Journal of Research in Science Teaching

WILEY

RESEARCH ARTICLE OPEN ACCESS

Preservice Teachers' Professional Development for Implementing Differentiated Instruction in Science Education: An Embedded Case Study

Tessa Slim[1,2] | Johanna E. van Schaik[3] | Anna C. G. Hotze[2,4] | Maartje E. J. Raijmakers[1,5]

[1]Vrije Universiteit Amsterdam, Educational Studies, De Boelelaan 1105, Amsterdam, HV, the Netherlands | [2]iPabo University of Applied Sciences, Jan Tooropstraat 136, Amsterdam, AD, the Netherlands | [3]Radboud University, Behavioural Science Institute, Nijmegen, XZ, the Netherlands | [4]Delft University of Technology, Delft, CD, the Netherlands | [5]University of Amsterdam, Amsterdam, NK, the Netherlands

Correspondence: Tessa Slim (t.slim@vu.nl)

Received: 9 January 2025 | **Revised:** 16 December 2025 | **Accepted:** 26 December 2025

Keywords: differentiated instruction | preservice teachers | professional development | science education

ABSTRACT

The current study explored how a professional development program for differentiated instruction in science education supported change in preservice teachers, following the Interconnected Model of Professional Growth by Clarke and Hollingsworth. Among eight preservice teachers, a self-report questionnaire (pre and post) and reflective logbooks were used to identify changes in the four domains of the Interconnected Model of Professional Growth. Three out of the eight preservice teachers were additionally analyzed as embedded units using semi-structured interviews and classroom observations. The professional development program consisted of seven 2-h meetings, which included theory about differentiated instruction and science education, relevant pedagogical content knowledge, and both individual and group reflection. Two science lessons, with a focus on supporting high-achieving students and students with reading difficulties, were implemented by the preservice teachers in their internship schools and evaluated during the professional development program. For the embedded units of analysis, the preservice teachers implemented an additional follow-up lesson after completion of the professional development program. Overall, results showed that preservice teachers reported increased practice of differentiated instruction in science education, higher self-efficacy, and positive experiences with the professional development program. The use of the Interconnected Model of Professional Growth helped to illustrate aspects of the professional development program that may have contributed to change, while also nuancing the self-reported progress. The domain of consequence—specifically the ability of noticing—seems to be an important catalyst for teacher change. Taken together, these findings provide detailed insights into how teachers can be better prepared in meeting the needs of all students in science education.

1 | Introduction

There is a need for teaching differentiated primary science education, given the great variation in students' science achievement and development (Lazonder et al. 2021). Differentiated instruction is a teaching approach that accommodates differences in learners' readiness, interest, or preferred modes of learning through proactive modification of curricula, teaching, and learning activities (Tomlinson et al. 2003). Research on differentiated instruction or other adaptive forms of teaching in science education is limited yet can offer important insights that help teachers in supporting their students more effectively than

Anna C. G. Hotze and Maartje E. J. Raijmakers contributed equally to this study.

ad-hoc support (Schlatter et al. 2022; Tobin and Tippett 2014). In general, differentiated instruction is a complex teaching skill that only a few teachers master; professional development and improvement of teacher education curricula is thus desirable (Van Geel et al. 2019). Understanding how preservice teachers can be better prepared for this complex task might ultimately result in a more tailored learning environment for all students and could reduce the currently large variation in students' achievement. Employing an embedded case study design, the current study therefore focuses on preservice teachers' professional development regarding differentiated instruction in science education. Through in-depth qualitative research, this study contributes to the understanding of how differentiated instruction can be best implemented in science education and what the possible barriers or opportunities are for preservice teachers' professional development in this matter.

2 | Literature Review

2.1 | Differentiated Instruction

Rather than implementing a one-size-fits-all teaching approach, teachers should accommodate individual differences in the classroom (Parsons et al. 2018; Tomlinson et al. 2003). Various terminologies exist for this instructional approach, such as differentiated instruction, instructional adaptations, or adaptive teaching (Bosker 2005; Bowers et al. 2020; Parsons et al. 2018; Roy et al. 2013). The terms are often lumped together, yet consensus is lacking about precise conceptualizations (Lazonder 2018). In this study, we adhere to the approach and term 'differentiated instruction' (DI). We follow Bosker's (2005) distinction of DI in relation to other adaptive teaching practices, in that it is essential that DI is planned and embedded within a cycle of monitoring, execution, and evaluation. We thereby build on previous research that has shown the positive effects of DI (see e.g., the meta-analysis by Deunk et al. 2018). Furthermore, DI is a common practice in mathematics and language education, both internationally (De Jesus 2012; Smit and Humpert 2012; Watts-Taffe et al. 2012; Whitley et al. 2019) and in the Dutch educational context (Prast et al. 2018; Van Geel et al. 2019), the setting of the current study. As DI is less common in elementary science education (Kahmann et al. 2022), the conceptualization of DI mentioned below is based on the context of mathematics and reading education.

The approach of DI can include practices such as providing additional instruction, using (flexible) heterogenous or homogenous ability grouping, or offering acceleration opportunities (Deunk et al. 2018). These practices are tailored to students' needs by making modifications in the content (i.e., the knowledge and skills that a student needs to learn), process (i.e., how students learn the content), or product (i.e., how students demonstrate what they learned) (Tomlinson and Imbeau 2010). Furthermore, the implementation of DI practices can be influenced by specific characteristics of different subject areas. For instance, additional instruction might be easier to provide in mathematics lessons, while whole-class instruction may be more suited for reading lessons (Pozas et al. 2020; Ritzema et al. 2016). Importantly, DI is more than a collection of practices that follow a certain fixed formula; it is a complex skill that warrants professional development (Frerejean et al. 2021; Van Geel et al. 2019). Teachers are not only required to have knowledge about individual students (e.g., their level of achievement, interests, or peer relations), but advanced subject-matter knowledge as well (Van Geel et al. 2019).

In addition to the knowledge and skills needed to successfully implement DI, beliefs that teachers hold about their students, themselves, or education in general, are also important to consider (De Vries et al. 2013). Teacher beliefs can be defined as "the concept that grasps the conceptions or claims teachers hold" (Roose et al. 2019, 141). Beliefs can contribute to the shaping of expectations teachers have for their students and may help explain why teachers do not always provide optimal support for each student (Hornstra et al. 2023). Previous studies have shown that the implementation of DI practices is related to teachers' beliefs, such as their views on student growth mindset (Gheyssens et al. 2020) and their own sense of self-efficacy (Whitley et al. 2019). Self-efficacy, according to the social learning theory by Bandura (1977), entails the judgment about one's capability to perform a task and strongly influences teachers' professional behavior (Herrington et al. 2016). Self-efficacy is expected to develop through mastery experience, vicarious experience (i.e., observing someone else), verbal persuasion, and physiological and affective experiences (Scarparolo and Subban 2021). In the context of DI (DI in general, not specific for science education), self-efficacy has been found to play a key role in teachers' professional development (De Neve et al. 2015). The literature concerning preservice teachers' self-efficacy for implementing DI, however, is limited (Scarparolo and Subban 2021). As teachers' self-efficacy is difficult to change once it is formed, it is fundamental that preservice teachers are supported in developing a high sense of self-efficacy during teacher education (Scarparolo and Subban 2021).

2.2 | The Context of Science Education and the Need for DI

The research on DI described so far has been mainly conducted in the context of mathematics or language education. A science lesson, however, can be quite different from a typical mathematics or language lesson. In the Netherlands, where the present study is situated, science education is known as "science and technology education", as curriculum guidelines emphasize the implementation of both inquiry and design-based learning (Van Graft and Klein Tank 2018). Nevertheless, we adhere to the term "science education" in this paper, as our focus is more specifically on the teaching of science, where students are engaged in scientific practices such as asking questions, planning and carrying out investigations, and evaluating evidence across various science domains (National Research Council 2012; Teig et al. 2022).

These numerous facets associated with inquiry-based education make science teaching a demanding task for teachers (Dobber et al. 2017), and at the same time, a crucial one. Research consistently points to the essential role of the teacher and the guidance they provide during science education (Aditomo and Klieme 2020; Dobber et al. 2017; Furtak et al. 2012; Lazonder and Harmsen 2016; Minner et al. 2010). In addition, more recent research highlights the necessity of matching teacher guidance to fit individual students' needs (i.e., DI), as many students follow different learning pathways (Schlatter et al. 2021; Slim et al. 2022).

Students' different learning pathways in science education have often been linked to individual differences in reading comprehension or mathematical achievement. In the longitudinal study by Lazonder et al. (2021), 43% of students only slowly progressed in scientific reasoning over the course of 3 years and did not optimally benefit from the instruction they received. This study investigated whether students' developmental pattern in scientific reasoning would match patterns of development in reading comprehension and mathematical achievement. This was only the case for reading comprehension, with higher levels of reading comprehension being related to higher levels of scientific reasoning. This led the authors to suggest that reading comprehension scores might be used to predict development in scientific reasoning. In other studies, science learning has been found to correlate to both reading comprehension and mathematical achievement (Slim et al. 2022; Schlatter et al. 2020). In the study by Van Dijk et al. (2016, p/152), prompts in an online inquiry task were (against authors' expectations) used by the "higher ability students" as opposed to the "lower ability students". It remained speculative as to why the "lower ability students" did not use the prompts.

Thus, although scores on academic measures such as reading comprehension and mathematics seem to be indicative of students' science learning, the appropriate way to accommodate these differences between students is not self-evident. Furthermore, in our previous work, although we found a relation between the academic scores and our assessment of science learning, we also found that students with lower academic scores were successfully engaged in the science learning environment in a variety of ways (Slim et al. 2022). Some students did indicate that the written language impeded their understanding of the content. Our proposed possibilities for differentiated instruction, therefore, included varying in product (i.e., letting students decide if they want to present their results orally or in writing) and offering more language support. Although not easy to implement, DI in science education might eventually offer a better experience for both teachers and students (Slim et al. 2022).

However, research on DI in science education is limited and presents mixed results as of yet. Schlatter et al. (2022) found no effect of pre-planned adaptive science instruction (which included prompts of varying levels), but the authors pointed out that the limited learning gains in their sample stress the need for support. As specific teacher support was not included in their study, a more comprehensive view of the possibilities of teacher adaptive instruction is needed. Positive learning outcomes of DI were found in the studies by Simpkins et al. (2009) and Tobin and Tippett (2014). Simpkins et al. (2009) examined effects of peer tutoring support, while Tobin and Tippett (2014) examined tools such as choice boards or graphic organizers. Although these studies provided valuable insights into various ways of DI in the context of science education, more understanding is needed regarding what makes teachers change their teaching and truly adopt DI in science education. For example, Eysink et al. (2017) investigated the effects of an intervention focused on supporting teachers with differentiated instruction. They found that there was considerable variation among teachers in their implementation of the intervention. This intervention included a jigsaw procedure: in a first lesson, students were grouped according to their performance ("general" and science), and each group worked on assignments and experiments specific to their topic. In a next lesson, students were regrouped heterogeneously, such that topics were mixed, and each student shared what they had learned in the first lesson. The findings of this study are complex, but key insights include that student learning outcomes were affected by the way teachers implemented DI, and that teachers did not transfer the DI practices into their regular science lessons. Since the teachers did not receive formal training in DI, an important remaining question is how teacher education can support the process of implementing DI in science education.

As discussed above, one facet known to affect teachers' actual implementation of DI is self-efficacy, but this has yet to be studied in the context of science education. Although literature on preservice teachers' self-efficacy beliefs regarding DI (in general) is limited (Scarparolo and Subban 2021), it is often established that preservice teachers experience low self-efficacy beliefs for science teaching (Gunning and Mensah 2011; Knaggs and Sondergeld 2015; Menon and Sadler 2016; Seung et al. 2019). As DI is a complex skill even for experienced teachers (Van Geel et al. 2019) and Dutch teachers' self-efficacy in science education is low (Velthuis et al. 2014), a logical conjecture would be that preservice teachers also experience low self-efficacy regarding DI in science education. To improve teacher education, more empirical insights are needed to understand preservice teachers' self-efficacy regarding DI in science education.

3 | The Present Study and Theoretical Framework

Following the conclusions of research thus far, it is important to understand how teachers can best be supported when implementing DI in science education. In the present study, we focus specifically on preservice teachers and their professional development in DI. Although Frerejean et al. (2021) believe that novice teachers might be too inexperienced for the task of DI, other studies have pointed out that teacher education should lay the foundation of professional development regarding DI. For example, Tomlinson et al. (1994) stress the importance of setting a precedent for teachers in making DI a priority and preventing a status quo of a one-size-fits-all approach. Furthermore, teacher beliefs, which influence the implementation of DI, develop during teacher education and are difficult to change once they are formed (Griful-Freixenet et al. 2021; Scarparolo and Subban 2021). De Neve and Devos (2016) showed that the more novice teachers indicated that their beliefs regarding DI had positively changed due to their teacher education, the more their practice in DI had changed.

