagogical conceptual change in a drama-based professional development program. *Teaching and Teacher Education*, *30*, 84-98. https://doi.org/10.1016/j.tate.2012.10.010
Leong, Y. H., Toh, T. L., Tay, E. G., Quek, K. S., Toh, P. C., & Jaguthsing, D. (2021). Scaling up of continual professional development for mathematics problem solving in Singapore schools. *International Journal of Science and Mathematics Education, 19*, 1291-1310. https://doi.org/10.1007/s10763-020-10097-3
Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational researcher, 35*(3), 3-14. https://doi.org/10.3102/0013189X035003003
Lindvall, J., Helenius, O., Eriksson, K., & Ryve, A. (2022). Impact and design of a national-scale professional development program for mathematics teachers. *Scandinavian*

Journal of Educational Research, 66(5), 744-759. https://doi.org/10.1080/00313831.2021.1910563

Lindvall, J., Kirsten, N., Eriksson, K., Brehmer, D., & Ryve, A. (2023). Does the duration of professional development programs influence effects on instruction? An analysis of 174 lessons during a national-scale program. *European Journal of Teacher Education*, 1-19. https://doi.org/10.1080/02619768.2023.2198101

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, *19*(6), 527-537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Little, C. A., Hauser, S., & Corbishley, J. (2009). Constructing Complexity for Differentiated Learning. *Mathematics Teaching in the Middle School, 15*(1), 34-42. https://doi.org/10.5951/MTMS.15.1.0034

Lortie, D. C. (1977). *Schoolteacher: A sociological study*. The University of Chicago Press.

Luna, M. J., & Sherin, M. G. (2017). Using a video club design to promote teacher attention to students' ideas in science. *Teaching and Teacher Education, 66*, 282-294. https://doi.org/10.1016/j.tate.2017.04.019

Lynch, D. (2014). Improving teaching through coaching, mentoring and feedback: a review of literature. *MIER Journal of Educational Studies Trends and Practices, 4*(2), 136-166. https://doi.org/10.52634/mier/2014/v4/i2/1467

Marishane, M. A., Marishane, R. N., & Mahlo, F. D. (2015). Teacher capacity for curriculum differentiation in teaching foundation phase mathematics. *International Journal of Educational Sciences, 11*(3), 253-262. https://doi.org/10.1080/09751122.2015.11890396

Marrongelle, K., Sztajn, P., & Smith, M. (2013). Scaling up professional development in an era of common state standards. *Journal of Teacher Education, 64*(3), 202-211. https://doi.org/10.1177/0022487112473838

Marshall, J. C., & Horton, R. M. (2011). The relationship of teacher-facilitated, inquiry-based instruction to student higher-order thinking. *School Science and Mathematics, 111*(3), 93-101. https://doi.org/10.1111/j.1949-8594.2010.00066.x

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. Routledge. https://doi.org/10.4324/9780203471876

Maulana, R., Helms-Lorenz, M., Moorer, P., Smale-Jacobse, A., & Feng, X. (2023). *Differentiated instruction in teaching from the international perspective: Methodological and empirical insights*. University of Groningen Press. https://doi.org/10.21827/62c5541759973

McChesney, K., & Aldridge, J. M. (2019). A review of practitioner-led evaluation of teacher professional development. *Professional development in education, 45*(2), 307-324. https://doi.org/10.1080/19415257.2018.1452782

McCormick, M. (2016). Exploring the cognitive demand and features of problem solving tasks in primary mathematics classrooms. In B. White, M. Chinnappan, & S. Trenholm (Eds.) *Opening up mathematics education research* (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australia, pp. 455-462). MERGA. https://files.eric.ed.gov/fulltext/ED572329.pdf

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, *64*(5), 378-386. https://doi.org/10.1177/0022487113493807

McQuarrie, L. M., & McRae, P. (2010). A provincial perspective on differentiated instruction: The Alberta Initiative for School Improvement (AISI). *Journal of Applied Research on Learning*, *3*(4), 1-18.

Mellroth, E., Bergwall, A., & Nilsson, P. (2021). Task design for differentiated instruction in mixed-ability mathematics classrooms: Manifestations of contradictions in a professional learning community. *Mathematics Teacher Education and Development, 23*(3), 78-96. http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-93642
Merchie, E., Tuytens, M., Devos, G., & Vanderlinde R. (2016). Evaluating teachers' professional development initiatives: Towards an extended evaluative framework. *Research Papers in Education, 33*(2),1-26. https://doi.org/10.1080/02671522.2016.1271003
Mertens, D. M. (2010). *Research and evaluation in education and psychology*. SAGE Publications Inc.
Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474-496. https://doi.org/10.1002/tea.20347
Mitchell, R., Beisiegel, M., & Barmore, J. (2022). How much does facilitation matter? Comparing externally and participant-facilitated, MQI-based video clubs. *Mathematics Education Research Journal*, *34*, 369-392. https://doi.org/10.1007/s13394-020-00348-5Mitchell, R.N., & Marin, K.A. (2015). Examining the use of a structured analysis framework to support prospective teacher noticing. *Journal of Mathematics Teacher Education, 18,* 551–575. https://doi.org/10.1007/s10857-014-9294-3
Mullis, I. V., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center, Boston College. https://www.iea.nl/sites/default/files/2020-12/TIMSS%202019-International-Results-in-Mathematics-and-Science.pdf
Munter, C., & Wilhelm, A. G. (2021). Mathematics teachers' knowledge, networks, practice, and change in instructional visions. *Journal of Teacher Education, 72*(3), 342-354. https://doi.org/10.1177/0022487120949836
Murawski, C. (2019). *The role of content-focused coaching in fostering ambitious mathematics teaching practices in elementary classrooms* [Doctoral dissertation]. University of Pittsburgh. https://d-scholarship.pitt.edu/38040/1/MurawskiCorinneMarko_ETD_2019.pdf
Murnane, R., Sawhill, I., & Snow, C. (2012). Literacy challenges for the twenty-first century: Introducing the issue. *The Future of Children, 22*(2), 3-15. https://www.jstor.org/stable/23317408
Muthoni, W. M., & Mbugua, Z. K. (2014). Effectiveness of differentiated instruction on secondary school students' achievement in mathematics. *International Journal of Applied Science and Technology, 4*(1), 116-122. https://www.ijastnet.com/journals/Vol_4_No_1_January_2014/12.pdf
National Council for Teaching Mathematics [NCTM]. (2014). *Principles to actions: Ensuring mathematical success for all*. The National Council of Teachers of Mathematics, Inc. https://www.nctm.org/Store/Products/Principles-to-Actions--Ensuring-Mathematical-Success-for-All/
Newmann, F., & Wehlage, G. G., (1996). *Authentic achievement: Restructuring schools for intellectual quality.* Jossey Bass. https://files.eric.ed.gov/fulltext/ED387925.pdf
Noddings, N. (2001). The caring teacher. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.). American Educational Research Association. https://www.aera.net/Publications/Books/Handbook-of-Research-on-Teaching-Fourth-Edition

