Interpretation of Students' Errors as Part of the Diagnostic Competence of Pre-Service Primary School Teachers

Macarena Larrain · Gabriele Kaiser (ID)

Abstract Understanding students' thinking and learning processes is one of the greatest challenges teachers face in the classroom. Misconceptions and errors have the potential to be a rich source of information for identifying students' thinking and reasoning processes. However, empirical studies show that pre-service teachers (PSTs) and teachers find it challenging to focus their interpretations and pedagogical decisions on students' thinking processes when they identify students' mathematical errors.

Based on the theoretical approach of noticing, the study described in this paper examines primary PSTs' diagnostic competence in error situations before and after they participated in a seminar sequence implemented at several Chilean universities. Our analyses focus on PSTs' competence with regard to formulating hypotheses about the causes of students' errors. The proposed hypotheses were categorized into those that attributed errors to students' lack of conceptual understanding, those that explained errors in terms of lack of procedural understanding, and those that assumed a failure of instructional strategies. In addition, the relationships between PSTs' diagnostic competence, their beliefs and university learning opportunities were examined. The results indicate that PSTs' diagnostic competence in error situations and the changes of this competence were related to PSTs' beliefs, practical experiences, and learning opportunities. Overall, the findings suggest that it is possible to promote changes on PSTs' diagnostic competence during initial teacher

Macarena Larrain
Faculty of Education, Universidad de los Andes, Santiago, Chile

Macarena Larrain · Gabriele Kaiser (✉)
Faculty of Education, Universität Hamburg, Hamburg, Germany
E-Mail: gabriele.kaiser@uni-hamburg.de

Gabriele Kaiser
Faculty of Education and Arts, Nord University, Bodø, Norway

education. The paper concludes with implications for teacher education and future research.

Interpretation von Fehlern der Schülerinnen und Schüler als Teil der diagnostischen Kompetenz von Lehramtsstudierenden für die Grundschule

Zusammenfassung Das Verstehen der Denk- und Lernprozesse von Lernenden ist eine der größten Herausforderungen, denen sich Lehrerinnen und Lehrer im Unterricht stellen müssen. Fehlvorstellungen und Fehler stellen eine reichhaltige Informationsquelle für die Identifikation der Denk- und Argumentationsprozesse der Lernenden dar. Empirische Studien zeigen jedoch, dass es für (angehende) Lehrerinnen und Lehrer herausfordernd ist, ihre Interpretationen und pädagogischen Entscheidungen auf die Denkprozesse der Lernenden auszurichten, wenn sie die mathematischen Fehler der Lernenden erkennen.

Basierend auf dem Ansatz der professionellen Unterrichtswahrnehmung untersucht die hier beschriebene Studie die diagnostische Kompetenz von angehenden Grundschullehrkräften vor und nach der Teilnahme an einer an mehreren chilenischen Universitäten durchgeführten Seminarsequenz. Unsere Analysen konzentrieren sich auf die Kompetenz der angehenden Lehrkräfte, Hypothesen über die Ursachen der Fehler der Lernenden zu formulieren. Die vorgeschlagenen Hypothesen wurden kategorisiert in solche, die die Fehler auf mangelndes begriffliches Verständnis der Lernenden zurückführten, in solche, die die Fehler mit mangelndem prozeduralem Verständnis erklärten und solchen, die ein Versagen im unterrichtlichen Vorgehen annahmen. Darüber hinaus wurde die Zusammenhänge der diagnostischen Kompetenz mit den *beliefs* und den universitären Lernmöglichkeiten der Lehramtsstudierenden untersucht. Die Ergebnisse deuten darauf hin, dass die diagnostische Kompetenz von angehenden Grundschullehrkräften in Fehlersituationen und die Veränderungen dieser Kompetenz mit ihren *beliefs*, praktischen Erfahrungen und universitären Lernmöglichkeiten zusammenhängen.

Insgesamt weisen die Ergebnisse darauf hin, dass es möglich ist, die diagnostische Kompetenz von angehenden Grundschullehrkräften während der universitären Lehrerausbildung zu fördern. Der Beitrag schließt mit Schlussfolgerungen für die Lehrerausbildung und die Durchführung zukünftiger Forschung.

1 Introduction

Teachers' diagnostic competence is acknowledged to be essential for understanding and evaluating students' thinking. In addition, it is of great importance in the

teaching-learning process because it is key for adapting pedagogical responses to students' learning needs (Helmke 2017). In their daily practice, teachers can find many sources of information about their students' thinking. Learning situations in which errors arise can be very informative (Ashlock 2010; Prediger and Wittmann 2009; Rach et al. 2013; Radatz 1979). By looking closely at students' errors during their learning process, teachers can uncover incorrect conceptualizations and gain valuable insights into individual students' understanding of mathematical concepts and procedures. Based on this insight, teachers can target their instructional responses and adapt their teaching strategies to support students in overcoming misconceptions and building further mathematical knowledge.

Although classroom contexts can provide a wealth of information about students' mathematical thinking, this information can only be used in pedagogy if teachers are able to identify it. Which incidents teachers pay attention to and how they interpret it can impact their ability to provide adequate support to students and promote learning (Kaiser et al. 2015; Sherin et al. 2011). The competence needed to carry out the process of perceiving, interpreting and making pedagogical decisions can be developed with time and experience. However, several studies have highlighted the need to create opportunities for competence development early in initial teacher education programs (Artelt and Gräsel 2009; Bartel and Roth 2017; Brandt et al. 2017). Pre-service teachers (PSTs) need support to learn to look at relevant details, ask specific questions, and understand students' explanations (Götze et al. 2019).