The goal of the current study was to develop and examine a professional development program (PDP) for preservice teachers regarding DI in science education. We sought to explore preservice teachers' experiences with a PDP and provide a detailed description of their professional development to help inform teacher education on how to prepare teachers for DI in science education.

We constructed our study within the theoretical framework of social constructivism and the model of teacher professional development by Clarke and Hollingsworth (2002).

First, both science education as well as the teaching approach of DI are often linked to Vygotsky's theory of social constructivism, in which learners are thought to actively and socially construct knowledge in interaction with others, including the teacher (Parsons et al. 2018). A central notion of this theory is that a teacher should ideally offer support attuned to a student's current level of understanding (Subban 2006). As levels of understanding are expected to differ between students, teachers should thus offer differentiated support. In science education, attuning to students' understanding often involves responding to students' initial conceptions of the subject and supporting them in building coherent scientific models or theories (Bächtold 2013). This requires the noticing of meaningful learning opportunities (Chan et al. 2021).

Second, for teachers' professional development, we first consulted Boylan et al. (2018) analytical framework of five significant contemporary models regarding teacher professional learning in order to support an informed selection of models. Using this analytical framework and corresponding key questions, we selected the theory of teacher professional growth by Clarke and Hollingsworth (2002) to guide our development and analysis of a PDP for preservice teachers. The main reason for selection was the detailed perspective on teachers' professional learning and the change processes within this theory. We were interested in these processes, since a better accommodation of students' needs in science education requires a change in teacher behavior. Changing teacher behavior is a complex process that involves teacher learning (Clarke and Hollingsworth 2002), and previous research shows that teachers have difficulty with implementing DI (Van Geel et al. 2019), specifically in science education (Eysink et al. 2017). A deeper understanding of how preservice teachers learn to implement DI in science education thus helps to understand when and how their instructional behavior might change. In Clarke and Hollingsworth's (2002) interconnected model on teacher professional growth (IMPG), depicted in Figure 1, change is suggested to occur through reflection and enactment in four distinct domains: external (sources of information, stimulus or support), practice (professional experimentation), consequence (salient outcomes), and personal (knowledge, beliefs, and attitude). The external domain is a different type of domain than the others, as it is located outside of the teacher's personal world. The other three domains, in combination, "constitute the individual teacher's professional world of practice, encompassing the teacher's professional actions, the inferred consequences of those actions, and the knowledge and beliefs that prompted and responded to those actions" (Clarke and Hollingsworth 2002, 951). When change in one domain leads to change in another domain, supported by empirical data, Clarke and Hollingsworth refer to this as a 'change sequence'. Change sequences can entail momentary or fleeting instances of change; if the change is more lasting, the term professional growth can be used.

In alignment with this theoretical framework, our main research question is: How does a professional development program for differentiated instruction in science education support change among preservice teachers in the four domains of the Interconnected Model of Professional Growth?

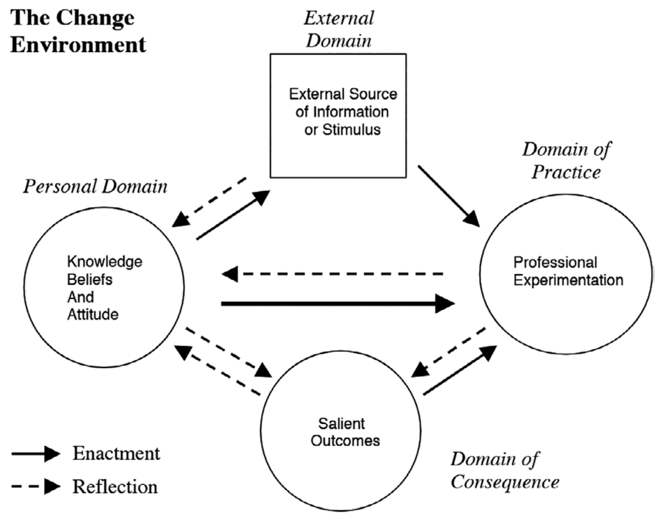

FIGURE 1 | The interconnected model of professional growth (Clarke and Hollingsworth 2002, 951).

4 | Methods

4.1 | Design and Sample

An embedded single-case study design, using mixed-methods analyses, was employed (Yin 2018). The single case in this study involves the professional development of preservice teachers who follow a PDP that is aimed at implementing DI in science education. We studied the phenomenon of professional development as a single case, rather than multiple cases of preservice teachers, as the professional development is bound within the specific context of our PDP. Including a subset of the preservice teachers as embedded units of analysis provided more in-depth sources of evidence, allowing for a more detailed exploration of our case.

To further bound our case (Cleland et al. 2021), we located our study in a class of fourth-year preservice teachers of a teacher education college in the Netherlands. This teacher education college is located in the capital city of the Netherlands and is attended by ~1200 students every year (Hogeschool IPABO 2021). In the Netherlands, teacher education colleges prepare preservice teachers to teach 4–12-year-olds in all subjects, including science. In this specific teacher education college, fourth-year students follow one of three specialized tracks. They are required to spend 80 h on professional development within their specialization. We offered the PDP as a means to fulfill this requirement. A class of 25 preservice teachers who followed the specialization of inquiry-based learning was asked whether they would like to participate in the PDP and corresponding case study, of which nine chose to do so. Participation was entirely voluntary and based on preservice teachers' interests and availability (there were specific date and time requirements). The remaining preservice teachers opted to fulfill their professional development requirement through one of the other available options. Recruitment was non-directive, with no targeted outreach to specific individuals, and there was no selection process. All preservice teachers provided informed consent before participating in the study.

In sum, nine female preservice teachers participated in this study[1] ($M_{age} = 21.88$ years). One preservice teacher did not complete the post-test measures for personal reasons and has therefore been excluded from analysis. Out of the eight remaining preservice teachers, three teachers were analyzed as embedded units of analysis (pseudonymized names: "Lisa", "Anne", and "Iris"). These teachers were purposefully selected based on their pre-test questionnaire scores. Specifically, we included teachers that maximally differed from one another in their self-efficacy, use of DI practice, and prior experience with science education. Their individual scores are presented in Table 1.

4.2 | The Professional Development Program

A timeline of the PDP and research design is depicted in Figure 2. The PDP ran from October 2022 through January 2023, and consisted of seven 2-h group meetings. The first author of this study taught the PDP. Consistent with the theory of Clarke and Hollingsworth (2002), the program was designed to provide input on all four domains of the IMPG. In Table 2, an overview is provided of the PDP's components, learning objectives, and relation to the four IMPG-domains. With respect to taught content, we presented the theories and findings regarding DI in general and DI in science education presented in the introduction of the present paper. As research on DI in science education is limited, we discussed the possibilities and difficulties of implementing DI in science education based on research on individual differences in science education. After discussing the theory, the preservice teachers prepared two differentiated science lessons in the group meetings, enacted them in their internship schools, and evaluated them together in a group meeting.

The two science lessons were based on existing lesson plans that were developed within the ENGINEER project in collaboration with a Dutch science museum (Anyfandi et al. 2016; Nemo Science Museum n.d.), and focused on the concepts of sound and balance. In the preparation of the science lessons during the PDP, attention was devoted to the pedagogical content knowledge of the preservice teachers regarding these concepts and students' initial concepts and conceptual change. The specific DI practices that preservice teachers were instructed to implement were targeted at supporting students with reading difficulties, that is, "reading support", and at challenging students who are high-achievers in math, that

TABLE 1 | Questionnaire results for embedded units of analysis.

	Lisa		Anne		Iris	
Age	20		20		22	
Number of science lessons taught before PDP	> 10×		1–3×		3–10×	
Science curriculum	Project-based		"Blink"		"Blink"	
DSAQ scales (range 1–5)	Pre	Post	Pre	Post	Pre	Post
Identifying educational needs	1.40	1.60	1.80	2.60	2.40	3.20
Differentiated goals	1.17	3.00	2.83	3.50	2.00	2.50
Differentiated instruction	2.00	4.00	3.00	3.00	3.43	3.71
Differentiated practice	1.25	2.13	2.63	2.88	2.50	3.38
Evaluation of progress and process	1.57	2.43	2.29	2.86	2.29	3.86
Self-efficacy	2.25	4.00	3.00	4.25	3.00	4.75
Attitude	3.25	4.50	3.25	3.75	4.50	3.75

Note: Names are pseudonyms. "Blink" is the name of a curriculum used in the Netherlands for science education.

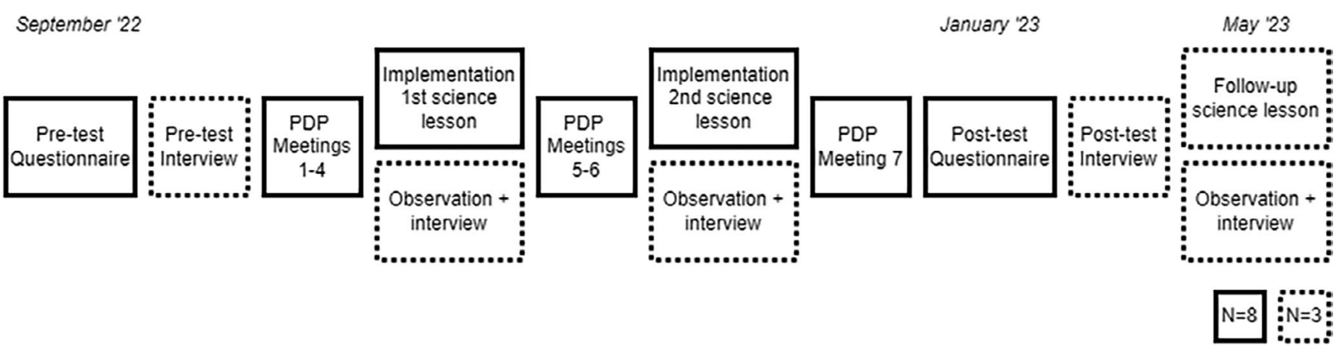

FIGURE 2 | Research design.

TABLE 2 | Overview of PDP components and learning objectives.

Meeting	Main components	Learning objectives
1	Theory on DI in general (external domain)	• Know what DI entails, which phases can be distinguished, and which teacher skills are important • Know several factors that determine the complexity of DI • Know what the ethical dilemma regarding DI entails and reflect on their own vision about it
2	Theory on DI in science education (external domain)	• Know individual differences that are important to consider in science education • Know several applications of DI and teacher guidance that can be implemented in science education • Know what the difficulties and possibilities regarding DI in science education are
3	Preparation for first science lesson regarding the concept of sound, including pedagogical content-knowledge regarding the concept of sound (external domain and personal domain)	• Being able to formulate (inquiry-based) learning goals regarding the concept of sound • Know what fair experimenting entails • Made a draft outline for the first science lesson
4	Continued preparation for first science lesson regarding the concept of sound (external domain and personal domain)	• Prepared for supporting students with reading difficulties when they are experimenting with sound • Prepared for asking enrichment questions to students who are experimenting with sound
Between 4 and 5	Enactment of first science lesson (domain of practice)	
5	Evaluation of first science lesson (external domain, personal domain, domain of consequence)	• Reflect on own DI skills • Evaluate whether students have met learning goals • Formulate learning goals regarding DI skills
6	Preparation for second science lesson, including pedagogical content-knowledge regarding the concept of balance (external domain and personal domain)	• Being able to formulate (inquiry-based) learning goals regarding the concept of balance • Know what different strategies (rules) students apply when reasoning about balance • Prepared for supporting students with reading difficulties when they are experimenting with balance • Prepared for asking enrichment questions to students who are experimenting with balance
Between 6 and 7	Enactment of second science lesson (domain of practice)	
7	Evaluation of second science lesson and PDP in total (external domain, personal domain, domain of consequence)	• Reflect on own DI skills • Evaluate whether students have met learning goals

is, "asking enrichment questions", based on previous research on individual differences in science education (Lazonder et al. 2021; Schlatter et al. 2020; Slim et al. 2022). Reading support consisted of explanations of the difficult vocabulary used in the worksheet. The enrichment questions were designed to deepen conceptual understanding and to stimulate reasoning about the experiments presented in the lesson. To allow for meaningful comparison, the implementation of DI was held consistent for all preservice teachers. In line with the principle of 'practice what you preach', the course content and materials of the PDP were continuously adapted to meet the needs of the preservice teachers.