OECD (2023). Education at a glance 2023: OECD Indicators. OECD Publishing, Paris. https://doi.org/10.1787/e13bef63-en.

OECD. (2017). *Trends shaping education 2017 Spotlight 8: Mind the gap: Inequity in education*. France, Paris: OECD Publications.

Ollerton, M. (2009). Sacred cows. *Mathematics Teaching (Association of Teachers of Mathematics), 215*, 31-34. https://eric.ed.gov/?id=EJ859746

Opfer, D and Pedder, D., 2011. Conceptualizing Teacher Professional Learning. *Review of educational research, 81*(3), 376-407. https://doi.org/10.3102/0034654311413609

Organisation for Economic Co-operation and Development [OECD] (2019), PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris. https://doi.org/10.1787/5f07c754-en.

Otten, S., de Araujo, Z., Candela, A. G., Vahle, C., Stewart, M. E., & Wonsavage, F. P. (2022, January). Incremental change as an alternative to ambitious professional development. In Lischka, A. E., Dyer, E. B., Jones, R. S., Lovett, J. N., Strayer, J., & Drown, S. *Proceedings of the forty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Middle Tennessee State University. https://par.nsf.gov/servlets/purl/10405088

Özdemir, G., & Clark, D. B. (2007). An overview of conceptual change theories. Eurasia *Journal of Mathematics, Science and Technology Education, 3*(4), 351-361. https://doi.org/10.12973/ejmste/75414

Palieraki, S., & Koutrouba, K. (2021). Differentiated Instruction in Information and Communications Technology Teaching and Effective Learning in Primary Education. *European Journal of Educational Research, 10*(3), 1487-1503. https://doi.org/10.12973/eu-jer.10.3.1487

Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., Pierczynski, M., & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research, 88*(2), 205–242. https://doi.org/10.3102/0034654317743198

Patton, M. Q. (2015). *Qualitative research and evaluation methods*. Sage Publications Inc.

Pettersen, A. & Nortvedt, G. (2017). Identifying competency demands in mathematical tasks: Recognising what matters. *International Journal of Science and Mathematics Education, 16*, 1-17. https://doi.org/10.1007/s10763-017-9807-5

Pfister, M., Opitz, E. M., & Pauli, C. (2015). Scaffolding for mathematics teaching in inclusive primary classrooms: A video study. *ZDM Mathematics Education*, *47*(1), 1079-1092. https://doi.org/10.1007/s11858-015-0713-4

Pham, H. (2011). Differentiated instruction and the need to integrate teaching and practice. *Journal of College Teaching & Learning, 9*(1), 13–20. https://doi.org/10.19030/tlc.v9i1.6710

Piaget, J. (1936). Origins of intelligence in the child. Routledge.

Pogrow, S. (1988). Teaching thinking to at-risk elementary students. *Educational Leadership, 45*(7), 79–85. https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_198804_pogrow.pdf

Polly, D., Neale, H., & Pugalee, D. K. (2014). How does ongoing task-focused mathematics professional development influence elementary school teachers' knowledge, beliefs and enacted pedagogies?. *Early Childhood Education Journal*, *42*(1), 1-10. https://doi.org/10.1007/s10643-013-0585-6

Ponte, J. P., & Quaresma, M. (2016). Teachers' professional practice conducting mathematical discussions. *Educational Studies of Mathematics, 93*(1), 51–66. https://doi.org/10.1007/s10649-016-9681-z

Porter, A. C., Garet, M. S., Desimone, L., Yoon, K. S., & Birman, B. F. (2000). *Does professional development change teaching practice? Results from a three-year study*. U.S. Department of Education, ED Pubs. https://files.eric.ed.gov/fulltext/ED455227.pdfPozas, M., Letzel, V., & Schneider, C. (2020). Teachers and differentiated instruction: exploring differentiation practices to address student diversity. *Journal of Research in Special Educational Needs*, *20*(3), 217-230. https://doi.org/10.1111/1471-3802.12481

Praetorius, A. K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction, 31*, 2-12. https://doi.org/10.1016/j.learninstruc.2013.12.002

Prast, E. J., Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. (2018). Differentiated instruction in primary mathematics: Effects of teacher professional development on student achievement. *Learning and Instruction*, *54*, 22-34. https://doi.org/10.1016/j.learninstruc.2018.01.009

Psycharis, G., Potari, D., Triantafillou, C., & Zachariades, T. (2019, February). Teachers' attempts to address both mathematical challenge and differentiation in whole class discussion. In the *Eleventh Congress of the European Society for Research in Mathematics Education* (No. 27). Freudenthal Group; Freudenthal Institute; ERME. https://hal.science/hal-02430171/document

Putnam, R. T., Heaton, R. M., Prawat, R. S., & Remillard, J. (1992). Teaching mathematics for understanding: Discussing case studies of four fifth-grade teachers. *Elementary School Journal, 93*(2), 213–228. https://doi.org/10.1086/461723

Puzio, K., Newcomer, S. N., & Goff, P. (2015). Supporting literacy differentiation: The principal's role in a community of practice. *Literacy research and instruction, 54*(2), 135-162. https://doi.org/10.1080/19388071.2014.997944

Ramos, J. L., Cattaneo, A. A., de Jong, F. P., & Espadeiro, R. G. (2022). Pedagogical models for the facilitation of teacher professional development via video-supported collaborative learning. A review of the state of the art. *Journal of Research on Technology in Education, 54*(5), 695-718. https://doi.org/10.1080/15391523.2021.1911720

Raundenbush, S. W., & Bryk, A.S. (2002). *Hierarchical linear components*. Sage Publications Inc.