Creating such complex opportunities is a challenge for teacher educators as it raises a number of questions related to the characteristics those opportunities should have and where they should be placed within teacher education programs. To answer those questions, it is necessary to extend our understanding of how diagnostic competence develops and the complex interactions between the dimensions of diagnostic competence and PSTs' knowledge, beliefs, and practical experience.

The present study aims to characterize PSTs' diagnostic competence and the development thereof. Due to its high relevance to the diagnosing process, we focus on one dimension of diagnostic competence, namely, the competence to formulate hypotheses about the causes of students' errors (Heinrichs and Kaiser 2018). In Sect. 2, we summarize the theoretical background on PSTs' diagnostic competence in error situations. Sects. 3 and 4 present the research questions and methods used, while Sect. 5 reports the results of the analyses. These are discussed in Sect. 6, providing a further outlook.

2 Theoretical Background

2.1 Teachers' Diagnostic Competence in Error Situations

Teachers' diagnostic competence is widely recognized as an important prerequisite for successful teaching, especially in heterogeneous classrooms (Helmke 2017). The individualization and differentiation of teaching strategies to meet students' needs in learning mathematics demands teachers who can effectively respond to a range of diagnostic challenges (Hoth et al. 2016). Several central tasks of teaching, such as

collecting information about individual students' abilities, recognizing and understanding various learning processes, and assessing and interpreting students' learning outcomes, involve diagnostic challenges for teachers (Bartel and Roth 2017). Despite challenging, these tasks are key for designing effective teaching strategies (Artelt and Gräsel 2009).

Heinrichs and Kaiser (2018) defined diagnostic competence in error situations as "the competence that is necessary to come to implicit judgments based on formative assessment in teaching situations by using informal or semiformal methods" (p. 81). They emphasized that the objective of diagnosis is to adapt teaching strategies and promote students' mathematical understanding. Identifying, interpreting and handling errors in complex classroom situations is a major challenge for teachers (Prediger and Wittmann 2009).

Empirical studies have taken different approaches to investigate the role of teachers' diagnostic competence in mathematics education. One approach is to conceptualize diagnostic competence as *accuracy of judgment*, which is determined by comparing teachers' predictions of student achievement with actual performance on some objective measure (Helmke and Schrader 1987). This approach has found large differences in teachers' ability to accurately judge student achievement (Südkamp et al. 2012) as well as various influences on these judgements (Kaiser et al. 2017). Other approaches focus on the diagnostic tasks that teachers perform in their daily activities to understand students' learning processes and make pedagogical decisions in the moment (Herppich et al. 2018; Praetorius et al. 2012). In the present study, we focus on PSTs' competence to gather information during instruction about students' mathematical understanding, learning difficulties, and possible misconceptions; to conduct ongoing analyses under complex circumstances; and to make in-the-moment pedagogical decisions that promote learning. This type of competence has been called *situation-based diagnostic competence* (Hoth et al. 2016). In particular, we focus on the diagnostic competence needed in learning situations in which students make errors.

In the context of mathematics, the errors students make when learning can be a valuable source for understanding their thinking, especially any flaws in students' mathematical reasoning and understanding of concepts (Radatz 1980). Analysis of errors can provide information about where students' knowledge and skills need further support and allow teachers to adapt their instructional strategies and arrange appropriate pedagogical resources (Brodie 2014; Brown and Burton 1978; McGuire 2013; Radatz 1979; Santagata 2005).

2.2 Hypothesizing About the Causes of Students' Errors

Several models have been developed to describe the diagnosis process in teaching situations. The present study adopts Heinrichs' (2015) model of the diagnostic process in error situations, which organizes the process into three phases: identification of the error, formulation of hypotheses about causes for the error, and decision-making aimed at supporting the student to overcome their error (Heinrichs and Kaiser 2018). Many approaches have emphasized the relevance of the interpretation phase of the diagnostic process because of the fundamental role of interpretation

of students' thinking in pedagogical decision-making. Moreover, most models of diagnostic processes in error situations include a phase in which hypotheses are made about students' behavior, sources of the errors they make, or their possible underlying difficulties (Klug et al. 2013; Reisman 1982; Wildgans-Lang et al. 2020). This requires from teachers what Heinrichs (2015) called *competence to hypothesize about causes of students' errors*, conceptualized as "the ability to find different hypotheses about causes for one specific error and especially being able to name causes for the specific error and not only the general reasons for an error to occur" (Heinrichs and Kaiser 2018, p. 85).

Teachers' hypotheses about students' mathematical procedures, their reasoning, and causes of their errors form the basis for diagnoses with a focus on students' learning processes (Scherer and Moser Opitz 2012). Ball et al. (2008) emphasized that only identifying wrong answers "does not equip a teacher with the detailed mathematical understanding required for a skillful treatment of the problems […] students face" (p. 397). Götze et al. (2019) argued that intervention strategies need to be based on a diagnosis to be targeted to the student's needs, and that diagnosis involves not only detecting an error or an error pattern but also identifying potential causes.