4.3 | Data Collection and Instruments

Within the main sample of eight preservice teachers, a pre- and post-test questionnaire and reflective journals were used as measurement instruments. For the three preservice teachers included as embedded units of analysis, additional interviews and observations were included. They were interviewed before and after the PDP and observed during the two science lessons they had to teach, with interviews conducted immediately after these lessons as well. Around 4 months after completing the PDP, they taught a self-prepared science lesson, where they were free to choose how to implement DI. This lesson was also observed and followed by an interview. The first author distributed the questionnaire and conducted all interviews and lesson observations. Duration of interviews was around 20–30 min and was audio-recorded.

4.3.1 | Instruments

4.3.1.1 | Questionnaire. The questionnaire consisted of two parts, providing input for different domains of the IMPG: the first part focused on the domain of practice and the second part on the personal domain. Only these domains were included in the questionnaire because (1) the domain of consequence was measured through reflective journals, observations, and interviews and (2) the external domain consisted of input provided throughout the PDP, as described in Table 2. No other sources of external input were questioned, as data collection was already extensive. The reflective journals, observations, and interviews did not provide any indication that preservice teachers sought other external sources of information apart from the PDP.

The first part contained the Differentiation Self-Assessment Questionnaire (DSAQ; developed for DI in math, Prast et al. 2015), which we used to measure preservice teachers' self-reported practice regarding differentiated instruction in science education, before and after the PDP, replacing the word 'math' with 'science' where needed. The DSAQ consists of 33 items across five subscales: identification of educational needs, differentiated goals, differentiated instruction, differentiated practice, and evaluation of progress and process. All items are included as Supporting Information (Table S1). Each subscale contains 5–8 items, for which preservice teachers could rate their agreement on a Likert scale ranging from 1 (does not apply to me at all) to 5 (fully applies to me). The subscale 'differentiated goals' consists of three items that refer to curriculum materials, such as 'I know the opportunities for differentiation offered by the curriculum materials'. The wording of these items causes some trouble for interpretation. For example, it is unclear if a low score means that a preservice teacher does not know about the opportunities provided by the curriculum, or that the curriculum does not offer any opportunities for differentiation. We therefore included the option for preservice teachers to leave these three items unanswered if they were certain the curriculum materials did not offer any opportunities for DI and a low score would not be fitting to the situation. Reliability statistics for the DSAQ were excellent for the pre-test items (Cronbach's $\alpha = 0.95$) and acceptable for the post-test items (Cronbach's $\alpha = 0.76$).

The second part of the questionnaire focused on self-efficacy beliefs and attitude, as components of the personal domain. For self-efficacy beliefs, 8 items derived from Tomlinson et al. (1994) were included, with the addition of "science education" or "science ability", such as 'How confident do you feel about accommodating varying levels of science ability in your class'. For attitude, 5 items from the Dimensions of Attitude toward Science (Van Aalderen-Smeets and Walma Van Der Molen 2013) were included, such as 'I enjoy teaching science very much'. All items consisted of a 5-point Likert scale. Pre-test self-efficacy items (Cronbach's $\alpha = 0.78$) and attitude items (Cronbach's $\alpha = 0.70$) were acceptable. Post-test reliability statistics were questionable for the self-efficacy items (Cronbach's $\alpha = 0.62$) and poor for attitude items (Cronbach's $\alpha = 0.59$). All items are included as Supporting Information (Table S1). While a self-report questionnaire has known limitations, such as susceptibility to bias, it is a valuable tool to assess processes of learning (Pekrun 2020). To strengthen the credibility of findings and minimize the limitations associated with self-reported data, triangulation of data was sought by including other (self-report) measures as well.

4.3.1.2 | Reflective Journals. To capture more detailed changes in the four domains of the IMPG, all preservice teachers filled out an online reflective journal throughout the PDP. These journals contained questions that prompted them to reflect on the provided theory and literature, their beliefs about teaching and learning, their self-efficacy, expectations of students, the learning goals of the group meetings, the two science lessons (or three for the embedded units of analysis) that they carried out in their classrooms, and finally, their experience with the PDP (including evaluative questions on a 7-point Likert scale). In addition, preservice teachers were asked to provide their lesson plans within the online journal. Questions in the reflective journal that were not relevant for our research question, for example regarding their DI practice in mathematics education, were not included in our analysis.

4.3.1.3 | Pre- and Post-Interviews. The three preservice teachers were interviewed before and after the start of the PDP (the first and the last group meeting). These interviews were intended to deepen and clarify the insights derived from the questionnaire and could provide information on possible changes for all four domains of the IMPG. The DSAQ items and questionnaire results were used to structure the interview protocol (semi-structured). The protocols for both the pre- and post-test are included as Supporting Information (methods_supplement2). Interviews took place at the college or online.

4.3.1.4 | Lesson Observations and Interview. Additionally, these three preservice teachers were observed during the two science lessons they taught during the PDP and the one after completion of the PDP. The observations were subsequently followed by another interview, which also took place in the preservice teachers' internship schools. Both the observation and interview were guided by the ADAPT instrument (Keuning et al. 2022). This instrument is based on a cognitive task analysis of DI (Van Geel et al. 2019) and was specifically developed to assess the match between teachers' practices and students' needs. It consists of 23 indicators that can be scored from 1 to 4, based on a lesson observation, interview, and relevant documents such as lesson preparation. Additional observational notes were made when necessary.

The instrument and indicators served as a helpful framework for our observations and subsequent interviews, as they provided input to establish possible changes in the four domains of the IMPG. The scores on the indicators, however, were not relevant for our research question and are therefore not included in analyses (included as Supporting Information, Table S2).

4.4 | Analyses

4.4.1 | Quantitative Analysis

The questionnaire data will be presented in a box plot graph, including error bars. Given the small sample size and risk of inflation of family-wise error due to the use of repeated statistical tests, we will not examine statistical significance in differences from pre to post. For a quantitative evaluation of how preservice teachers felt the PDP supported them in implementing DI in science education, we will describe the descriptive results of seven Likert-scale questions that were asked within the online journal. For all five DSAQ-subscales, preservice teachers indicated how the PDP supported them in the implementation of DI. For the subscale 'implementing differentiated instruction', we additionally discerned between 'the intensive group' (i.e., students who have difficulty obtaining the learning objective) and 'the enrichment group', resulting in a total of seven items.

4.4.2 | Qualitative Analysis

Our qualitative analysis aimed to extract themes that focus on preservice teachers' experiences with the PDP, "to identify how these relate to changes in particular domains" (Wilkie and Clarke 2015, 100). We used a reflexive thematic analysis approach (Braun and Clarke 2019), which is inherently shaped by our own subjectivity and theoretical assumptions. Our analysis included the journals of all eight preservice teachers, clean verbatim interview transcripts, and observational data of the three embedded cases. The software package Atlas.ti 23 was used to organize the data. The first step, conducted by the first author, involved familiarization with the journals and the interviews by reading them iteratively (in chronological order, per preservice teacher, but also per question or per meeting), while making notes and mind maps. A summary was made for each PDP meeting, lesson evaluation, and embedded unit of analysis. Holistic coding was applied in the next step to the journals of all eight preservice teachers and to the interview transcripts of the three preservice teachers as embedded cases to further organize and structure the data (Saldaña 2009). The four domains of IMPG were applied as additional codes to the existing segments. Themes were constructed in consultation with all authors, while going back and forth in the data and discussing the segments within the holistic codes, the summaries of notes, mind maps, questionnaire data, and observational notes. Comparing the three selected preservice teachers, for which rich examples were available, allowed us to refine the themes and discuss specific changes along the four IMPG domains. In that way, the IMPG served as a lens in the interpretation of the data (Ketelhut et al. 2020). We specifically looked at salient (lack of) changes in the four domains of the IMPG in relation to each theme to understand when and why teachers' professional development may have occurred through these changes in the four domains. We discussed the possible pathways of changes while reviewing the data as a whole. Again, we went back and forth over all eight participants to compare changes in the different domains (personal, external, practice, or consequence). We examined the interactions of the domains more closely by analyzing the individual change sequences of the embedded cases.

4.5 | Reflexivity: Researcher Position in the Field

The practice of reflexivity entails a self-conscious evaluation of how subjectivity and context influence the qualitative research process (Olmos-Vega et al. 2023). The first author of this article was also the instructor of the PDP. For transparency purposes, it is important to briefly consider what this dual role embodied and the ways in which it may have shaped the study's research process (Nikkanen 2019).

Together with one of the co-authors of this paper, AH, the first author was an instructor for the fourth-year preservice teachers (the other two co-authors were not involved in any way with the preservice teachers). Both instructors were involved in grading the preservice teachers' assignments for their teacher education program. The PDP was part of a larger range of activities preservice teachers could follow as their professionalization requirement. This research study and the corresponding PDP were explained in the same manner as the other activities. If preservice teachers were interested, they received an information letter about the research studies and informed consent. This information letter was reviewed by the local ethics committee. Data collection was conducted by the first author. Before lesson observations and interviews took place, it was emphasized that the goal of data collection was to capture the preservice teachers' experiences rather than scoring or grading teacher behavior. However, power dynamics inherent to the dual role of instructor and researcher may have influenced preservice teachers' participation. For example, preservice teachers may have been less inclined to reflect about their shortcomings as a teacher.

Furthermore, challenges faced by the preservice teachers, as indicated by the preservice teachers themselves or by the first author during classroom observations, were used to tailor the content of the PDP and accommodate the preservice teachers' needs. The role of the first author as instructor was thus intricately related to the role as researcher. An example of how this influenced one another was that the preservice teachers expressed their concern about lacking content knowledge, which may have influenced the first author during subsequent observations. Another example is that the observation of both a preservice teacher who adequately responded to a thinking error made by students as well as a preservice teacher who did not respond to students' surprise during an experiment made the first author focus and reflect more on noticing as a means for teachers' professional development (Preminger et al. 2024). The first author discussed these observations with the other authors of this study, read more literature regarding this topic, and incorporated it into the PDP meetings. If the observations of the first author would have focused on other aspects of teaching, it is likely that a different change process would have been described. Nevertheless, by including rich descriptions of data and triangulation of sources (journals, questionnaires, interviews, and observations), the authors contend to have provided a comprehensive and valid overview of the underlying change processes in this study.

5 | Results

5.1 | Quantitative Results

Results for the pre-test to post-test differences on the questionnaire are depicted in Figure 3. Results for the three preservice teachers are shown in Table 1. The results for the evaluative questions about how the preservice teachers believed the PDP supported them are depicted in Table 3. Averages for

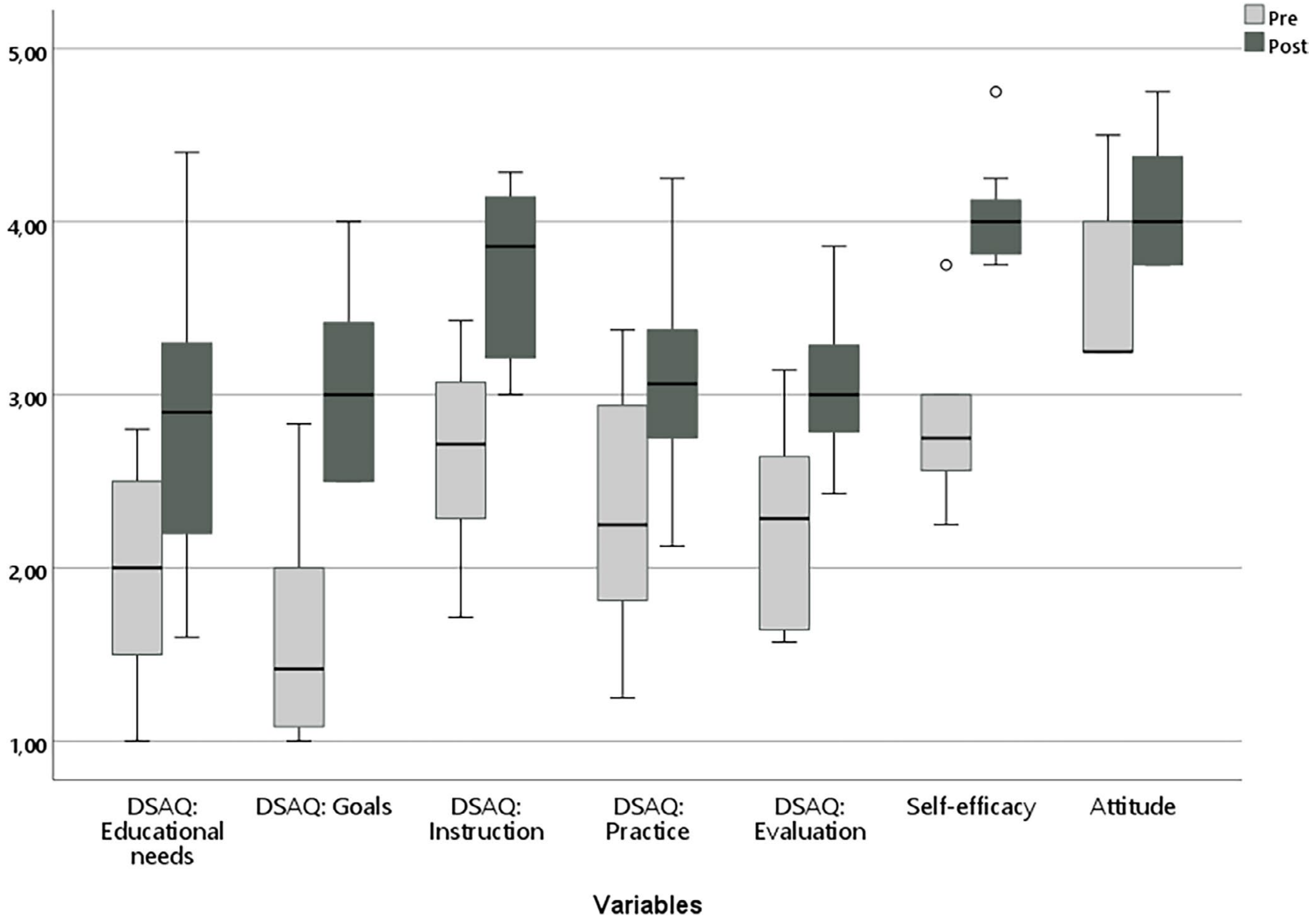

FIGURE 3 | Questionnaire results pre-post.