Reed, G. G. (2011). The complexity of moral learning: Diversity, deprovincialisation and privilege. *Journal of Moral Education, 40*(3), 359–367. doi: 10.1080/03057240.2011.596337

Reio Jr, T. G., Rocco, T. S., Smith, D. H., & Chang, E. (2017). A critique of Kirkpatrick's evaluation model. *New Horizons in Adult Education and Human Resource Development, 29*(2), 35-53. DOI: 10.1002/nha3.20178

Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientations toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Education, 35*(5), 352-388. https://doi.org/10.2307/30034820

Resnick, L. B. (1987). The 1987 Presidential Address Learning In School and Out. *Educational Researcher, 16*(9), 13-54. https://doi.org/10.3102/0013189X016009013

Resnick, L. B., & Zurawsky, C. (2006). Do the Math: Cognitive Demand Makes a Difference. *American Educational Research Association (AERA) Research Points, 4*(2), 1-4. https://files.eric.ed.gov/fulltext/ED497645.pdf

Rich, P. J., & Hannafin, M. (2009). Video annotation tools: Technologies to scaffold, structure, and transform teacher reflection. *Journal of Teacher Education*, *60*(1), 52-67. https://doi.org/10.1177/0022487108328486

Riger, S., & Sigurvinsdottir, R. (2016). Thematic analysis. In L. A. Jason & D. S. Glenwick (Eds.), *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods* (pp. 33-41). Oxford University Press.

Ritzema, E. S., Deunk, M. I., & Bosker, R. J. (2016). Differentiation practices in grade 2 and 3: Variations in teacher behavior in mathematics and reading comprehension lessons. *Journal of Classroom Interaction, 51*(2), pp. 50-72.

Robinson, L., Maldonado, N., & Whaley, J. (2014, November). *Perceptions about implementation of differentiated instruction* [Paper presentation]. Presented at the Annual Mid-South Educational Research (MSERA) Conference (pp. 1–22). https://files.eric.ed.gov/fulltext/ED554312.pdf

Roche, A., & Clarke D. (2015). Describing the nature and effect of teacher interactions with students during seat work on challenging tasks. In M. Marshman, V. Geiger, & A. Bennison (Eds.), *Mathematics education in the margins (Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 532–539). MERGA. https://www.merga.net.au/documents/RP2015-59.pdf

Roche, A., Clarke, D., Sullivan, P., & Cheeseman, J. (2013). Strategies for encouraging students to persist on challenging tasks: Some insights from work in classrooms. *Australian Primary Mathematics Classroom, 18*(4), 27-32. https://files.eric.ed.gov/fulltext/EJ1093218.pdf

Roth McDuffie, A. M., & Mather, M. (2006). Reification of instructional materials as part of the process of developing problem-based practices in mathematics education. *Teachers and Teaching: Theory and Practice*, *12*(4), 435-459. https://doi.org/10.1080/13450600600644285

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal on Research in Science Teaching, 48*(2),117–148. https://doi.org/10.1002/tea.20408

Russo, J., & Hopkins, S. (2019). Teachers' perceptions of students when observing lessons involving challenging tasks. *International Journal of Science and Mathematics Education*, *17*(4), 759-779. https://doi.org/10.1007/s10763-018-9888-9

Russo, J., Bobis, J., Downton, A., Hughes, S., Livy, S., McCormick, M., & Sullivan, P. (2019). Teaching with challenging tasks in the first years of school: What are the obstacles and how can teachers overcome them?. *Australian Primary Mathematics Classroom*, *24*(1), 11-18. https://www.researchgate.net/profile/James-Russo/publication/332139139_Teaching_with_challenging_tasks_in_the_first_years_of_school_What_are_the_obstacles_and_how_can_teachers_overcome_them/links/5ca2d178299bf1116956b6ca/Teaching-with-challenging-tasks-in-the-first-years-of-school-What-are-the-obstacles-and-how-can-teachers-overcome-them.pdf

Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A. K., Yang, X., & Kaiser, G. (2021). Mathematics teacher learning to notice: A systematic review of studies of video-based programs. *ZDM–Mathematics Education, 53*(1), 119-134. https://doi.org/10.1007/s11858-020-01216-z

Santamaria, L. J. (2009). Culturally responsive differentiated instruction: Narrowing gaps between best pedagogical practices benefiting all learners. *Teachers College Record, 111*(1), 214-247. https://doi.org/10.1177/016146810911100105

Santana, C. C. (2020). *Teacher conceptualization and implementation of differentiated instruction in the elementary reading classroom* [Doctoral dissertation]. The College of William and Mary in Virginia. https://doi.org/10.25774/w4-hbxg-jt80

Santangelo, T., & Tomlinson, C. A. (2012). Teacher educators' perceptions and use of differentiated instruction practices: An exploratory investigation. *Action in Teacher Education*, *34*(4), 309-327. https://doi.org/10.1080/01626620.2012.717032

Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. OECD Publishing.

Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM, 48*, 29-40. https://doi.org/10.1007/s11858-016-0765-0

Schoenfeld, A. H. (2023). A theory of teaching. In C. Y. Charalambous, & A. K. Praetorius (Eds.), *Theorizing teaching: Current status and open issues* (pp. 159-187).Springer Cham.https://doi.org/10.1007/978-3-031-25613-4

Schwab, S., Sharma, U., & Hoffmann, L. (2022). How inclusive are the teaching practices of my German, Maths and English teachers? – psychometric properties of a newly developed scale to assess personalisation and differentiation in teaching practices. *International Journal of Inclusive Education, 26*(1), 61–76. https://doi.org/10.1080/13603116.2019.1629121

Schwarzer, R., & Hallum, S. (2008). Perceived teacher self-efficacy as a predictor of job stress and burnout: Mediation analyses. *Applied Psychology*, *57*(1), 152-171. https://doi.org/10.1111/j.1464-0597.2008.00359.x

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454-499. https://doi.org/10.3102/0034654307310317

Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development. *English Language Teaching*, *3*(4), 237-248. https://files.eric.ed.gov/fulltext/EJ1081990.pdf

Sherin, M. G. (2001). Developing a professional vision of classroom events. In T. Wood, B. S. Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 75–93). Erlbaum. https://doi.org/10.4324/9781410612335-12

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education, 20*(2), 163-183. https://doi.org/10.1016/j.tate.2003.08.001

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20-37. https://doi.org/10.1177/0022487108328155

Sherin, M. G., & van Es., E. A. (2005). Using video to support teachers' ability to notice classroom interactions*. Journal of Technology and Teacher Education, 13*(3), 475-491. https://www.learntechlib.org/p/4824/

Shumway, J. F., Bundock, K., King, J., Burnside, M., Gardner, H., & Messervy, F. (2020). Visualizing Number: Instruction for Number System Knowledge in Second-Grade Classrooms. *Investigations in Mathematics Learning, 12*(2), 142–161. https://doi.org/10.1080/19477503.2020.1740383

Siegel, S., & Castellan, N.J. (1988). *Non-parametric statistics: For the behavioural sciences.* (2nd ed.). McGraw-Hill Book Company. https://psycnet.apa.org/record/1988-97307-000

Silver, E. A., & Kenney, P. A. (2016). *More lessons learned from research, Volume 2: Helping all students understand important mathematics*. National Council of

Teachers of Mathematics. https://www.nctm.org/Store/Products/More-Lessons-Learned-from-Research,-Volume-2-(Download)/
Silver, E. A., & Stein, M. K. (1996). The QUASAR Project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education, 30*(4), 476-521. https://doi.org/10.1177/0042085996030004006
Silver, E. A., Clark, L. M., Ghousseini, H. N., Charalambous, C. Y., & Sealy, J. T. (2007). Where is the mathematics? Examining teachers' mathematical learning opportunities in practice-based professional learning tasks. *Journal of Mathematics Teacher Education, 10*, 261-277. https://doi.org/10.1007/s10857-007-9039-7
Silver, E. A., Mesa, V. M., Morris, K. A., Star, J. R., & Benken, B. M. (2009). Teaching mathematics for understanding: An analysis of lessons submitted by teachers seeking NBPTS certification. *American Educational Research Journal, 46*(2), 501-531. https://doi.org/10.3102/0002831208326559
Simpkins, P., Mastropieri, A., & Scruggs, T. (2009). Differentiated curriculum enhancements in inclusive fifth-grade science classes. *Remedial and Special Education, 30*(5), 300–308. https://doi.org/10.1177/0741932508321011
Sims, S., Fletcher-Wood, H., O'Mara-Eves, A., Cottingham, S., Stansfield, C., Van Herwegen, J., Anders, J. (2021). *What are the Characteristics of Teacher Professional Development that Increase Pupil Achievement? A systematic review and meta-analysis.*Education Endowment Foundation. https://educationendowmentfoundation.org.uk/education-evidence/evidence-reviews/teacherprofessional-development-characteristics
Slade, M. L., Dettmer, P. A., & Miller, T. N. (2006). Professional development for the education of secondary gifted students. In F. A. Dixon & S. M. Moon (Eds.), *Handbook of secondary gifted education* (pp. 611–648). Prufrock Press.
Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated instruction in secondary education: A systematic review of research evidence. *Frontiers in psychology, 10*, 1-23. | https://doi.org/10.3389/fpsyg.2019.02366
Smets, W., De Neve, D., & Struyven, K. (2022). Responding to students' learning needs: how secondary education teachers learn to implement differentiated instruction. *Educational Action Research, 30*(2), 243-260. https://doi.org/10.1080/09650792.2020.1848604
Smit, R., & Humpert, W. (2012). Differentiated instruction in small schools. *Teaching and Teacher Education, 28*(8), 1152–1162. https://doi.org/10.1016/j.tate.2012.07.003
Smith, A., & Kelly, A. (2015). Cognitive processes. In S. Whitbourne, T*he Encyclopaedia of Adulthood and Aging*. Wiley. https://doi.org/10.1002/9781118521373.wbeaa213
Smith, C., & Gillespie, M. (2023). Research on professional development and teacher change: Implications for adult basic education. In *Review of Adult Learning and Literacy*, Volume 7 (pp. 205-244). Routledge.
Smith, M. S., & Stein, M. K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3*(5), 344–350. https://doi.org/10.5951/MTMS.3.5.0344
Smith, M. S., & Stein, M. K. (2011). *Five practices for orchestrating productive mathematical discussions*. National Council of Teacher of Mathematics. https://www.nctm.org/Store/Products/5-Practices-for-Orchestrating-Productive-Mathematics-Discussions,-2nd-edition-(Download)/

Smith, M. S., Bill, V., & Hughes, E. K. (2008). "Thinking Through a Lesson: Successfully Implementing High-Level Tasks." *Mathematics Teaching in the Middle School, 21*(9), 532-539. https://doi.org/10.5951/MTMS.14.3.0132

Smith, M., & Stein, M. K. (2023). Mathematical Tasks: The Lasting Legacy of the QUASAR Project. In J., Cai, G. J., Stylianides, & P. A., Kenney, *Research Studies on Learning and Teaching of Mathematics: Dedicated to Edward A. Silver* (pp. 275-297). Springer Cham. https://doi.org/10.1007/978-3-031-35459-5_12

Spangler, D. A., & Wanko, J. J. (Eds.). (2017). *Enhancing classroom practice with research behind the principles to actions*. The National Council of Teachers of Mathematics, Inc. https://www.nctm.org/Store/Products/Enhancing-Classroom-Practice-with-Research-behind-Principles-to-Actions/