Results from empirical studies focusing on teachers' competence to identify, analyze, and respond to student errors confirm that this is a challenging process for teachers. In Cooper's (2009) study with PSTs, although all participants were able to identify computational error patterns in students' work, finding possible rationales for misconceptions was more difficult. Similarly, Seifried and Wuttke (2010) found that teachers faced difficulties in recognizing reasons for student errors during lessons. In a subsequent study, they found that PSTs (in contrast to in-service teachers) showed a low ability to identify and correct student errors (Türling et al. 2012; Wuttke and Seifried 2013).

Son (2013) analyzed the connection between PSTs' understanding of an error, interpretation of that error, and suggested pedagogical responses. She found that even for errors with a strong conceptual origin, most prospective teachers would explain it with procedural causes and suggest a remediation strategy based on telling the student the right procedure. Son ascribed this disconnection to the PSTs' insufficiently developed mathematical and professional knowledge. Overall, studies have indicated that teachers face challenges when interpreting students' errors, finding the underlying causes of these errors, and making coherent pedagogical decisions. Moreover, research suggests that teachers' and PSTs' competences in this area are related to their characteristics, such as their professional knowledge.

2.3 Types of Causes of Students' Errors

Teachers can assign an error to different categories and thus may make several hypotheses about its causes (Scherer and Moser Opitz 2012). Numerous categorizations have been developed to differentiate between types of errors. For example, a distinction can be made between random or careless errors and systematic errors resulting from an underlying incorrect idea, usually called a misconception (Cox 1975; Radatz 1980; Smith et al. 1993; Weimer 1925). Of particular interest to this

study are systematic errors, which exhibit a particular pattern and can be used as evidence of an underlying misconception of mathematical concepts or procedures (Ashlock 2010).

When formulating hypotheses about the causes of student errors, teachers can identify a wide variety of sources. On the one hand, it is possible to attribute the error to causes external to the student, such as deficiencies in teaching strategies, lack of clarity in task formulation, or contextual reasons. On the other hand, it is possible to find causes internal to the student. These, in turn, can be general or specific. General causes are those that refer to general learning or psychological difficulties faced by the student, lack of overall subject knowledge, or motivational aspects. Specific causes relate to the student's mathematical thinking regarding the error under analysis (Heinrichs and Kaiser 2018).

Previous studies distinguished between procedural and conceptual knowledge when discussing their role in the development of mathematical competence (Lenz et al. 2020; Rittle-Johnson and Alibali 1999; Son 2013). *Conceptual knowledge* is defined as the "explicit or implicit understanding of the principles that govern a domain and the interrelations between pieces of knowledge in a domain" (Rittle-Johnson and Alibali 1999, p. 175). It refers to abstract concepts, relations, and principles that are independent of specific problems. *Procedural knowledge* is defined as the "action sequences for solving problems" (Rittle-Johnson and Alibali 1999, p. 175). Although the relationship between the two types of knowledge is controversial, Rittle-Johnson and Koedinger (2009) point out that they do not develop independently or sequentially. Rather, they influence each other. The distinction between conceptual and procedural understanding helps to identify and understand student errors. In this way, it also helps to provide instructional responses that effectively address students' errors and promote understanding (Son 2013). Therefore, this distinction was used as the main framework in the present study to analyze PSTs' hypotheses about the causes of students' errors.

2.4 Development of Diagnostic Competence

Given the crucial role of diagnostic competence in effective teaching, the development of this competence needs to be fostered in initial teacher education (Artelt and Gräsel 2009; Brandt et al. 2017). Research on teacher expertise reveals differences in the knowledge structure of novices compared to experts (Berliner 2001). Experts are able to process more complex information, recognize relevant information faster, and represent this information more deeply and completely. They are also more flexible and able to adapt to different situations. However, to date, it has been unclear how teachers develop this expertise, and whether classroom experience is sufficient to promote development.

The structures of competence models provide certain insights into how teachers develop this competence. Blömeke et al.'s (2015) model of competence as a continuum considers cognitive and affective aspects to be dispositional elements, including professional knowledge, beliefs and motivation, among other features. Professional knowledge is an important part of most existing models of teachers' diagnostic competence (e.g., the COSIMA model: Heitzmann et al. 2019; NeDiKo model: Herppich

et al. 2018; DiaKom model: Loibl et al. 2020). Although the role and interaction of different components of professional knowledge in the diagnostic process remain unclear, various studies have revealed the influence of professional knowledge on diagnosis of students' understanding of mathematics (Bartel and Roth 2017; Ostermann et al. 2018; Kron et al. 2021). However, these studies have also indicated that knowledge alone is not sufficient. PSTs need opportunities to engage in evaluations of students' thinking and develop their content and pedagogical content knowledge. In the context of affective dispositions, teachers' beliefs are crucial for teachers' perceptions of situations and the decisions they make about how to act (Felbrich et al. 2012). Moreover, beliefs have been associated with teachers' choice of teaching methods and with student achievement (Bromme 2005; Leder et al. 2002). Thus, in the present study, we explore the relationship of PSTs' diagnostic competence to (1) their beliefs about the nature of mathematics and mathematics teaching and learning and (2) several variables showing PSTs' professional knowledge.