TABLE 3 | Individual results and mean scores of evaluative questions concerning the PDP.

On a scale of 1–7, how much did the PDP support you in learning to…	101	103	104	105 (Lisa)	106	107	108 (Anne)	109 (Iris)	M
Identify educational needs	5	6	5	6	5	5	6	6	5.5
Set differentiated goals	6	5	5	6	6	7	6	6	5.9
Implement differentiated instruction	5	5	6	3	6	7	7	6	5.6
… for the intensive group	6	5	4	4	7	7	6	6	5.6
… for the enrichment group	5	6	5	4	6	7	6	6	5.6
Implement differentiated practice	5	6	5	3	5	7	5	5	5.1
Evaluate progress and process	5	5	5	2	5	7	6	5	5.0
Mean	5.3	5.4	5.0	4.0	5.7	6.7	6.0	5.7	5.5

the individual subscales are all 5 or above, indicating that preservice teachers valued the PDP and felt supported in improving their implementation of DI practice in science education. They felt most supported in setting differentiation goals (5.9), which corresponds to the DSAQ scores, where this subscale showed the most progression.

5.2 | Qualitative Results

Five main themes were derived from the thematic analysis: (1) Lack of experience, (2) Asking enrichment questions, (3) Identifying needs, (4) Perception of DI practices, and (5) Student involvement. We will first discuss the themes for our overall case and show how the PDP supported changes in domains of the IMPG, with illustrative quotes from the total sample. Next, we will explore this in more detail by illustrating the different change sequences of the three selected preservice teachers. We conclude with a summary of changes for all three preservice teachers.

5.2.1 | Changes in the Domains of the IMPG for the Overall Case

5.2.1.1 | Lack of Experience. A general underlying theme is a lack of experience in teaching science and therefore also in implementing DI in science education. This lack of experience is reflected in the personal domain, domain of practice, and domain of consequence. Regarding a change in the personal domain, preservice teachers indicated they gained knowledge about DI in science education, which they lacked at the start of the PDP. This theme is supported by the quantitative self-reported score for self-efficacy beliefs (found in Figure 3), which increased substantially: the average was below 3 (with 1 representing no confidence and 5 very confident) at the start of the PDP, but more than 4 at the end of the PDP.

With respect to changes in the domain of practice, many preservice teachers, according to the questionnaire results, had little experience with formulating (differentiated) learning goals. The PDP seemed to elicit a change in this. Anne explained:

> I've learned to always take a moment to review the learning goal. (…). I didn't do that before. It used to be just the task, and that was it.

When more closely examining how the preservice teachers formulated learning goals throughout the PDP, there seemed to be a shift from abstract or vague goals to more concrete goals regarding student learning, which also relates to the domain of consequence (see next paragraph). For example, for the first lesson, most preservice teachers formulated general learning goals regarding skills, such as "The students conduct an experiment together" (104, journal 4) or "The students investigate materials and scientific phenomena" (Anne, journal 4). For the second lesson, more concrete learning goals were formulated such as: "Students make use of formulations such as 'If …, then …' or 'Because …'" (108, journal 7).

A more specific focus on student learning, as a result of gaining experience, was also related to the domain of consequence. It seemed that some preservice teachers were not yet accustomed to assessing what students had actually learned. For example, after lesson 1, one preservice teacher initially reflected negatively on what students had learned:

> The students didn't have any prior knowledge. (…) They didn't have the right inquiry attitude. Conducting an experiment is hard enough as it is. They're not ready for enrichment yet
>
> (107, journal 5).

After examining the students' worksheets, which was part of a group meeting within the PDP, she altered her perception:

> It was not so bad after all. (…) Learning goals have been met, now that I have looked at students' worksheets
>
> (107, journal 6).

Interestingly, in the evaluation of the next lesson, this teacher's initial reflection was more focused on her own actions rather than the belief that students lacked prior knowledge:

> The students had to dig deep for their prior knowledge. I could have prepared better questions to follow up with
>
> (107, journal 8).

5.2.2 | Asking Enrichment Questions

One of the specific types of DI that preservice teachers had to implement, as part of the PDP, was challenging high-achieving students by asking enrichment questions. This practice was designated as a theme, as it was an important stimulus for reflection and change. First, it was a change in practice for many preservice teachers, since they had no prior experience with this. They often struggled:

> Asking challenging questions. I still find this difficult to implement. Is this question not too hard? I only think about the questions I want to ask the "low" achieving students, in order for them to have an experience of success
>
> (104, journal 2).

Varying reasons seemed to be the cause of this struggle: lack of content knowledge or low self-efficacy (related to personal domain), insufficient time, too much focus on maintaining an orderly class rather than student learning ("*I didn't want to break their concentration*" (Iris, journal 6)), or not being able to recognize who is ready for challenging questions, that is, identifying needs (see next theme). Several preservice teachers stated that the PDP supported them in this practice:

> We learned a lot about asking enrichment questions. We've also practiced this further, through role-play. This really helped me and I realise that I'm becoming more successful at this
>
> (106, journal 9).

> I've gotten to know a lot about enrichment questions and DI for high-achieving students. I still have plenty of room to grow in this, but thanks to this PDP I now know how to formulate enrichment questions for these students
>
> (Iris, journal 9).

For some preservice teachers, a change in the domain of consequence (salient outcomes) was noticeable as a result of this changed practice. For example, Lisa states:

> During the lessons about sound and balance, I actually noticed that there are students in my class who are ready for enrichment questions
>
> (journal 9).

This may have led to a (strengthened) belief that asking enrichment questions is of added value, which is explained in more detail within the embedded unit of analysis Lisa.

5.2.2.1 | Identifying Needs. A recurrent theme throughout the PDP was identifying needs, which preservice teachers also struggled with. This was related to the domain of practice, as preservice teachers changed their practice in using mathematics and reading achievement as indicators of who might need additional support or challenge. Their reflection about this practice could lead to certain beliefs about how indicative mathematics and reading scores are for science education. The general consensus seemed to be that these scores provide some direction for assessing needs, but are not absolute:

> Math and reading ability are definitely useful for inquiry-based learning. But still, you'll see some students with other strong points that emerge
>
> (107, journal 8).

This is also reflected in the DSAQ results (as shown in Figure 3); identifying educational needs improved but the low score still indicated that preservice teachers do not implement this practice often.

5.2.2.2 | Perception of DI Practices. Reflection on salient outcomes (domain of consequences), as a result of changed practice, showed how some preservice teachers perceived certain DI practices negatively (personal domain). Preservice teachers practiced with ability grouping and providing additional instruction to students with reading difficulties and high-achieving students. For the students with reading difficulties, some preservice teachers felt that providing additional instruction was demotivating. In the first lesson that preservice teachers implemented, they used this separate additional instruction to provide reading support while other students already started working on the experiments. Some preservice teachers felt uncomfortable with this practice, inferring that students had a negative experience:

> *The students didn't enjoy having to wait before starting with the assignment. One of the reactions was: "Why always me?"*
>
> *(106, journal 5).*

After discussing this in a following PDP session, it was decided for the second lesson that the preservice teachers would provide the reading support separately to pairs of students during experimenting or at an earlier time in the day (pre-teaching). Some preservice teachers preferred this method, even though it cost more time, as they inferred that students felt less "left-out". A preservice teacher that contrarily preferred a separate instruction group based on reading levels, did so because of time and to be able to help students more efficiently. She also mentioned, however, that students might prefer figuring it out for themselves first, before they receive support.

5.2.2.3 | Student Involvement. Changes in student involvement were part of the consequences regularly discussed by the preservice teachers (domain of consequence). These outcomes were not always specifically related to new DI practices, but more generally related to the teaching of the two science lessons as part of the PDP. Overall, preservice teachers mentioned that students were actively involved and enthusiastic as a result of the science lesson. However, preservice teachers also often used terms as 'chaotic', 'rowdy', or 'messy' to describe students' behavior. They felt science lessons were too "free" (i.e., unstructured) for some students:

> He finds it difficult to participate in activities that have a lot of freedom. He will then act disorderly or distract other students
>
> (101, journal 5).

These are important changes in the domain of consequence for DI in science, as they might influence how teachers subsequently accommodate the perceived differences. Sometimes, preservice teachers would expect some students to act disorderly, who would then surprise them:

> My expectation was that he would slack off, which was not the case. He was very curious and asked a lot of questions. He was also motivated while experimenting. His reasoning wasn't always correct, but he really tried and enjoyed doing it!
>
> (106, journal 6).

5.2.3 | Changes in the Domains of the IMPG of the Three Preservice Teachers

We will explore these five themes and (absence of) changes in each of the IMPG domains, and their interaction, in more detail, by describing the change sequences of the three preservice teachers throughout the PDP. We will first report each individual trajectory during the PDP, after which we will provide an overall summary of the follow-up lessons the preservice teachers taught 3 or 4 months after the PDP. We conclude with a cross-case analysis. All names are pseudonyms.

5.2.3.1 | Lisa. Lisa is a 20-year-old female who has taught a relatively high (compared with the other preservice teachers) number of science lessons (more than 10). Her pre-test scores on the DSAQ are relatively low, almost all under 2 ("does not apply to me"), as well as her self-efficacy for DI practices, while her attitude toward teaching science is reasonably positive (as shown in Table 1). One of the things she struggled with before the start of the PDP was identifying educational needs and setting differentiated goals. In the pre-interview she said:

> The best thing to do would be to adapt [the goals] to students' needs. Some students will be ready for a more challenging goal, and others will have to take a step back. Right now, I'm not doing that yet, because I haven't figured out what the division could be.

In the post-interview, there was a clear change:

> During the lessons that I had to teach (...), I really looked into what do I think is a main goal? (...) Then you can always look at: how can I maybe simplify it for students who will not reach that goal? And more challenging goals (...).

She now felt she had sufficient insight into the knowledge, skills and attitude of the students to be able to establish different groupings. This is also reflected in her self-efficacy score on the post-test questionnaire (from 2.25 to 4).

Lisa valued using mathematics and reading scores as a tool for identifying needs. She felt that generally only students who are high-achievers in mathematics are ready for enrichment in science, with some exceptions for students who are very enthusiastic or have an attitude of inquiry. When asked if she feels the enrichment is too difficult for other students, she responded:

> Yes. Today as well, when I look at the questions they had to answer for the quiz, I really think: 'Wow, suddenly it's 10 steps too far for most of them. While the high-achieving kids (...), they are really ready for that enrichment. If you change everything all of a sudden, and start to ask all those questions, I think they get stuck in their heads, the middle group and especially the children with reading difficulties. I think it is too complex for them, or too far from their experience
>
> (interview; lesson 2).

In the first lesson that had to be carried out as part of the PDP, Lisa provided the additional instruction to students with reading difficulties while they were still seated at their own desk, with other students beside them already working on the assignment. She felt that this didn't work, as she couldn't engage the students with her additional instruction, so she forfeited the instruction sooner than she had intended. Furthermore, she asked enrichment questions to two pairs of students with high mathematics achievement, which was effective in her opinion;

> I think that maybe they considered their process again, like 'What did we actually do? Was this a fair experiment?' (...) The other [high-achieving math] students, where I focused more on enrichment regarding knowledge, I made them think more (...) Making connections between concepts or stuff they had to do last time
>
> (interview; lesson 1).