Spillane, J. P., Hopkins, M., & Sweet, T. M. (2018). School district educational infrastructure and change at scale: Teacher peer interactions and their beliefs about mathematics instruction. *American Educational Research Journal*, *55*(3), 532-571. https://doi.org/10.3102/0002831217743928

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. *Review of Educational Research, 72*(3), 387–431. https://doi.org/10.3102/00346543072003387

Stein, M. K., & Kaufman, J. H. (2010). Selecting and supporting the use of mathematics curricula at scale. *American Educational Research Journal, 47*(3), 663-693. https://doi.org/10.3102/0002831209361210

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation*, *2*(1), 50-80. https://doi.org/10.1080/1380361960020103

Stein, M. K., Baxter, J. A., & Leinhardt, G. (1990). Subject-matter knowledge and elementary instruction: A case from functions and graphing. American *Educational Research Journal, 27*(4), 639–663. https://doi.org/10.3102/00028312027004639

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313-340. https://doi.org/10.1080/10986060802229675

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455-488. https://doi.org/10.3102/00028312033002455

Stein, M. K., Remillard, J., & Smith, M. S. (2007). In F. K. Lester (Ed.), Second handbook of research on mathematics teaching and learning (pp. 319–369). Information Age Publishing.

Stein, M. K., Smith, M. S., Henningsen, M. A., & Silver, E.A. (2009). *Implementing standards-based mathematics instruction: A Casebook for professional development (2nd Ed)*. Teachers College Press.

Stigler, J. W. & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership*, *61*(5), 12-17. https://www.researchgate.net/profile/James-Stigler/publication/228731157_Improving_mathematics_teaching/links/02e7e529e9b1081f6f000000/Improving-mathematics-teaching.pdf

Stipek, D. J., Salmon, J. M., Givvin, K. B., Kazemi, E., Saxe, G., & Mac-Gyvers, V. L. (1998). The value (and convergence) of practices suggested by motivation research and promoted by mathematics education reformers. *Journal for Research in Mathematics Education, 29*(4), 465–488. https://doi.org/10.5951/jresematheduc.29.4.0465

Stosich, E. L. (2016). Building teacher and school capacity to teach to ambitious standards in high-poverty schools. *Teaching and Teacher Education*, *58*, 43-53. https://doi.org/10.1016/j.tate.2016.04.010

Stradling, B., & Saunders, L. (1993). Differentiation in practice: Responding to the needs of all pupils. *Educational Research, 35*(2), 127-137. https://doi.org/10.1080/0013188930350202

Stylianides, A. J., & Stylianides, G. J. (2014). Impacting positively on students' mathematical problem-solving beliefs: An instructional intervention of short duration. *The Journal of Mathematical Behavior*, *33*, 8-29. https://doi.org/10.1016/j.jmathb.2013.08.005

Stylianides, G. J., & Stylianides, A. J. (2020). Research-based interventions in the area of proof: The past, the present, and the future. *Educational Studies in Mathematics, 104*(1), 9-24. https://doi.org/10.1007/s10649-017-9782-3

Suh, J., Gallagher, M. A., Capen, L., & Birkhead, S. (2021). Enhancing teachers' noticing around mathematics teaching practices through video-based lesson study with peer coaching. *International Journal for Lesson & Learning Studies, 10*(2), 150-167. https://doi.org/10.1108/IJLLS-09-2020-0073

Sullivan, P. (2011). *Teaching mathematics: Using research-informed strategies*. Australian Council for Educational Research (ACER). https://research.acer.edu.au/cgi/viewcontent.cgi?article=1022&context=aer

Sullivan, P., & Mornane, A. (2014). Exploring teachers' use of, and students' reactions to, challenging mathematics tasks. *Mathematics Education Research Journal, 26*(2), 193–213. https://doi.org/10.1007/s13394-013-0089-0

Sullivan, P., Askew, M., Cheeseman, J., Clarke, D., Mornane, A., Roche, A., & Walker, N. (2015). Supporting teachers in structuring mathematics lessons involving challenging tasks. *Journal of Mathematics Teacher Education, 18*, 123-140. https://doi.org/10.1007/s10857-014-9279-2

Sullivan, P., Aulert, A., Lehmann, A., Hislop, B., Shepherd, O., & Stubbs, A. (2013). Classroom culture, challenging mathematical tasks and student persistence. In V. Steinle, L. Ball & C. Bardini (Eds.), *Mathematics education: Yesterday, today and tomorrow (Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia).* (pp. 618-625). MERGA. https://files.eric.ed.gov/fulltext/ED573047.pdf

Sullivan, P., Borcek, C., Walker, N., & Rennie, M. (2016a). Exploring a structure for mathematics lessons that initiate learning by activating cognition on challenging tasks. *The Journal of Mathematical Behavior, 41*, 159–170. https://doi.org/10.1016/j.jmathb.2015.12.002

Sullivan, P., Clarke, D., & Clarke, B. (2009). Converting mathematics tasks to learning opportunities: An important aspect of knowledge for mathematics teaching. *Mathematics Education Research Journal, 21*(1), 85–105. https://doi.org/10.1007/BF03217539

Sullivan, P., Holmes, M., Ingram, N., Linsell, C., Livy, S., & McCormack, M. (2016b). The Intent and Processes of a Professional Learning Initiative Seeking to Foster Discussion around Innovative Approaches to Teaching. In White, B., Chinnappan, M. & Trenholm, S. (Eds.). *Opening up mathematics education research (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia,* pp. 669–672). MERGA. https://files.eric.ed.gov/fulltext/ED572341.pdf

Sullivan, P., Mousley, J., & Zevenbergen, R. (2006). Teacher actions to maximize mathematics learning opportunities in heterogeneous classrooms. *International Journal of Science and Mathematics Education, 4*(1), 117–143. https://doi.org/10.1007/s10763-005-9002-y

Sun, J., & van Es, E. A. (2015). An exploratory study of the influence that analyzing teaching has on preservice teachers' classroom practice. *Journal of Teacher Education, 66*(3), 201-214. https://doi.org/10.1177/0022487115574103