In order to develop PSTs' diagnostic competence, it seems promising to provide PSTs with opportunities to analyze and discuss real classroom situations and students' understanding without the overwhelming complexity of real situations and the pressure to make in-the-moment decisions (Barth and Henninger 2012; Heitzmann et al. 2019). Jacobs and Philipp (2004) emphasize that the value of analyzing students' work in teacher education lies in the discussions it can elicit. Videos have been used to develop PSTs' competences (i.e., McDuffie et al. 2014; Santagata and Guarino 2011; van Es et al. 2017; for an overview, see Santagata et al. 2021). In university courses, videos can be used to facilitate discussions about students' thinking and relevant issues about teaching and learning mathematics. Since everyone in the course is observing the same situation and videos can be paused or shortened, they are helpful to focus PSTs' attention on significant aspects (Star and Strickland 2008; Santagata and Guarino 2011). They provide an opportunity to focus attention on mathematics and students' thinking (van Es et al. 2017) and to shift PSTs' analyses of pedagogical situations to an interpretative stance (Sherin and van Es 2005).

3 Aims and Research Questions

The purpose of this study was to describe PSTs' competence to formulate hypotheses about the causes of students' errors. Our goal was to test predictors of this aspect of diagnostic competence and its development. Based on Blömeke et al.'s (2015) model, we expect that knowledge and beliefs will be positively correlated with PSTs' competence to formulate hypotheses. We are also interested in exploring the role of practical experience in the characteristics and development of PSTs' competence. We posed the following research questions:

1. What are the characteristics of PSTs' competence to hypothesize about the causes of students' errors? To which types of causes do they attribute students' errors?
2. Are beliefs, knowledge, and teaching experience related to PSTs' competence to hypothesize about the causes of students' errors?

3. Are beliefs, knowledge, and teaching experience related to increases in PSTs' competence to hypothesize about the causes of students' errors after participating in a seminar sequence within initial teacher education?

4 Methods

Our study investigated the characteristics of PSTs' competence to hypothesize about the causes of students' errors and the development of this competence after participating in a university seminar sequence. The seminar sequence aimed at developing PSTs' competence to identify, interpret, and respond to students' errors.

The participating PSTs were 131 undergraduates (127 female, 4 male, $M = 22.2$ years old, $SD = 3.69$, from 18 to 42) from 11 Chilean universities. They were in their first to tenth semester ($M = 5.6$ semesters, $SD = 2.53$) of their initial teacher education for primary schools. In Chile, teacher education for primary school teachers lasts between eight and ten semesters and focuses on school grades one to four or one to six. PSTs are educated as generalists, though mathematics is a core subject. In addition, some universities offer the option to take a set of supplementary courses (equivalent in duration to two semesters) focusing on content and pedagogical issues related to lower secondary grades. This path allows PSTs to teach in grades five to eight. Of the 41 participants of the study who chose this path for mathematics education, 26 had already taken some courses. Usually, the courses include school-based activities from the beginning of the program. During the first semesters, PSTs adopt an observant and passive role. As they progress through the program, they adopt a more active role that includes several hours of teaching.

4.1 The University Seminar Sequence

The university seminar sequence consisted of four 90-minute sessions. During these sessions, PSTs engaged in individual and collective discussions about students' mathematical thinking and learning in classroom situations in which students' errors arise. Heinrichs and Kaiser's (2018) model of the diagnostic process in error situations was used to structure the seminar sequence. It was introduced to PSTs as a three-step process to be used to support their analyses of students' errors. In addition, an overarching goal of the seminar sequence was to make participants aware of the relevance and potential of errors for mathematics teaching and learning.

The seminar sequence aimed at developing PSTs' diagnostic competence in error situations. It did not intend to cover the wide variety of errors that could arise in primary classrooms or to study any particular error in depth. To limit the breadth of knowledge required to analyze errors, all errors used during the sessions and in the assessment instruments were related to numeracy and operations. As suggested by Jacobs and Philipp (2004), to bring PSTs closer to real classroom situations, the errors were presented using short video vignettes of primary school classrooms and samples of students' written work. The videos showed students working on tasks and making errors. Some of them also included brief interactions with the teacher. These materials were used as prompts for discussions about students' thinking. Discussions

were supported by the teacher educator and a series of questions and stimuli that directed PSTs' focus to students' learning processes and understanding of mathematics. In addition, excerpts from literature related to the errors were provided to support and enrich analyses. The sessions, activities, and errors used as subjects for discussion are described in detail by Larrain and Kaiser (2019).

The activities for the seminar sequence were planned in detail and conducted by the same university lecturer in all settings. The plan included a step-by-step description of all activities and questions to guide the discussions. This ensured consistency between the participating groups. At three universities, the seminar sequence was embedded in PSTs' regular academic activities. Their participation in surveys was voluntary. At the fourth university, the seminar sequence was offered as an extracurricular activity, in which PSTs from eight different universities participated. Due to the great difficulty in finding settings that allowed us to implement the seminar sequence as part of academic activities and the need to gain a large number of participants to conduct the quantitative analyses planned for the study, it was not possible to create a control group to compare outcomes. Moreover, the universities that agreed to participate were specifically interested in including error analysis in their curriculum, so other topics would not have been accepted for a control seminar sequence.