Lisa seemed to have a focus on students' learning of concepts and how enrichment questions might support this. This was especially noticeable in lesson 2, where students were experimenting with balancing a ruler, which hung horizontally off the side of a table on a thread, while placing clothespins on both sides of the ruler. By placing a clothespin on one side of the ruler closer to the fulcrum than the other side, the balance of the ruler is skewed to one side. When asked to explain this, many students stated: "because one side is heavier", ignoring the influence of distance. Lisa asked two students, who had provided that answer, to put the clothespins in their hands, one on each side. "Is one side heavier or are they both the same weight?", she asked. When the students answered that both are the same weight, she followed-up with the question: "But if both clothespins are the same weight, how can one side of the ruler be heavier?". This question seemed to stimulate these students to further experiment with the ruler. In reflecting on this occurrence, Lisa says:

> These girls eventually came to the insight that distance also has something to do with balance. (...) I definitely saw them in deep thought. That's nice. That's the goal of enrichment, that they will think some more about what they are doing. So it's a nice addition.
>
> (interview; lesson 2).

Lisa's noticing of the consequences of the practice of asking enrichment questions thus seemed to lead to (or strengthen) a belief about the added value of this practice.

5.2.3.2 | Anne. Anne is a 20-year-old female, with, compared with Lisa, less experience but higher scores on the pre-DSAQ (between 2 and 3 on average, as shown in Table 1). Pre to post changes in DSAQ scores were small, while Anne's evaluation scores in her reflective journal on how the PDP had supported her learning to implement DI were high (average of 6 on 7-point Likert scale items). In the first interview, Anne indicated that she had only taught one science lesson so far. She felt she needed to gain more experience with science education in general before implementing DI practices. However, she did also feel confident in being able to apply DI in science education, for example, using additional instruction groups as she does with mathematics education. In order to achieve this, she felt she needed more insight into students' needs, examples of differentiated lessons, and more time during a lesson to see and support all students.

The support she envisioned seemed to be mainly focused on students' process and inquiry skills as opposed to conceptualization, which is something that did not seem to change over the course of the PDP. In the pre-interview, she mentioned:

> I don't necessarily think they need to learn, for example, why a paper clip floats or something like that. I don't think that's important. (...) It's about their ability to work with such a teaching method, I think. I believe that's what my class needs most.

After the first lesson, in which Anne implemented the DI-practices that were part of the PDP, she was surprised at her students' capability:

> I asked the two pairs enrichment questions. They really came up with an answer, that sound needs something to bounce off from, in order for you to hear it. I found that quite nice, that they came up with that. I did not expect it actually.

In the post-interview, however, her focus was still mainly on skills:

> For me, especially in my class, it's really the skills that the children don't possess yet because they're not accustomed to doing inquiry-based learning. (...) Some of them have the knowledge, and some don't, but it's mainly the skills that they still lack in my class.

The enrichment that she offered in the third science lesson (the follow-up) was also focused on skills, where high-achieving students were instructed to use multiple sources of information instead of one, and to compare sources to answer their research question.

Regarding her own self-efficacy in asking enrichment questions, she stated:

> At first I thought, how on earth can you do that if you yourself don't even know how the concept works exactly... Now we just did some things together [in the PDP], and I can ask questions more easily because of that.
>
> (Interview; post-test).

During the observations of lesson 1 and 2, however, it seemed that the enrichment questions that she asked were still somewhat limited. She did ask the questions as discussed in the PDP, but did not respond much to the answers provided by the students. One student was on the verge of understanding the importance of distance in the experiment with the ruler.

> He did not reason about multiplying distance with weight (...) I thought: he is not going to get there anyway, and I might just confuse him (...). It might work, but I would need more time (...) Otherwise, he will just be frustrated
>
> (interview; lesson 2).

After both lessons, when asked what the added value was of asking enrichment questions, Anne answered that it will get students more enthusiastic about the lesson, but that it's not really necessary for their learning.

5.2.3.3 | Iris. Iris is a 22-year-old female, with somewhat less experience in science education than Lisa, but somewhat more than Anne. Her DSAQ-scores were between 2 and 3 for most scales on the pre-test, and between 3 and 4 on the post-test (as shown in Table 1). In the pre-interview, Iris found it hard to envision what DI in science education would look like because of lack of experience. She did think it is possible and necessary to implement DI in science education.

In the first lesson, Iris had trouble with implementing the DI practice of asking enrichment questions. She felt that the students were concentrated and working "nicely", and she didn't want to disturb that by asking them questions. In reflecting on the question if these students should have had enrichment, she said:

> I don't think they will go home with the feeling 'oh, this was so simple, I didn't get anything out of it'. (...) I think they all really enjoyed it... (...) They all learned something today. So, I don't think they missed out
>
> (interview; lesson 1).

In the second lesson, she did ask some enrichment questions, but still found it difficult:

> When they started trying it, I asked 'What is happening now? Articulate what is happening.' [Thinks] Yeah... Oh, I do find those enrichment questions a bit tricky, you know
>
> (interview; lesson 2).

It is possible that she didn't ask the enrichment questions in lesson 1 because she found that difficult to do, because of low self-efficacy or lack of content knowledge, rather than not wanting to disturb the students.

Similar to the practice of Anne, Iris seemed to lack a focus on students' learning of content or concepts. But while Anne seemed to be focused on inquiry skills, Iris was mainly concerned about whether students showed on-task behavior, rather than if they were learning new skills. This was also reflected in her answer to the question how she determined that students had a good understanding of the concept 'balance':

> That they're not looking at that worksheet thinking, 'What am I supposed to do now?' (…) So they got to work right away and didn't have all these question marks above their heads
>
> (interview, lesson 2).

Another incident that supported this notion was during the observation of lesson 2. Iris had provided a pair of students with an enrichment question, by placing two clothespins on one side of the fulcrum and asking them where a third clothespin should go to make the ruler balanced. The students made an incorrect guess before placing the clothespin (the clothespin had to be placed much further away from the fulcrum than the students had guessed). This made them exclaim: "Huh? But how can that be?!" Iris responded with something along the lines of: "Good job, so where does the clothespin go? Write it down", and walked away. She did not support these students in using their discovery and amazement to reason about the role of distance.

Different from Lisa and Anne, Iris' practice of identifying who needs support or enrichment was not only informed by mathematics and reading scores. In the interview after lesson 1, she mentioned that she feels there are also other indicators that determine whether a student struggles, such as concentration. She also felt, after lesson 1, that students working in homogenous pairs (i.e., ability grouping) were disadvantageous for students with reading difficulties; there is no one 'taking the lead'. In the post-interview and the interview after lesson 3, she mentioned using mathematics and reading scores as identifiers for needs, but also her own observations. Grouping decisions were made based on 'strengthening each other's character', for instance pairing high mathematical achievers with creative students.

Furthermore, Iris mentioned in the post-interview that she does not want to take children aside for additional instruction during a science lesson. Pre-teaching would be an option she would consider, but during the lesson itself she feels the additional instruction is perceived negatively by students.

> I don't know why. With math, for example, you do sometimes take students aside to provide additional instruction. That is perceived really differently from when all other students get to work on a worksheet or craft-assignment
>
> (interview; post-test).

The term 'worksheet' was used by Iris to refer to something other than the students' regular workbook, such as the worksheet that was used to guide students' experiments.

5.2.3.4 | Follow-Up Lessons. We will discuss the main differences or similarities between the three follow-up lessons. A first similarity of all three was that the lessons were not typical or high-quality science lessons, which evidently influences the DI practices enacted by the preservice teachers. Suitable for the Dutch context of science education, some of their lessons were more focused on design-based learning, yet still substandard. Lisa taught (part of) a design-based learning lesson about a communication tool for the future, Anne a history themed lesson with "inquiry" (consulting online sources) regarding a conflict between two countries, and Iris carried out a design challenge where students had to build a bridge that could hold a toy car, using three sheets of paper. Although all lessons had some connection to inquiry or design, they lacked some important elements of science education, such as active construction of conceptual knowledge or systematically testing and refining a solution for a well-defined problem. Iris purposely chose to simplify her lesson, a decision she first ascribed to the students' ability level. After further questioning, however, she said:

> I also find it challenging to explain. So, it's a bit for myself as well, like 'okay, I don't understand it. Then I can't convey it to the children either' (…) I think if I had confidently said 'Hey, guys, you can build a suspension bridge like this or that, if I can convey it well and answer questions about it, then I believe it would become clear for the students'
>
> (interview; lesson 3).

Notably, Iris' self-efficacy scores for DI in science education were almost maximum.

Both Anne and Iris chose not to provide support for students who had trouble attaining the lesson objective, while Lisa provided pre-teaching for all students. Anne and Lisa indicated they had trouble identifying students who might need support. In Anne's case, some general support related to her follow-up lesson was already provided for students with reading difficulties at a different time, by a different teacher. Iris indicated that there were no students who had trouble attaining the learning objective. She mentioned that she usually starts with instruction for everybody, and adjusts her teaching on the spot if necessary:

> I do make sure I know, 'Okay, how can I take a step back, in order for them to understand it? So I don't have to come up with it on the spot
>
> (Interview; lesson 3).

However, upon further questioning, Iris could only come up with examples for mathematics education. She realized that she finds it difficult to determine how to adjust her science instruction to attune to a student's current understanding, which is something Lisa mentioned as well.

For enrichment, all three preservice teachers selected students who are high-achievers in mathematics or other subject areas. Both Anne and Lisa specifically felt that mathematics scores are a good indication of who needs enrichment in science. Lisa also mentioned strong communication skills as indicative. Lisa and Iris grouped students heterogeneously according to their general academic ability, reasoning that students can be complementary to each other. Lisa said:

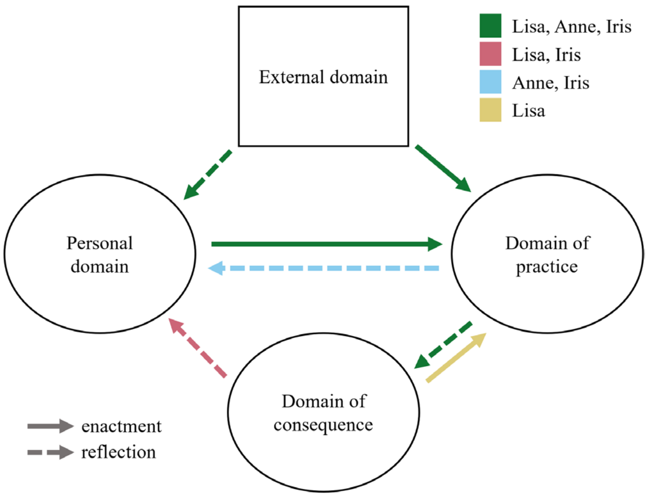

FIGURE 4 | Change sequences embedded units of analysis.

> Often time those kids who may not be very strong in math or grammar are very creative and can inspire other children (…) to come up with more creative ideas
>
> (interview; lesson 3).

In Lisa's lesson, the group of students that turned out to be 'strongest' in their work was not what she had expected beforehand, which struck her as very positive. The type of enrichment provided was different for all three teachers. Lisa asked questions, Anne had an additional learning objective (comparing multiple sources of information) and Iris provided a restriction in learning material (building the bridge with one instead of three sheets of paper).

5.2.3.5 | Cross-Case Analysis. The change sequences for all three embedded cases are depicted in Figure 4. The green arrows represent change sequences that were evident for all three preservice teachers, although they might represent different practices. Notably, for all three, there were no arrows going *to* the external domain, meaning that apart from the input provided in the PDP, there were no changes regarding external sources of information. For Lisa, changes in the domain of practice as a result of the PDP (e.g., using mathematics and reading scores as indicators, asking enrichment questions) led to changes in the domain of consequence (e.g., noticing how students are ready for enrichment), which led to a (strengthened) belief about who is ready for enrichment and who is not. There was also a continuance of a practice, as she asked enrichment questions in the follow-up lesson as well. She still felt insecure about supporting students who have trouble with the learning material in science education, which she also didn't specifically implement in the follow-up lesson (she provided pre-teaching for everybody). Additionally, her expectations for the follow-up lesson about who would be a high-achieving student did not match with the outcome, which might lead to new beliefs and expectations for the following lesson.