Sun, M., Wilhelm, A. G., Larson, C. J., & Frank, K. A. (2014). Exploring colleagues' professional influence on mathematics teachers' learning. *Teachers College Record, 116*(6), 1-30. https://doi.org/10.1177/016146811411600604

Superfine, A. & Bragelman, J. (2018). Analyzing the Impact of Video Representation Complexity on Preservice Teacher Noticing of Children's Thinking. *Eurasia Journal of Mathematics, Science and Technology Education, 14*(11). https://doi.org/10.29333/ejmste/99501

Superfine, A.C., & Superfine, B.M. (2023). From Mathematical Tasks to Research–Practice Partnerships: A Look at Edward Silver's Influence on Our Efforts at Mathematics Instructional Improvement. In: Cai, J., Stylianides, G.J., Kenney, P.A. (eds) *Research Studies on Learning and Teaching of Mathematics. Research in Mathematics Education*. Springer. https://doi.org/10.1007/978-3-031-35459-5_6

Sykes, G., Bird, T., & Kennedy, M. (2010). Teacher education: Its problems and some prospects. *Journal of Teacher education, 61*(5), 464-476. https://doi.org/10.1177/0022487110375804

Tam, A. C. F. (2015). The role of a professional learning community in teacher change: A perspective from beliefs and practices. *Teachers and Teaching, 21*(1), 22-43. https://doi.org/10.1080/13540602.2014.928122

Taras, A., Smit, R., Hecht, P., & Matic, M. (2022). Noticing inclusive teaching practices in tandems–results from cross-national video clubs at two different school levels. *International Journal of Inclusive Education*, 1-17. https://doi.org/10.1080/13603116.2022.2119489

Tarr, J., Reys, R., Reys, B., Chávez, Ó, Shih, J., & Osterlind, S. (2008). The Impact of Middle-Grades Mathematics Curricula and the Classroom Learning Environment on Student Achievement. *Journal for Research in Mathematics Education, 39*(3), 247-280. https://doi.org/10.2307/30034970

Taylor, B. K. (2015). Content, process, and product: Modeling differentiated instruction. *Kappa Delta Pi Record, 51*(1), 13-17. https://doi.org/10.1080/00228958.2015.988559

Tekkumru-Kisa, M. T., & Stein, M. K. (2015). Learning to see teaching in new ways a foundation for maintaining cognitive demand. *American Educational Research Journal, 52*(1), 105-136. https://doi.org/10.3102/0002831214549452

Tekkumru-Kisa, M., Schunn, C., Stein, M. K., & Reynolds, B. (2019). Change in thinking demands for students across the phases of a science task: An exploratory study. *Research in Science Education*, *49*(3), 859-883. https://doi.org/10.1007/s11165-017-9645-z

Tekkumru-Kisa, M., Stein, M. K., & Doyle, W. (2020). Theory and research on tasks revisited: task as a context for students' thinking in the era of ambitious reforms in mathematics and science. *Educational Researcher*, *49*(8), 606–617. https://doi.org/10.3102/0013189X20932480

The New Teacher Project [TNTP] (2015). *The mirage: Confronting the hard truth about our quest for teacher development.* TNTP. https://files.eric.ed.gov/fulltext/ED558206.pdf

Thompson, J., Windschitl, M., & Braaten, M. (2013). Developing a theory of ambitious early-career teacher practice. *American Educational Research Journal, 50*(3), 574–615. https://doi.org/10.3102/0002831213476334

Tieso, C. (2005). The effects of grouping practices and curricular adjustments on achievement. *Journal for the Education of the Gifted*, *29*(1), 60-89. https://doi.org/10.1177/016235320502900104
Timperley, H. (2008). *Teacher professional learning and development* (Educational Practices Series 18). International Academy of Education. www.ibe.unesco.org/fileadmin/user_upload/Publications/Educational_Practices/EdPractices_18.pdf
Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Ministry of Education. https://educationcounts.edcentre.govt.nz/goto/BES
Timperley, H.S. & Alton-Lee, A. (2008). Reframing teacher professional learning: An alternative policy approach to strengthening valued outcomes for diverse learners. In G. Kelly, A. Luke and J. Green (Eds.), *Disciplines, knowledge and pedagogy: review of research in education Vol. 32* (pp. 328-369). Sage Publications Inc. https://doi.org/10.3102/0091732X07308968
Tomlinson C. A. & Allan, S. (2000). Leadership for differentiating schools and classrooms. Association for Supervision & Curriculum Development.
Tomlinson, C. A. (1999). Mapping a route toward differentiated instruction. *Educational leadership, 57*(1), 12-17. Tomlinson, C. A. (2005a). Traveling the road to differentiation in staff development. *Journal of Staff Development, 26*(4), 8-12.
Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). ASCD.
Tomlinson, C. A. (2005a). Traveling the road to differentiation in staff development. *Journal of Staff Development, 26*(4), 8-12. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4a1bfc061f59ac08d2087203c5d48f5a5a827923
Tomlinson, C. A. (2005b). Grading and differentiation: Paradox or good practice?. *Theory into Practice, 44*(3), 262-269. https://doi.org/10.1207/s15430421tip4403_11
Tomlinson, C. A. (2008). The goals of differentiation. *Educational leadership, 66*(3), 26-30. https://victesol.vic.edu.au/wp-content/uploads/wpforo/default_attachments/1498800479-differentiationtomlinson.pdf
Tomlinson, C. A. (2015). Teaching for excellence in academically diverse classrooms. *Society, 52*, 203-209. https://doi.org/10.1007/s12115-015-9888-0
Tomlinson, C. A. (2017). *How to differentiate instruction in academically diverse classrooms* (3rd ed.). Association for Supervision & Curriculum Development.
Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). *Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: a review of literature. Journal for the Education of the Gifted, 27*(2-3), 119-145. https://doi.org/10.1177/016235320302700203
Tomlinson, C. A., Moon, T. R., & Imbeau, M. B. (2013). *Assessment and student success in a differentiated classroom*. Association for Supervision and Curriculum Development. https://files.ascd.org/staticfiles/ascd/pdf/siteASCD/publications/assessment-and-di-whitepaper.pdf
Tomlinson, C., Brimijoin, K., & Narvaez, L. (2008). *The differentiated school: Making revolutionary changes in teaching and learning.* Association for Supervision and