4.2　Instruments

Data were collected before and after the seminar sequence. At the first measurement point, primary PSTs completed an online survey that included a survey about demographic data and information about their learning opportunities, two beliefs questionnaires, and two error analysis tasks. Additionally, participants answered the mathematical knowledge for teaching (MKT) assessment in paper-and-pencil format. The instrument at the second measurement point included two error analysis tasks and a short survey about participation during the sessions.

4.2.1　Demographics and Learning Opportunities Survey

The demographics and learning opportunities survey collected general data about the PSTs, such as their gender, age, and university entrance exam score. Information was also collected on the opportunities they had to develop professional knowledge and gain practical teaching experience. Specifically, the survey asked about their progress within the program (as measured by the number of semesters they had completed), completed coursework on mathematics or mathematics education, school practicum, experience teaching as generalists in primary classrooms, experience teaching mathematics in primary classrooms, and experience with private tutoring. As these are the usual questions, we did not include the instrument in the online appendix.

4.2.2 Beliefs Questionnaires

To collect data about PSTs' beliefs, we used the beliefs items from the international Teacher Education and Development Study in Mathematics (TEDS-M) questionnaires (Tatto et al. 2008), as Chile had participated in the TEDS-M. Participants were asked to express their level of agreement with a series of statements on a six-point Likert scale. These statements concerned two constructs: beliefs about the nature of mathematics and beliefs about the teaching and learning of mathematics. The Chilean translations of the items were used (available in Ávalos and Matus 2010) and given to participants in an online format. As these are standardized instruments not developed by us, they are not included in the online appendix.

The questionnaire of beliefs about the nature of mathematics gathered information about how mathematics was perceived as a subject. It consisted of two scales. Items in the first scale reflected a view of mathematics as a strongly structured subject in which practice, memory, and precise application of procedures are fundamental for correctly solving tasks. The items on the second scale reflected a view of mathematics as an inquiry process, in which rules and concepts can be flexibly applied to discover connections and solve mathematical problems with practical relevance.

The questionnaire on beliefs about learning mathematics focused on views about teaching strategies, students' thinking processes, and the purposes of learning the subject. It also had two scales. One scale reflected the view that learning mathematics is a strongly teacher-directed process in which students learn by following instructions and memorizing formulas and procedures. The second scale reflected a student-centered view of learning mathematics in which mathematics learning occurs when students are involved in inquiries and develop their own strategies to solve mathematical problems. In this view, importance is given to the understanding of concepts and why procedures are correct.

The internal consistency of the scales was confirmed in the present study (Cohen et al. 2007). For beliefs about the nature of mathematics, the scale of mathematics as a set of rules and procedures showed acceptable reliability (0.78) and the scale of mathematics as a process of inquiry displayed very high reliability (0.91). For beliefs about mathematics learning, the teacher-centered scale showed high reliability (0.84) and the active learning scale presented very high reliability (0.91).

4.2.3 MKT Questionnaire

To measure PSTs' professional knowledge, we used a version of the MKT test developed within the Learning Mathematics for Teaching Project (Hill et al. 2008) that was adapted and validated for use with pre-service Chilean teachers within the Refip project (Martínez et al. 2014). The selected test booklet focuses on evaluation of specialized content knowledge (i.e., the knowledge about mathematics required by teachers), specifically numbers and operations for primary school students. This version was chosen because all the examples of errors used during the sessions and in the error analysis tasks at the pre- and post-test were from this domain of mathematics. The instrument consisted of 15 multiple-choice items and 9 complex

multiple-choice questions with 3 or 4 items each. In total, the test included 45 items that the participants had to answer in 60 min.

Participants' answers were analyzed using IRT methods. The estimated Rasch model suggested a good fit. All items showed a weighted MNSQ (Infit) within the 0.8–1.2 range suggested by Linacre (2002), with a mean of 0.997 ($SD = 0.067$). The EAP/PV value of 0.825 is above 0.7 (Neumann et al. 2011), confirming sufficient reliability. Item difficulty ranged from –2.057 to 2.066, with a mean of 0.1 ($SD = 1.04$). Taking the estimated difficulties as a base, individual persons' ability parameters were estimated, showing a mean of –0.09 ($SD = 0.861$). For ease of further interpreting the data, ability parameters were converted into T-scores by standardizing them to a mean of 0 and a standard deviation of 1, and then scaling them to a mean of 50 and a standard deviation of 10 as this was used in studies of the TEDS-M research program.

4.2.4 Diagnostic Competence in Error Situations

To evaluate PSTs' diagnostic competence in error situations, four error analysis tasks were developed. Each task was based on a mathematical error made by a primary school student and consisted of several items that followed the model of the diagnostic process in error situations (Heinrichs 2015). Participants were asked to respond to two tasks at each testing time. In order to randomly assign the tasks and ensure that no one received the same task twice, four versions of the instrument were created with different combinations of two tasks. In the pre-test, versions were randomly assigned to participants. In the post-test, each participant was assigned the version that contained the two tasks they had not yet completed. This approach incorporates methodological limitations, as the tasks may differ in difficulty and implies a limitation of the study.

The four error situations selected for the tasks concerned numbers and operations taught in primary school (original tasks and questions available in the annexes). Each task began with contextual information that included the grade level, the content of the lesson, and what had been done previously in the class in which the error occurred. The error situation was then presented in a short video vignette. The video showed the student making the error and explaining the procedure they used. Videos could be viewed several times until the respondent moved to the next page of the test.