For Iris, a change sequence also starts with asking enrichment questions and providing support to ability groups in practice, as a result of the PDP. Regarding support, the most salient outcome for her was that it feels uncomfortable to work with additional instruction in science education for ability groups. It seems that she also does not feel the students with reading difficulties are in need of extra support, as she does not provide this in the follow-up lesson. She did continue with enrichment and also uses her own observations in addition to students' mathematics ability to determine who might need enrichment. She challenged students by limiting their use of material rather than asking them questions. Asking enrichment questions is something she might still feel insecure about, even though her post-test self-report self-efficacy score was almost maximum. In terms of a change sequence, it is possible that the PDP did increase her self-efficacy (arrow from external to personal), but that trying it out in practice, after the PDP, decreased it (arrow from practice to personal).

For Anne, important change sequences were also related to the practice of asking enrichment questions. By doing so, she discovered that students were able to answer these questions (domain of consequence), which she had not expected. This did not seem to change the belief she had about knowledge or conceptualization being less important than skill acquisition in science education (personal domain). The learning goals that she formulated within the PDP were focused on conceptualization, but in the follow-up lesson, they were mostly focused on inquiry skills. Her continued practice thus seems to be mainly influenced by existing beliefs, rather than the outcomes she experienced throughout the PDP. Her self-efficacy beliefs about being able to ask enrichment questions did change over the course of the PDP, which may be a result of her practice with it or as a result of training within the PDP. Her practice of asking enrichment questions in lesson 2 also seemed to be influenced by beliefs about certain students and whether or not her practice would lead them to a better understanding of the concept, given the available time.

6 | Discussion

Employing an embedded single case study, we aimed to explore how a professional development program (PDP) for differentiated instruction (DI) in science education supports change in preservice teachers. Specifically, we looked at preservice teachers' change in the personal domain, domain of consequence, domain of practice, and the external domain, following the Interconnect Model of Professional Growth (IMPG) by Clarke and Hollingsworth (2002). By including preservice teachers as embedded units of analysis, for which we collected more in-depth sources of evidence, we could examine in closer detail how changes in the different domains interact with each other (i.e., change sequences) to further illustrate our case of professional development.

Our PDP was designed to support change in the four domains of the IMPG, by, for example, including theory on DI and DI in science education (external domain), deliberate practice in both group meetings and in the preservice teachers' internship schools (domain of practice), support for pedagogical content knowledge prior to the lessons the preservice teachers taught (personal domain), and reflecting on student outcomes (domain of consequence). In our results section, we provided an extended summary of preservice teachers' professional development based on questionnaire data and reflective journals, and for three embedded units of analysis also observations and interviews. Overall, preservice teachers reported increased practice of DI in science education, higher self-efficacy in implementing DI in science education, and positive experiences with the PDP. With the use of the IMPG, we illustrated aspects of the PDP that may have contributed to change, while also nuancing the self-reported progress. To answer our research question, we will discuss the most important changes in three domains of the IMPG: the personal domain, domain of consequence, and the domain of practice. For the external domain, which consists of external sources of information or stimuli, there were no changes other than the input provided in the PDP. We therefore focus our discussion on changes in the other three domains.

6.1 | Personal Domain

The personal domain consists of changes in preservice teachers' knowledge, beliefs, and attitude (Clarke and Hollingsworth 2002). In general, the preservice teachers lacked knowledge about science education, DI in science education, and specific content knowledge. The findings show that the PDP supported the preservice teachers in, for example, knowing how to set differentiated goals or how to formulate enrichment questions. While the PDP also supported the preservice teachers in content knowledge for the two lessons that were part of the PDP, the follow-up lessons for the embedded cases were free of choice and were not supported or prepared during the PDP. Notably, all three follow-up lessons did not include teaching science content knowledge. A likely reason why these preservice teachers chose to abstain from teaching science content knowledge is a lack of self-efficacy in this area, which became apparent in the embedded analysis of Iris. She initially claimed she decided to simplify her science lesson about constructing a bridge because of the students' ability but later acknowledged that she felt it was too difficult for her in terms of content knowledge. She did think that if she had invested more time in understanding the material herself, she could convey it to the students.

The interrelatedness of lower self-efficacy beliefs, lack of content knowledge, and teaching practice has been previously shown as well (Bleicher and Lindgren 2005; Knaggs and Sondergeld 2015; Menon and Sadler 2016). In the current study, however, quantitative self-efficacy scores were quite high after the PDP, including for Iris (4.75 out of 5), which does not seem to correspond with our findings in the follow-up lesson. This discrepancy was also found in the study by Eysink et al. (2017), regarding DI in science education among in-service teachers, and by Herrington et al. (2016), regarding science teaching among preservice teachers. These authors hypothesize that a limited view of what DI or science teaching entails may help explain the inflated self-efficacy perceptions. Kazempour and Sadler (2015) also speculate that preservice teachers who are interested in science may overestimate their ability to teach it. Alternatively, preservice teachers might feel discomfort in admitting they lack knowledge (Wilkie and Clarke 2015).

Changing beliefs about students was important within this domain as well. Part of the PDP was to write down expectations about students and reflect on them, which led to surprising outcomes for several preservice teachers; students were unexpectedly competent or engaged. In a study about teaching algebra, Wilkie and Clarke (2015) found that teachers were similarly surprised that their "usual suspects", that is, students with high mathematical achievement, were not the most competent students when it came to functional thinking. This further points to the importance of identifying educational needs, which is discussed under the domain of practice.

6.2 | The Domain of Consequence

The domain of consequence constitutes the outcomes of a classroom practice that a teacher considers to be salient (Clarke and Hollingsworth 2002). Related to this domain, we found that preservice teachers seemed to lack a focus on student learning yet were able to develop it more through the PDP. Focusing on student learning has been linked to the ability of noticing, which can be conceptualized as teachers' ability to identify and attend to student ideas, to make sense of what they observe or hear students say, to connect this to their pedagogical content knowledge, and to shape their following interactions with a student (Kang and Anderson 2015; Preminger et al. 2024; van Es and Sherin 2021). In science education, noticing is relevant in order to respond to students' (initial) conceptions (Chan et al. 2021).

A first indication of a change in noticing was that one preservice teacher revised her initial reflection that students had not learned anything, which was induced by examining the worksheets together in a group session. Hence, letting the preservice teachers reflect on *actual* outcomes rather than what they *thought* the outcome was seemed to lead to a changed belief. This finding is similar to the study by Kang and Anderson (2015), where the use of high-quality assessments and supporting subsequent interpretation of these assessments was related to preservice teachers' noticing and responsiveness to student thinking.

Second, we described how Lisa asked enrichment questions to help students reframe their reasoning when she noticed that students did not yet integrate distance in their reasoning about balance. She also noticed the consequence of her practice, which seemed to lead to a (strengthened) belief about the value of it, as she continued with asking questions in her follow-up lesson. Both Anne and Iris, on the other hand, did not seem to be convinced that asking questions could improve learning outcomes and did not continue with it in their follow-up lessons. Their ability to notice also seemed less developed; Iris was more focused on maintaining an "orderly" classroom and did not respond to students' amazement during their experimentation, while Anne mentioned that she previously did not review the learning objectives of a lesson. Lisa was also most experienced in science teaching, which might help explain the differences between these teachers. In previous studies, novice teachers are found to be less focused on student learning than more experienced teachers (Preminger et al. 2024; Richards et al. 2020; Widjaja et al. 2017; Wolff et al. 2015).

Another salient outcome preservice teachers reflected on was students' perception of certain DI practices. Within the PDP, we chose to let the preservice teachers practice with additional instruction for ability groups; a form of DI that is very common in many (Dutch) schools for language and mathematics education (Denessen 2017; Prast et al. 2023). Most preservice teachers in the current study felt uncomfortable implementing this practice in their science lessons. In the first science lesson, additional instruction was provided to a small group of students with reading difficulties, while other students were already working on their experiments. Some students had expressed their dislike of being taken aside after whole-class instruction ("Why always me?"), leading preservice teachers to conclude that it was demotivating for students. Most preservice teachers did feel comfortable with providing the additional instruction through pre-teaching or to pairs of students already experimenting. Iris reflected on her different perceptions of additional instruction for science versus mathematics education and attributed it to the difference in type of learning activities (e.g., experimenting versus working in a textbook). It may also have to do with the preservice teachers' objective for the science lesson; Anne, for example, indicated that she did not think it was necessary that students learn conceptual knowledge and was more focused on teaching skills. The disinclination of preservice teachers to use this DI practice points to the need to research DI in the specific context of science education.

Furthermore, for the follow-up lesson, two of the three selected preservice teachers used heterogeneous ability grouping. These preservice teachers reasoned that, by grouping generally high-achieving students with academically low-achieving but creative students, everyone could learn from each other. The research done so far regarding the effects of ability grouping has been ambiguous (e.g., Denessen 2017; Prast et al. 2023), which is probably due to the complexity of grouping. Important distinctions exist between, for instance, fixed or flexible grouping, grouping for specific subjects or not, and whether or not the teacher actually modifies instruction to accommodate the needs (Deunk et al. 2018). Further research is needed for both the effects of ability grouping in the context of science education, and the beliefs teachers hold about grouping and how they affect their instructional decisions.

6.3 | The Domain of Practice

Practicing with the implementation of DI resulted in changes in the domain of practice, as the preservice teachers in this study had little experience with DI in science education. Considering the complexity of DI, their lack of experience is not surprising. The in-service teachers in the study by Eysink et al. (2017) were found to only implement DI in their science lessons during an intervention, but not before or after. Frerejean et al. (2021) have also pointed out that mastering basic classroom management skills might be a prerequisite for learning to implement DI. Our results about preservice teachers' inability to notice meaningful interactions endorse this claim, yet our findings also show that only minor prompts provided in our PDP made preservice teachers more aware of student learning. It is important to further understand the relationship between classroom management skills and noticing in the context of implementing DI in science education. For example, are preservice teachers unable to identify meaningful learning opportunities or is it that they are uncomfortable acting upon those opportunities because of having to manage the rest of the classroom?

The DI practice of identifying educational needs posed difficulties for preservice teachers in the context of science teaching. We chose to include mathematics and reading achievement as indicators for support or enrichment, as these individual differences have been linked to achievement in science education. However, we also discussed with the preservice teachers how using these scores can be problematic, as low prior achievement has been shown to lower teachers' expectations of students and negatively affect students' subsequent achievement (Rubie-Davies 2015). Furthermore, students who usually struggle academically can be high achievers in science education (Slim et al. 2022), which the preservice teachers in the current study mentioned as well. It thus seems vital to determine how these students can best be supported or challenged while cherishing their positive learning experience with science, and how (preservice) teachers can further be supported in identifying needs.

6.4 | Limitations

A first limitation of the current study is that we could only provide limited insight into the professional growth of the preservice teachers. By including follow-up observations for the embedded cases, which is an advantage of the current study, we were able to determine how preservice teachers implemented DI in their science lessons when given more freedom of choice. Lasting professional growth, however, should be established through a longitudinal study. Second, we have accounted for the validity of the DSAQ in the context of science education by using Cronbach's alpha and triangulation of data. Since self-report measures are prone to deviate from actual DI practice (Sougari and Mavroudi 2019), which our results underline, a validated measurement instrument regarding DI in science education is an important gap that must be addressed in future research. Preferably an instrument that includes lesson observations, or one that accounts for teachers' actual knowledge about DI, rather than mere self-report. Third, our PDP offered limited support for students who have trouble attaining the lesson goals. Our focus on providing reading support was based on previous research, but further research into the various difficulties students face in science and how these can be supported by teachers is much needed (Slim et al. 2022).

7 | Conclusion and Implications

The current study has presented a detailed look into how a PDP for DI in science education supported change in preservice teachers, thereby providing important implications for future research and science teacher education.

First, our results provide more insight into how DI can be implemented in science education and endorse earlier calls for researching DI in specific contexts (e.g., Pozas et al. 2020). Research on the implementation of DI in science education is limited (Sun and Xiao 2021; Tobin and Tippett 2014), yet professional development for subject-specific DI is most effective (Kahmann et al. 2022). Our study is therefore an important contribution to the field, as it provides process descriptions of what DI in science education looks like and what possible barriers or opportunities there are for starting teachers. Preservice teachers' inclination to provide additional instruction to students with reading difficulties, for example, illustrates the necessity to advance our understanding of DI in science education; relying on knowledge about DI in other educational contexts is not enough to fully understand preservice teachers' professional development.