Curriculum Development. https://www.ascd.org/books/the-differentiated-school?variant=105005
Tomlinson., C.A., & Imbeau, M.B. (2023). *Leading and managing a differentiated classroom* (2nd Ed.). Association for Supervision and Curriculum Development. https://files.ascd.org/pdfs/publications/books/Leading-and-Managing-A-Differentiated-Classroom-2ed-sample-pages.pdf
Trotter, Y. D. (2006). Adult learning theories: Impacting professional development programs. *Delta Kappa Gamma Bulletin*, *72*(2), 8-13.
Tulbure, C. (2011). Do different learning styles require differentiated teaching strategies?. *Procedia-Social and Behavioral Sciences*, *11*, 155-159. https://doi.org/10.1016/j.sbspro.2011.01.052
Turner, W. D., & Solis, O. J. (2017). The misnomers of differentiating instruction in large classes. Journal of Effective Teaching, 17(3), 64–76. https://files.eric.ed.gov/fulltext/EJ1176034.pdf
Tzur, R. (2008). Profound Awareness of the Learning Paradox (PALP): a journey towards epistemologically regulated pedagogy in mathematics teaching and teacher education. In T. Woods, & B. Jaworski (Eds.), *Handbook of research in mathematics teacher education* (vol. 4, pp. 1–2).Brill Sense. https://doi.org/10.1163/9789087905521_009
Valiandes, A.S., Kyriakides, L., & Koutselini, M. (2011). Investigating the Impact of Differentiated Instruction in Mixed Ability Classrooms: It's Impact on the Quality and Equity Dimensions of Education Effectiveness. In the *Proceedings of the 24th International Congress for School Effectiveness and Improvement* (pp. 1-19). University of Cyprus. https://www.researchgate.net/profile/Stavroula-Valiandes/publication/262566283_Valiandes_S_Koutselini_M_Kyriakides_L_2011_Investigating_the_Impact_of_Differentiated_Instruction_in_Mixed_Ability_Classrooms_its_impact_on_the_Quality_and_Equity_Dimensions_of_Education_Effectiveness/links/5b4b768145851519b4c01955/Valiandes-S-Koutselini-M-Kyriakides-L-2011-Investigating-the-Impact-of-Differentiated-Instruction-in-Mixed-Ability-Classrooms-its-impact-on-the-Quality-and-Equity-Dimensions-of-Education-Effect.pdf
Valiandes, S. (2015) Evaluating the impact of differentiated instruction on literacy and reading in mixed ability classrooms: Quality and equity dimensions of education effectiveness. *Studies in Educational Evaluation, 45*, 17-26. https://doi.org/10.1016/j.stueduc.2015.02.005
Valiandes, S., & Koutselini, M. I. (2009, June). Application and evaluation of differentiation instruction in mixed ability classrooms [Paper presentation]. In *4th Hellenic Observatory PhD Symposium* (pp. 25-26). London School of Economics. https://websites.ucy.ac.cy/release/documents/Publications/English/DifferentiationInstructionInMixedAbilityClassrooms.pdf
Valiandes, S., & Neophytou, L. (2018). Teachers' professional development for differentiated instruction in mixed-ability classrooms: Investigating the impact of a development program on teachers' professional learning and on students' achievement. *Teacher Development, 22* (1), 123–138. https://doi.org/10.1080/13664530.2017.1338196Valiandes, S., Neophytou, L., Leban, T., Paparistodemou, E., Bermúdez, M. M. & Boghian, I., Mâţă, L., & Núñez-Delgado, P. (2017). *Teachers' guide to differentiated instruction*. Cyprus Pedagogical Institute. https://www.didesu.cy.net/docs/1_ENGLISH_SELIDOSH.pdf
Van Driel, J. H., & Berry, A. (2012). Teacher professional development focusing on pedagogical content knowledge. *Educational researcher, 41*(1), 26-28. https://doi.org/10.3102/0013189X11431010

van Es, E. A. (2012). Using video to collaborate around problems of practice. *Teacher Education Quarterly, 39*(2), 103-116. https://www.jstor.org/stable/23479674
van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, *10*(4), 571-596. https://www.learntechlib.org/primary/p/9171/
van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and teacher education, 24*(2), 244-276. https://doi.org/10.1016/j.tate.2006.11.005
van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education, 13*, 155-176. https://doi.org/10.1007/s10857-009-9130-3
van Es, E. A., Cashen, M., Barnhart, T., & Auger, A. (2017a). Learning to notice mathematics instruction: Using video to develop preservice teachers' vision of ambitious pedagogy. *Cognition and Instruction*, *35*(3), 165-187. https://doi.org/10.1080/07370008.2017.1317125
van Es, E. A., Hand, V., & Mercado, J. (2017b). Making visible the relationship between teachers' noticing for equity and equitable teaching practice. In E. Schack, M. Fisher, & J. Wilhelm (eds.) *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks. Research in Mathematics Education* (pp. 251-270). Springer. https://doi.org/10.1007/978-3-319-46753-5_15
van Es, E. A., Tunney, J., Goldsmith, L. T., & Seago, N. (2014). A framework for the facilitation of teachers' analysis of video. *Journal of Teacher Education, 65*(4), 340-356. https://doi.org/10.1177/0022487114534266
Van Geel, M., Keuning, T., & Safar, I. (2022). How teacher develop skills for implementing differentiated instruction: Helpful and hindering factors. *Teaching and Teacher education: Leadership and Professional Development, 1*, 1–11. https://doi.org/10.1016/j.tatelp.2022.100007
van Geel, M., Keuning, T., Frèrejean, J., Dolmans, D., van Merriënboer, J., & Visscher, A. J. (2019). Capturing the complexity of differentiated instruction. *School effectiveness and school improvement, 30*(1), 51-67. https://doi.org/10.1080/09243453.2018.1539013
van Tassel-Baska, J. & Stambaugh, T. (2005). Challenges and possibilities for serving gifted learners in the regular classroom. *Theory into Practice, 44*(3), 211-217. https://doi.org/10.1207/s15430421tip4403_5
Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education, 61*, 47-59. https://doi.org/10.1016/j.tate.2016.10.001
VanTassel-Baska, J., Xuemei Feng, A., Brown, E., Bracken, B., Stambaugh, T., French, H., McGowan, S., Worley, B., Quek, C., & Wenyu Bai. (2008). A Study of Differentiated Instructional Change Over 3 Years. *Gifted Child Quarterly, 52*(4), 297-312. https://doi.org/10.1177/0016986208321809
Villegas-Reimers, E. (2003). *Teacher professional development: An international review of the literature.* International Institute for Educational Planning. https://www.teachersity.org/files/PDF/UNESCO%20-%20Teacher%20Professional%20Development.pdf
Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2013). Mathematics teachers' beliefs. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, M. Neubrand (eds.) *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 249-271). Springer. https://doi.org/10.1007/978-1-4614-5149-5_12