To analyze their competence to hypothesize about the causes of students' errors, PSTs were asked to suggest three possible causes for the student's error. Participants provided their hypotheses in an open-response item format (Fig. 1). These are the focus of the present article. There was no time pressure to answer the items. However, to reduce the overall time needed to answer the questionnaires and limit the complexity of the item, the number of hypotheses was restricted to three.

4.3 Data Analysis

Data analysis was performed in several phases and used different methods. Responses to the open-ended items were coded using qualitative text analysis. Closed

It can be stated that Isabel has added in the following way:

	T	U
	c	a
+	d	b
	c+d	a+b

$$98 \\ + 87 \\ \overline{1715}$$

Please suggest, from a mathematics didactics perspective, three possible causes for this error.

Cause 1

Cause 2

Cause 3

Fig. 1 Open-ended item measuring PSTs' competence to hypothesize about the causes of students' errors

items and the relationships between features of PSTs' diagnostic competence in error situations and background variables were analyzed using quantitative methods.

4.3.1 Competence to Hypothesize About the Causes of Students' Errors

Open-ended items were used to measure PSTs' competence to hypothesize about the causes of students' errors. The responses to these items were short pieces of text in which PSTs suggested three possible causes for the error presented in the video. Qualitative text analysis was chosen to analyze these complex data. This method provides a systematic framework to interpret and reduce the complexity of answers, so their most relevant characteristics can be considered when developing an answer to the research question (Kuckartz 2019).

Thematic categories were used to code the hypotheses suggested by PSTs. The categories were constructed in a multi-stage process based on the thematic qualitative text analysis process suggested by Kuckartz (2014). First, responses were coded using major categories derived from the research interests and then explored to identify new categories that emerged from the data. Each of the three hypotheses was considered a single unit to be coded. Where respondents referred to their other responses, these were kept together and coded sequentially to allow for full interpretation.

Three categories of hypotheses were defined in the first coding stage. One category was created for hypotheses attributing the error to a failure of teaching strategies or inadequate instructional decisions. Two categories were created for hypotheses attributing the error to aspects of students' mathematical thinking: one in which hypotheses suggested a lack of conceptual understanding of the concepts or processes

Table 1 Categories of hypotheses about the causes of students' errors

Category	Description	Examples
Lack of *conceptual* understanding	The error is attributed to a lack of conceptual understanding of mathematical issues related to the error situation. The error may indicate, for instance, that the student does not conceptually understand the algorithm and the related idea of regrouping, or that the student fails to understand place value	*"He didn't understand place value."* *"She does not understand that the ones and the tens of the same number are related and belong together."* *"He does not understand the decimal number system."*
Lack of *procedural* understanding	The error is attributed to a lack of procedural understanding of mathematical issues. For instance, it may be attributed to difficulties handling an algorithm	*"Does not know how to carry over."* *"Doesn't know that in the units column there must continue being only units, thus she makes the sum and puts a number of 2 digits in that space."*
Instructional wrong decisions	The error is the consequence of inappropriate pedagogical decisions or the use of inadequate or insufficient strategies. The teacher or curriculum is to blame for the error	*"She has not been taught what hundreds, tens, and units are, that is, place value."*
Ambiguous	A reasonable explanation of a cause is given, but because it is too general or not sufficiently described, it is not possible to classify it within the previous categories	*"She doesn't know addition with regrouping."* It is not clear whether what the student does not know is the procedure (algorithm) or the meaning of regrouping, or whether she has not been taught addition with regrouping
Not valid	The response does not give a valid cause for the error: the statement is incorrect, incomplete, or contradictory; it provides a description of the error instead of a cause for it; it suggests a very general cause unrelated to the learning of mathematics	*"Didn't pay attention to the teacher's explanation."* *"It is the easiest way to add up."*

involved in the task, and one in which errors were attributed to a lack of understanding of mathematical procedures. Initial examination of the responses confirmed that the data reflected these three categories. However, it was clear that another two categories were necessary. First, we created a category for responses that did not provide a possible cause for the error, contained incorrect statements, or attributed the error to general causes, such as students' inattentiveness. Second, we created a category for responses that, although plausible and correct, were ambiguous and could not be coded into any of the first three categories. The descriptions and examples of the categories shown in Table 1 were used to analyze the answers to the four error analysis tasks.

A coding manual was developed before starting the data evaluation phase. This manual included precise definitions of each category, anchor examples extracted from the data illustrating the responses that belonged to each category, and coding rules that referred to border cases. To test the coding manual, 20% of the responses were coded by two coders. Critical points were identified, and more precise delineations of categories, and further examples were included in a second version of the manual. In the next stages, parts of the data were independently evaluated by two coders and then compared using a consensual coding approach. In the third cod-

ing cycle, an independent coding approach was adopted, and data were evaluated separately by both coders. Intercoder agreement was controlled after the complete dataset was evaluated. Cohen's Kappa values for each category in the four error tasks were found to be over 0.7 and were interpreted as adequate (Mayring 2000).