Second, by building on the theory of Clarke and Hollingsworth (2002), our findings contribute to the understanding of teachers' professional growth in terms of changes in multiple domains working together in a non-linear process. To our knowledge, our case study is the first to specifically examine preservice teacher professional development regarding implementation of DI in elementary science education. Earlier studies that have used the IMPG highlight the domain of consequence as central for teacher learning (e.g., Justi and Van Driel 2006; Preminger et al. 2024; Widjaja et al. 2017; Witterholt et al. 2012), which is also in accordance with Guskey (2002), who proposes that evidence of student learning contributes to changes in teachers' attitudes and beliefs. Particularly for teacher development regarding DI, Schipper et al. (2017) have shown as well that increasing the focus on student learning can help preservice teachers' development of DI-related skills. Our results about the important role of noticing fit in with this line of reasoning. They show how it can be worthwhile for teacher education programs to specifically support preservice teachers in the skills associated with noticing, by analyzing students' work together, using high-quality assessment tasks, or helping them to set goals and review them (Kang and Anderson 2015; Preminger et al. 2024).

A third conclusion is the possible fundamental role of the personal domain of the IMPG, likely due to this study's distinctive context of science education. Within the structured, pre-planned science lessons that were part of the PDP, most preservice teachers were able to implement DI. However, for the science lesson that was free of choice, we saw that the three preservice teachers abstained from teaching content knowledge, which affected their use of DI as well. Preservice teachers' lack of content knowledge seemed to play a role, either in relation to beliefs about the objectives of science education or with low self-efficacy. Given that vicarious experiences contribute to the development of self-efficacy, it is important that teacher education provides preservice teachers with effective role models (Kazempour and Sadler 2015; Scarparolo and Subban 2021). Role models who are effective in both science teaching as well as implementing DI in science education.

Future research using the IMPG should further disentangle the interplay of the different domains in the context of science education and implementing DI. For teacher education programs, supporting content knowledge development and self-efficacy seems a promising catalyst for improving science teaching overall and preparing teachers to meet the needs of all students.

Funding

There are no sources of funding for this study.

Ethics Statement

This research was approved by the ethics committee of the Vrije Universiteit Amsterdam (VCWE-2022-121).

Consent

All participants have provided informed consent.

Data Availability Statement

The datasets presented in this article are not readily available because of privacy restrictions. Requests to access the datasets should be directed to t.slim@vu.nl.

Endnotes

[1] An additional context sample was included to compare participating and non-participating preservice teachers. This sample differed significantly from the included sample, which we elaborate on in the Supporting Information (methods_supplement1).

References

Aditomo, A., and E. Klieme. 2020. "Forms of Inquiry-Based Science Instruction and Their Relations With Learning Outcomes: Evidence From High and Low-Performing Education Systems." *International Journal of Science Education* 42, no. 4: 504–525. https://doi.org/10.1080/09500693.2020.1716093.

Anyfandi, G., A. Filippoupoliti, and I. Alexopoulous. 2016. "ENGINEER Project: Breaking New Ground in the Science Education Realm." In *Electronic Proceedings of the ESERA 2015 Conference. Science Education Research: Engaging Learners for a Sustainable Future, Part 9*, edited by J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto, and K. Hahl, 145–157. University of Helsinki.

Bächtold, M. 2013. "What Do Students "Construct" According to Constructivism in Science Education?" *Research in Science Education* 43, no. 6: 2477–2496. https://doi.org/10.1007/s11165-013-9369-7.

Bandura, A. 1977. "Self-Efficacy: Toward a Unifying Theory of Behavioral Change." *Psychological Review* 84, no. 2: 191–215.

Bleicher, R. E., and J. Lindgren. 2005. "Success in Science Learning and Preservice Science Teaching Self-Efficacy." *Journal of Science Teacher Education* 16, no. 3: 205–225. https://doi.org/10.1007/s10972-005-4861-1.

Bosker, R. J. 2005. "De Grenzen Van Gedifferentieerd Onderwijs." The Limits of Differentiated Education Rijksuniversiteit Groningen. https://pure.rug.nl/ws/portalfiles/portal/14812458/bosker.pdf.

Bowers, N., E. Merritt, and S. Rimm-Kaufman. 2020. "Exploring Teacher Adaptive Expertise in the Context of Elementary School Science Reforms." *Journal of Science Teacher Education* 31, no. 1: 34–55. https://doi.org/10.1080/1046560X.2019.1651613.

Boylan, M., M. Coldwell, B. Maxwell, and J. Jordan. 2018. "Rethinking Models of Professional Learning as Tools: A Conceptual Analysis to Inform Research and Practice." *Professional Development in Education* 44, no. 1: 120–139. https://doi.org/10.1080/19415257.2017.1306789.

Braun, V., and V. Clarke. 2019. "Reflecting on Reflexive Thematic Analysis." *Qualitative Research in Sport, Exercise and Health* 11, no. 4: 589–597. https://doi.org/10.1080/2159676X.2019.1628806.

Chan, K. K. H., L. Xu, R. Cooper, A. Berry, and J. H. Van Driel. 2021. "Teacher Noticing in Science Education: Do You See What I See?"

Studies in Science Education 57, no. 1: 1–44. https://doi.org/10.1080/03057267.2020.1755803.

Clarke, D., and H. Hollingsworth. 2002. “Elaborating a Model of Teacher Professional Growth.” *Teaching and Teacher Education* 18, no. 8: 947–967. https://doi.org/10.1016/S0742-051X(02)00053-7.

Cleland, J., A. MacLeod, and R. H. Ellaway. 2021. “The Curious Case of Case Study Research.” *Medical Education* 55, no. 10: 1131–1141. https://doi.org/10.1111/medu.14544.

De Jesus, O. N. 2012. “Differentiated Instruction: Can Differentiated Instruction Provide Success for All Learners?” *National Teacher Education Journal* 5, no. 3: 5–11.

De Neve, D., and G. Devos. 2016. “How Do Professional Learning Communities Aid and Hamper Professional Learning of Beginning Teachers Related to Differentiated Instruction?” *Teachers and Teaching* 23: 262–283. https://doi.org/10.1080/13540602.2016.1206524.

De Neve, D., G. Devos, and M. Tuytens. 2015. “The Importance of Job Resources and Self-Efficacy for Beginning Teachers' Professional Learning in Differentiated Instruction.” *Teaching and Teacher Education* 47: 30–41. https://doi.org/10.1016/j.tate.2014.12.003.

De Vries, S., E. P. W. A. Jansen, and W. J. C. M. Van De Grift. 2013. “Profiling Teachers' Continuing Professional Development and the Relation With Their Beliefs About Learning and Teaching.” *Teaching and Teacher Education* 33: 78–89. https://doi.org/10.1016/j.tate.2013.02.006.

Denessen, E. 2017. “Dealing Responsibly With Differences.” Socio-Cultural Backgrounds and Differentiation in Education (Inaugural Lecture) Universiteit Leiden. https://openaccess.leidenuniv.nl/handle/1887/51574.

Deunk, M. I., A. E. Smale-Jacobse, H. De Boer, S. Doolaard, and R. J. Bosker. 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.

Dobber, M., R. Zwart, M. Tanis, and B. Van Oers. 2017. “Literature Review: The Role of the Teacher in Inquiry-Based Education.” *Educational Research Review* 22: 194–214. https://doi.org/10.1016/j.edurev.2017.09.002.

Eysink, T. H. S., M. Hulsbeek, and H. Gijlers. 2017. “Supporting Primary School Teachers in Differentiating in the Regular Classroom.” *Teaching and Teacher Education* 66: 107–116. https://doi.org/10.1016/j.tate.2017.04.002.

Frerejean, J., M. Van Geel, T. Keuning, D. Dolmans, J. J. G. Van Merriënboer, and A. J. Visscher. 2021. “Ten Steps to 4C/ID: Training Differentiation Skills in a Professional Development Program for Teachers.” *Instructional Science* 49, no. 3: 395–418. https://doi.org/10.1007/s11251-021-09540-x.

Furtak, E. M., T. Seidel, H. Iverson, and D. C. Briggs. 2012. “Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis.” *Review of Educational Research* 82, no. 3: 300–329. https://doi.org/10.3102/0034654312457206.

Gheyssens, E., C. Coubergs, J. Griful-Freixenet, N. Engels, and K. Struyven. 2020. “Differentiated Instruction: The Diversity of Teachers' Philosophy and Praxis to Adapt Teaching to Students' Interests, Readiness and Learning Profiles.” *International Journal of Inclusive Education* 26, no. 14: 1383–1400. https://doi.org/10.1080/13603116.2020.1812739.

Griful-Freixenet, J., K. Struyven, and W. Vantieghem. 2021. “Exploring Pre-Service Teachers' Beliefs and Practices About Two Inclusive Frameworks: Universal Design for Learning and Differentiated Instruction.” *Teaching and Teacher Education* 107: 103503. https://doi.org/10.1016/j.tate.2021.103503.

Gunning, A. M., and F. M. Mensah. 2011. “Preservice Elementary Teachers' Development of Self-Efficacy and Confidence to Teach Science: A Case Study.” *Journal of Science Teacher Education* 22, no. 2: 171–185. https://doi.org/10.1007/s10972-010-9198-8.

Guskey, T. R. 2002. “Professional Development and Teacher Change.” *Teachers and Teaching* 8, no. 3: 381–391. https://doi.org/10.1080/135406002100000512.

Herrington, D. G., E. J. Yezierski, and S. F. Bancroft. 2016. “Tool Trouble: Challenges With Using Self-Report Data to Evaluate Long-Term Chemistry Teacher Professional Development.” *Journal of Research in Science Teaching* 53, no. 7: 1055–1081. https://doi.org/10.1002/tea.21323.

Hornstra, L., K. Stroet, C. Rubie-Davies, and A. Flint. 2023. “Teacher Expectations and Self-Determination Theory: Considering Convergence and Divergence of Theories.” *Educational Psychology Review* 35, no. 3: 76. https://doi.org/10.1007/s10648-023-09788-4.

Justi, R., and J. Van Driel. 2006. “The Use of the Interconnected Model of Teacher Professional Growth for Understanding the Development of Science Teachers' Knowledge on Models and Modelling.” *Teaching and Teacher Education* 22, no. 4: 437–450. https://doi.org/10.1016/j.tate.2005.11.011.

Kahmann, R., M. Droop, and A. W. Lazonder. 2022. “Meta-Analysis of Professional Development Programs in Differentiated Instruction.” *International Journal of Educational Research* 116: 102072. https://doi.org/10.1016/j.ijer.2022.102072.

Kang, H., and C. W. Anderson. 2015. “Supporting Preservice Science Teachers' Ability to Attend and Respond to Student Thinking by Design: RESPONDING TO STUDENT THINKING.” *Science Education* 99, no. 5: 863–895. https://doi.org/10.1002/sce.21182.

Kazempour, M., and T. D. Sadler. 2015. “Pre-Service Teachers' Science Beliefs, Attitudes, and Self-Efficacy: A Multi-Case Study.” *Teaching Education* 26, no. 3: 247–271. https://doi.org/10.1080/10476210.2014.996743.

Ketelhut, D. J., K. Mills, E. Hestness, L. Cabrera, J. Plane, and J. R. McGinnis. 2020. “Teacher Change Following a Professional Development Experience in Integrating Computational Thinking Into Elementary Science.” *Journal of Science Education and Technology* 29, no. 1: 174–188. https://doi.org/10.1007/s10956-019-09798-4.

Keuning, T., M. van Geel, and M. Dobbelaer. 2022. “ADAPT Krijg Zicht op Differentiatievaardigheden.” ADAPT Gives Insight Into Differentiation Skills. Uitgeverij Pica.

Knaggs, C. M., and T. A. Sondergeld. 2015. “Science as a Learner and as a Teacher: Measuring Science Self-Efficacy of Elementary Preservice Teachers.” *School Science and Mathematics* 115, no. 3: 117–128. https://doi.org/10.1111/ssm.12110.

Lazonder, A. W. 2018. “De Jonge Onderzoeker.” Radboud Universiteit.

Lazonder, A. W., and R. Harmsen. 2016. “Meta-Analysis of Inquiry-Based Learning: Effects of Guidance.” *Review of Educational Research* 86, no. 3: 681–718. https://doi.org/10.3102/0034654315627366.

Lazonder, A. W., N. Janssen, H. Gijlers, and A. Walraven. 2021. “Patterns of Development in Children's Scientific Reasoning: Results From a Three-Year Longitudinal Study.” *Journal of Cognition and Development* 22, no. 1: 108–124. https://doi.org/10.1080/15248372.2020.1814293.

Menon, D., and T. D. Sadler. 2016. “Preservice Elementary Teachers' Science Self-Efficacy Beliefs and Science Content Knowledge.” *Journal of Science Teacher Education* 27, no. 6: 649–673. https://doi.org/10.1007/s10972-016-9479-y.

Minner, D. D., A. J. Levy, and J. Century. 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, no. 4: 474–496. https://doi.org/10.1002/tea.20347.