Vygotsky, L. S. (1978). M*ind in society: The development of higher psychological processes*. Harvard University Press.
Waddell, L. R. (2014). Using culturally ambitious teaching practices to support urban mathematics teaching and learning. *Journal of Praxis in Multicultural Education, 8*(2), 2-22. https://doi.org/10.9741/2161-2978.1069
Wæge, K., & Fauskanger, J. (2021). Teacher time outs in rehearsals: In-service teachers learning ambitious mathematics teaching practices. *Journal of Mathematics Teacher Education, 24*(6), 563-586. https://doi.org/10.1007/s10857-020-09474-0
Wæge, K., & Fauskanger, J. (2023). Supporting in-service teachers' collective learning of ambitious teaching practices through teacher time outs. *Scandinavian Journal of Educational Research, 67*(4), 505-520. https://doi.org/10.1080/00313831.2022.2042730
Wager, A. A., Pietz, B., & Klehr, M. (2017). Providing access to equitable mathematics learning. In D. A. Spangler and J. J. Wanko (Eds.), *Enhancing classroom practice with research behind the principles to actions* (pp. 99-112). The National Council of Teachers of Mathematics, Inc.
Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education, 18*(1), 375–400. https://doi.org/10.1007/s10857-014-9286-3
Watson, A., & Sullivan, P. (2008). Teachers learning about tasks and lessons. In D. Tirosh & T. Wood (Eds.), *Tools and resources in mathematics teacher education* (pp. 109-135). Sense Publishers. https://doi.org/10.1163/9789004418967
Webb, J. N. (2018). *Conceptualizing and investigating mathematics teacher learning of practice*. [Doctoral Dissertation]. The University of North Carolina at Greensboro. https://libres.uncg.edu/ir//listing.aspx?id=23484
Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. National Staff Development Council. -
Weiss, I. R., & Pasley, J. D. (2004). What is high-quality instruction?. *Educational Leadership*, *61*(5), 24-28. https://www.ascd.org/el/articles/what-is-high-quality-instruction
Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Harvard Business School Press.
Wilhelm, A. G. (2014). Mathematics teachers' enactment of cognitively demanding tasks: Investigating links to teachers' knowledge and conceptions. *Journal for Research in Mathematics Education, 45*(5), 636–674. https://doi.org/10.5951/jresematheduc.45.5.0636
Windschitl, M., Thompson, J., & Braaten, M. (2011). Ambitious pedagogy by novice teachers: Who benefits from tool-supported collaborative inquiry into practice and why?. *Teachers college record, 113*(7), 1311-1360. https://doi.org/10.1177/016146811111300702
Winkler, I. (2020). Cognitive activation in L1 literature classes. A content-specific framework for the description of teaching quality. *L1-Educational Studies in Language and Literature, 20*(1), 1-32. https://doi.org/10.17239/L1ESLL-2020.20.01.03
Witherspoon, E. B., Ferrer, N. B., Correnti, R. R., Stein, M. K., & Schunn, C. D. (2021). Coaching that supports teachers' learning to enact ambitious instruction. *Instructional Science, 49*, 877-898. https://doi.org/10.1007/s11251-021-09536-7
Wolthuis, F., van Veen, K., de Vries, S., & Hubers, M. D. (2020). Between lethal and local adaptation: Lesson study as an organizational routine. *International journal of educational research, 100*, 1-12. https://doi.org/10.1016/j.ijer.2020.101534

Wu, S. C., & Lin, F. L. (2016). Inquiry-based mathematics curriculum design for young children-teaching experiment and reflection. *Eurasia Journal of Mathematics, Science and Technology Education*, *12*(4), 843-860. https://doi.org/10.12973/eurasia.2016.1233a

Xia, Y., Patthoff, A., Bravo, M., & Téllez, K. " We don’t observe other teachers”: Addressing professional development barriers through lesson study and video clubs. *Teacher Learning and Professional Development*, *7*(1), *1-22*. https://journals.sfu.ca/tlpd/index.php/tlpd/article/view/88

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4) 458-477. https://doi.org/10.5951/jresematheduc.27.4.0458

Yap, R.A.S.; Leong, Y.H. Using video clubs to learn for mathematical problem-solving instruction in the philippines: The case of teaching extensions. In S. Ng (ed.), *Cases of Mathematics Professional Development in East Asian Countries*: Mathematics Teacher Education (Vol. 10) (pp. 83–106). Springer. https://doi.org/10.1007/978-981-287-405-4_6

Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement.* Institute of Education Sciences, U.S. Department of Education.

Yoshida, M. (2012). Mathematics lesson study in the United States: Current status and ideas for conducting high quality and effective lesson study. *International Journal for Lesson and Learning Studies, 1*(2), 140-152. https://doi.org/10.1108/20468251211224181

Zhang, M., Lundeberg, M., Koehler, M. J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education, 27*(2), 454–462. https://doi.org/10.1016/j.tate.2010.09.015

Zimmerman, D. W. (2000). Statistical significance levels of nonparametric tests biased by heterogeneous variances of treatment groups. *The Journal of General Psychology, 127*(4), 354-364. https://doi.org/10.1080/00221300009598589