For each error, task respondents were asked to suggest three different hypotheses about causes. Thus, a maximum of six causes from each participant were available for coding. It was possible for a single respondent to make suggestions from different perspectives. This yielded various combinations of categories for each PST. Some PSTs showed a tendency to attribute the error to a lack of conceptual understanding, a lack of procedural understanding, or inappropriate instructional decisions, whereas others did not show any particular tendency and provided varied hypotheses. Thus, we considered the number of valid hypotheses about causes (i.e., causes with any of the first four codes in Table 1) when performing further correlation analyses. In other words, we related PSTs' background variables to the number of valid hypotheses they were able to make.

4.3.2 Associations Between PSTs' Competence to Hypothesize About the Causes of Errors and Background Characteristics

Pearson correlations and independent-samples t-tests were used to examine the associations between PSTs' competence to hypothesize about the causes of students' errors and individual characteristics. The number of valid hypotheses before participation in the seminar sequence was treated as a quasi-continuous dependent variable. Pearson correlation coefficients were calculated for continuous independent variables, such as beliefs, mathematical knowledge for teaching, study progress, completed coursework on mathematics or mathematics education, and school practicum. For dichotomous independent variables, independent-samples t-tests were used to evaluate differences in the means of the hypothesized causes for the two groups defined by the variable. These groups were defined by their experience teaching, for example, a group with experience teaching in primary classrooms and a group without such experience.

To further examine the associations between the independent variables and PSTs' competence, a multiple regression analysis was conducted. Due to the large number of available predictor variables, we were interested in developing a simple, exploratory model that could be used to predict PSTs' competence to hypothesize about the causes of errors. To do so, after analysis of each independent variable and the dependent variable, the most relevant predictors were identified. Several potential models, including these predictors, were calculated and then compared using the global goodness-of-fit criterion, the Akaike information criterion (AIC). A final and most parsimonious model that represented the data was selected using these criteria (Fahrmeir et al. 2013).

4.3.3 Development of PSTs' Competence to Hypothesize About the Causes of Errors and Background Characteristics

We expected PSTs to be able to make more valid hypotheses about the causes of students' errors in the post-test than in the pre-test, after their participation in the seminar sequence. To evaluate whether these changes could be associated with the independent variables, further analyses were conducted. Multiple regression analyses were run for the independent continuous variables. In each of these analyses, the pre-test value and one independent variable were included as predictors. Several analyses of covariance (ANCOVAs) were conducted to examine the associations between changes in competence and categorical independent variables. As an extension of the one-way analysis of variance, the ANCOVA incorporates a covariate variable to adjust the means of each group on the dependent variable because the covariate influences the dependent variable and needs to be statistically controlled for (Hatcher 2013; O'Connell et al. 2017).

Finally, multiple regression analyses were conducted to further study the variables that may predict greater changes in PSTs' competence to hypothesize about the causes of students' errors. Various models, including different predictors, were examined following a similar procedure to the one described at the end of Sect. 4.3.2.

5 Results

5.1 Characteristics of PSTs' Competence to Hypothesize About the Causes of Students' Errors

In the following section, we attempt to answer research question 1. Before characterizing PSTs' tendencies to attribute students' errors to instructional decisions, lack of conceptual understanding, or lack of procedural understanding, it was necessary to verify that the types of proposed causes did not depend on a particular error task. In our complex testing design, PSTs answered different error tasks at the two measurement time points, and not all PSTs answered the same error tasks at a given time point. Cross-tabulation of the four error tasks with the five categories of hypotheses and its corresponding test of independence yielded a statistically significant association between the error tasks and the types of hypotheses suggested by PSTs, χ^2 (12,

Table 2 Frequencies of types of hypotheses about the causes of errors at the pre- and post-tests

Hypothesis category	Pre-test		Post-test	
	Frequency	%	Frequency	%
Lack of *conceptual* understanding	109	27.7	125	31.8
Lack of *procedural* understanding	50	12.7	87	22.1
Wrong *instructional* decisions	36	9.2	24	6.1
Ambiguous	31	7.9	32	8.1
Not valid	167	42.5	125	31.8
Total	393	100.0	393	100.0

N = 1572) = 48.073, $p < 0.01$. Subsequent analyses permitted allocation of the error tasks that caused the dependency and supported the decision to select two error tasks for further analyses. Both selected tasks were related to subtraction with whole numbers and were not combined in any version of the test, so respondents were given one at the pre-test and the other at the post-test. Cross-tabulation and the test of independence between these tasks and the types of hypotheses indicated that there was no statistically significant association between them, χ^2 (4, N = 786) = 9.648, $p > 0.01$. Therefore, we considered only these two error tasks when characterizing the types of hypotheses that PSTs proposed about the causes of students' errors.

Table 2 displays the frequencies of the categories of hypotheses at both measurement times. Before the university seminar sequence, a large proportion (42.5%) of the responses were not valid hypotheses. On average, participants were able to formulate 1.7 hypotheses about the causes of students' errors (out of the three that were requested). Most valid causes (27.7%) attributed the error to a lack of conceptual understanding.

Changes from the pre-test to post-test indicate that the proportion of invalid hypotheses decreased considerably (10.7%). Hypotheses attributing errors to inappropriate instructional decisions also decreased (3.1%) after participation in the seminar sequence. Hypotheses ascribing the errors to a lack of conceptual understanding or a lack of procedural understanding increased (4.1 and 9.4%, respectively). This suggests that, at the post-test, PSTs could better focus on students' thinking when interpreting errors.