National Research Council. 2012. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, 13165. National Academies Press. https://doi.org/10.17226/13165.

Nemo Science Museum. n.d. "Maakkunde." https://maakkunde.nl/lesmateriaal/.

Nikkanen, H. 2019. "Double Agent: Ethical Considerations in Conducting Ethnography as a Teacher." In *Implementing Ethics in Educational Ethnography: Regulation and Practice*, 1st ed., 379–394. Routledge.

Olmos-Vega, F. M., R. E. Stalmeijer, L. Varpio, and R. Kahlke. 2023. "A Practical Guide to Reflexivity in Qualitative Research: AMEE Guide No. 149." *Medical Teacher* 45, no. 3: 241–251. https://doi.org/10.1080/0142159X.2022.2057287.

Parsons, S. A., M. Vaughn, R. Q. Scales, et al. 2018. "Teachers' Instructional Adaptations: A Research Synthesis." *Review of Educational Research* 88, no. 2: 205–242. https://doi.org/10.3102/0034654317743198.

Pekrun, R. 2020. "Self-Report is Indispensable to Assess Students' Learning." *Frontline Learning Research* 8, no. 3: 185–193. https://doi.org/10.14786/flr.v8i3.637.

Pozas, M., V. Letzel, and C. Schneider. 2020. "Teachers and Differentiated Instruction: Exploring Differentiation Practices to Address Student Diversity." *Journal of Research in Special Educational Needs* 20, no. 3: 217–230. https://doi.org/10.1111/1471-3802.12481.

Prast, E. J., K. Stroet, A. Koornneef, and T. Wilderjans. 2023. "What Do Students Think About Differentiation and Within-Class Achievement Grouping?" *Frontline Learning Research* 11, no. 1: 57–93. https://doi.org/10.14786/flr.v11i1.1079.

Prast, E. J., E. Van De Weijer-Bergsma, E. H. Kroesbergen, and J. E. H. Van Luit. 2015. "Readiness-Based Differentiation in Primary School Mathematics: Expert Recommendations and Teacher Self-Assessment." *Frontline Learning Research* 3, no. 2: 90–116. https://doi.org/10.14786/flr.v3i2.163.

Prast, E. J., E. Van De Weijer-Bergsma, E. H. Kroesbergen, and J. E. H. Van Luit. 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.

Preminger, L., K. N. Hayes, C. L. Bae, and D. O'Connor. 2024. "Why Do Teachers Vary in Their Instructional Change During Science PD? The Role of Noticing Students in an Iterative Change Process." *Science Education* 108, no. 3: 701–733. https://doi.org/10.1002/sce.21853.

Richards, J., A. Elby, M. J. Luna, A. D. Robertson, D. M. Levin, and C. G. Nyeggen. 2020. "Reframing the Responsiveness Challenge: A Framing-Anchored Explanatory Framework to Account for Irregularity in Novice Teachers' Attention and Responsiveness to Student Thinking." *Cognition and Instruction* 38, no. 2: 116–152. https://doi.org/10.1080/07370008.2020.1729156.

Ritzema, E. S., M. I. Deunk, and R. J. Bosker. 2016. "Differentiation Practices in Grade 2 and 3: Variations in Teacher Behavior in Mathematics and Reading Comprehension Lessons." *Journal of Classroom Interaction* 51, no. 2: 50–72.

Roose, I., W. Vantieghem, R. Vanderlinde, and P. Van Avermaet. 2019. "Beliefs as Filters for Comparing Inclusive Classroom Situations. Connecting Teachers' Beliefs About Teaching Diverse Learners to Their Noticing of Inclusive Classroom Characteristics in Videoclips." *Contemporary Educational Psychology* 56: 140–151. https://doi.org/10.1016/j.cedpsych.2019.01.002.

Roy, A., F. Guay, and P. Valois. 2013. "Teaching to Address Diverse Learning Needs: Development and Validation of a Differentiated Instruction Scale." *International Journal of Inclusive Education* 17, no. 11: 1186–1204. https://doi.org/10.1080/13603116.2012.743604.

Rubie-Davies, C. M. 2015. *Becoming a High Expectation Teacher: Raising the Bar*. Routledge.

Saldaña, J. 2009. *The Coding Manual for Qualitative Researchers*. Sage.

Scarparolo, G., and P. Subban. 2021. "A Systematic Review of Pre-Service Teachers' Self-Efficacy Beliefs for Differentiated Instruction." *Teachers and Teaching* 27, no. 8: 753–766. https://doi.org/10.1080/13540602.2021.2007371.

Schipper, T., S. L. Goei, S. De Vries, and K. Van Veen. 2017. "Professional Growth in Adaptive Teaching Competence as a Result of Lesson Study." *Teaching and Teacher Education* 68: 289–303. https://doi.org/10.1016/j.tate.2017.09.015.

Schlatter, E., I. Molenaar, and A. W. Lazonder. 2020. "Individual Differences in Children's Development of Scientific Reasoning Through Inquiry-Based Instruction: Who Needs Additional Guidance?" *Frontiers in Psychology* 11: 904. https://doi.org/10.3389/fpsyg.2020.00904.

Schlatter, E., I. Molenaar, and A. W. Lazonder. 2021. "Learning Scientific Reasoning: A Latent Transition Analysis." *Learning and Individual Differences* 92: 102043. https://doi.org/10.1016/j.lindif.2021.102043.

Schlatter, E., I. Molenaar, and A. W. Lazonder. 2022. "Adapting Scientific Reasoning Instruction to Children's Needs: Effects on Learning Processes and Learning Outcomes." *International Journal of Science Education* 44, no. 17: 2589–2612. https://doi.org/10.1080/09500693.2022.2140019.

Seung, E., S. Park, and M.-A. Lee. 2019. "The Impact of a Summer Camp-Based Science Methods Course on Preservice Teachers' Self-Efficacy in Teaching Science as Inquiry." *Journal of Science Teacher Education* 30, no. 8: 872–889. https://doi.org/10.1080/1046560X.2019.1635848.

Simpkins, P. M., M. A. Mastropieri, and T. E. Scruggs. 2009. "Differentiated Curriculum Enhancements in Inclusive Fifth-Grade Science Classes." *Remedial and Special Education* 30, no. 5: 300–308. https://doi.org/10.1177/0741932508321011.

Slim, T., J. E. van Schaik, M. Dobber, A. C. Hotze, and M. E. Raijmakers. 2022. "Struggling or Succeeding in Science and Technology Education: Elementary School Students' Individual Differences During Inquiry- and Design-Based Learning." *Frontiers in Education* 7: 1–20. https://doi.org/10.3389/feduc.2022.842537.

Smit, R., and W. Humpert. 2012. "Differentiated Instruction in Small Schools." *Teaching and Teacher Education* 28, no. 8: 1152–1162. https://doi.org/10.1016/j.tate.2012.07.003.

Sougari, A.-M., and A. Mavroudi. 2019. "Differentiated Instruction in the EFL Classroom: Discrepancies Between Teachers' Self-Report Questionnaires and Actual Practices." *Selected Papers on Theoretical and Applied Linguistics* 23: 398–413. https://doi.org/10.26262/ISTAL.V23I0.7356.

Subban, P. 2006. "Differentiated Instruction: A Research Basis." *International Education Journal* 7, no. 7: 935–947.

Sun, Y., and L. Xiao. 2021. "Research Trends and Hotspots of Differentiated Instruction Over the Past Two Decades (2000-2020): A Bibliometric Analysis." *Educational Studies* 50: 1–17. https://doi.org/10.1080/03055698.2021.1937945.

Teig, N., R. Scherer, and R. V. Olsen. 2022. "A Systematic Review of Studies Investigating Science Teaching and Learning: Over Two Decades of TIMSS and PISA." *International Journal of Science Education* 44, no. 12: 2035–2058. https://doi.org/10.1080/09500693.2022.2109075.

Tobin, R., and C. D. Tippett. 2014. "Possibilities and Potential Barriers: Learning to Plan for Differentiated Instruction in Elementary Science." *International Journal of Science and Mathematics Education* 12, no. 2: 423–443. https://doi.org/10.1007/s10763-013-9414-z.

Tomlinson, C. A., C. Brighton, H. Hertberg, et al. 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, no. 2–3: 119–145. https://doi.org/10.1177/016235320302700203.

Tomlinson, C. A., and M. B. Imbeau. 2010. *Leading and Managing a Differentiated Classroom*. ASCD.

Tomlinson, C. A., E. M. Tomchin, and C. M. Callahan. 1994. "Preservice Teachers' Perceptions of and Responses to the Differential Needs." Paper presented at The annual meeting of the American Educational Research Association.

Van Aalderen-Smeets, S., and J. Walma Van Der Molen. 2013. "Measuring Primary Teachers' Attitudes Toward Teaching Science: Development of the Dimensions of Attitude Toward Science (DAS) Instrument." *International Journal of Science Education* 35, no. 4: 577–600. https://doi.org/10.1080/09500693.2012.755576.

Van Dijk, A. M., T. H. S. Eysink, and T. De Jong. 2016. "Ability-Related Differences in Performance of an Inquiry Task: The Added Value of Prompts." *Learning and Individual Differences* 47: 145–155. https://doi.org/10.1016/j.lindif.2016.01.008.

van Es, E. A., and M. G. Sherin. 2021. "Expanding on Prior Conceptualizations of Teacher Noticing." *ZDM Mathematics Education* 53, no. 1: 17–27. https://doi.org/10.1007/s11858-020-01211-4.

Van Geel, M., T. Keuning, J. Frèrejean, D. Dolmans, J. Van Merriënboer, and A. J. Visscher. 2019. "Capturing the Complexity of Differentiated Instruction." *School Effectiveness and School Improvement* 30, no. 1: 51–67. https://doi.org/10.1080/09243453.2018.1539013.

Van Graft, M., and M. Klein Tank. 2018. "Wetenschap & Technologie in Het Basis- en Speciaal Onderwijs: Richtinggevend Leerplankader Bij Het Leergebied Oriëntatie op Jezelf en de Wereld." Science and Technology in Primary Education: Guiding Curriculum Framework of the Subject Orientation to Yourself and the World. SLO.

Velthuis, C., P. Fisser, and J. Pieters. 2014. "Teacher Training and Pre-Service Primary Teachers' Self-Efficacy for Science Teaching." *Journal of Science Teacher Education* 25, no. 4: 445–464. https://doi.org/10.1007/s10972-013-9363-y.

Watts-Taffe, S., B. P. Barbara Laster, L. Broach, B. Marinak, C. McDonald Connor, and D. Walker-Dalhouse. 2012. "Differentiated Instruction: Making Informed Teacher Decisions." *Reading Teacher* 66, no. 4: 303–314. https://doi.org/10.1002/TRTR.01126.

Whitley, J., S. Gooderham, C. Duquette, S. Orders, and J. B. Cousins. 2019. "Implementing Differentiated Instruction: A Mixed-Methods Exploration of Teacher Beliefs and Practices." *Teachers and Teaching* 25, no. 8: 1043–1061. https://doi.org/10.1080/13540602.2019.1699782.

Widjaja, W., C. Vale, S. Groves, and B. Doig. 2017. "Teachers' Professional Growth Through Engagement With Lesson Study." *Journal of Mathematics Teacher Education* 20, no. 4: 357–383. https://doi.org/10.1007/s10857-015-9341-8.

Wilkie, K. J., and D. Clarke. 2015. "Pathways to Professional Growth: Investigating Upper Primary School Teachers' Perspectives on Learning to Teach Algebra." *Australian Journal of Teacher Education* 40, no. 2: 87–118. https://doi.org/10.14221/ajte.2015v40n4.6.

Witterholt, M., M. Goedhart, C. Suhre, and A. Van Streun. 2012. "The Interconnected Model of Professional Growth as a Means to Assess the Development of a Mathematics Teacher." *Teaching and Teacher Education* 28, no. 5: 661–674. https://doi.org/10.1016/j.tate.2012.01.003.

Wolff, C. E., N. Van Den Bogert, H. Jarodzka, and H. P. A. Boshuizen. 2015. "Keeping an Eye on Learning: Differences Between Expert and Novice Teachers' Representations of Classroom Management Events." *Journal of Teacher Education* 66, no. 1: 68–85. https://doi.org/10.1177/0022487114549810.

Yin, R. K. 2018. *Case Study Research and Applications: Design and Methods*. Sixth ed. SAGE.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** Supporting Information. **Data S2:** Supporting Information. **Data S3:** Supporting Information. **Table S1:** Mean scores and standard deviations for pre- and post-test questionnaire. $N = 8$. **Table S2:** ADAPT scores for the three embedded units of analysis.