To explore how many hypotheses PSTs were able to suggest, responses to all four error tasks were considered because the total number of hypotheses was not dependent on the error analyzed. The mean number of hypotheses produced at the pre-test was 3.24 ($SD = 1.67$), whereas the mean at the post-test was 3.92 ($SD = 1.48$). This is an average increase of 0.69 hypotheses from before to after participation in the seminar sequence. Further analysis indicates that 52.7% of the participants generated more hypotheses about the causes of students' errors at the post-test than at the pre-test. Of those who produced the same number of hypotheses at both testing times, more than two-thirds were able to produce four or more hypotheses each time. Altogether, PSTs' competence to hypothesize about the causes of students' errors improved significantly from before to after their participation in the seminar sequence: $t(130) = -4.647$, $p < 0.001$, $d = 0.436$.

5.2 Associations with PSTs' Beliefs, Knowledge, and Teaching Experience

To answer research question 2, which aims to better understand PSTs' competence to hypothesize about the causes of students' errors, the association of this competence with several background variables was explored. Various features of PSTs' beliefs, knowledge, and teaching experience were included. Several bivariate analyses were carried out, taking the number of valid hypotheses produced at the pre-test as the dependent variable. The conducted correlation analyses showed that all the explored independent variables were significantly positively related to competence to make hypotheses about the causes of errors. The number of one-semester school practicums showed the smallest correlation, but was significant and in the expected

Table 3 Description statistics of the number of valid hypotheses by teaching experience group

Experience variable	n	M	SD	SE
Teaching experience—generalist				
Without experience	38	2.32	1.66	0.27
With experience	93	3.61	1.53	0.16
Teaching experience—mathematics				
Without experience	49	2.65	1.74	0.25
With experience	82	3.59	1.54	0.17
Private lesson experience				
Without experience	66	2.91	1.61	0.10
With experience	65	3.57	1.69	0.21

direction ($r = 0.293$, $p < 0.001$). Moderate correlations were found for PSTs' beliefs about the nature of mathematics as a process of inquiry ($r = 0.367$, $p < 0.001$), their mathematical knowledge for teaching ($r = 0.328$, $p < 0.001$), their completed coursework on mathematics or mathematics education ($r = 0.326$, $p < 0.001$), their university entrance test score ($r = 0.326$, $p < 0.001$), and their study progress ($r = 0.346$, $p < 0.001$). The strongest correlation was found with PSTs' beliefs about the learning of mathematics as an active and student-centered process ($r = 0.436$, $p < 0.001$).

To explore the association between competence and PSTs' practical teaching experiences, the means of the provided hypotheses were calculated for groups with and without practical experience (see Table 3). Independent-samples t-tests revealed that PSTs with experience teaching as generalists in primary schools provided significantly more hypotheses about the causes of students' errors than PSTs without experience, $t(129) = -4.289$, $p < 0.001$. Cohen's d revealed a large effect size, $d = 0.808$. Similarly, significant differences were found between PSTs with and without experience in teaching, specifically mathematics in primary classrooms, $t(129) = -3.195$, $p = 0.002$. There was a medium effect size, Cohen's $d = 0.572$. For the groups with and without experience giving private lessons to primary students, a statistically significant difference with a medium size effect was found, $t(129) = -2.295$, $p = 0.023$, Cohen's $d = 0.400$. This indicates that PSTs with experience were able to provide more hypotheses than those without it.

5.2.1 Predictors of PSTs' Competence to Hypothesize About the Causes of Students' Errors

A multiple regression analysis was conducted to better understand the relationship between PSTs' competence to hypothesize about the causes of students' errors and background characteristics. As all the independent variables showed significant associations in the bivariate initial analyses, they were entered into the model simultaneously as predictor variables. After controlling for multi-collinearity, eight independent variables were entered into the model as predictors (see Model 1 in Table 4). The model significantly predicted competence to hypothesize about the causes of students' errors, $F(8, 122) = 5.623$, $p < 0.001$, adj. $R^2 = 0.269$. In this model, most predictors contributed only slightly to the regression, and only beliefs about math-

Table 4 Regression models for the competence to hypothesize about the causes of students' errors

Model Predictors	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Beliefs—Learning of mathematics	x**	x***	x**	x**	x**	x***	x***	x***	x***	x***	x***	**x****	x***	x**	x***
Beliefs—Nature of mathematics	x	x	x	x	x	–	–	–	–	–	–	**x**	–	x	–
Teaching experience—Mathematics	x	x	x	–	–	–	–	–	–	–	–	**–**	x	–	–
Teaching experience—Generalist	x	x	x*	x	x**	x	x**	x**	x*	x*	x**	**x****	x*	–	–
Private lesson experience	x	–	–	–	–	–	–	–	–	–	x	**–**	–	–	–
School practicums	x	x	–	–	–	–	–	–	–	x	–	**–**	–	–	–
Mathematics coursework	x	x	x	x	–	x	–	–	x	–	–	**–**	–	–	–
MKT	x	x	x	x	x	x	x	–	–	–	–	**–**	–	x*	x**
AIC	110.64	108.74	106.88	105.76	104.82	104.63	104.29	103.95	103.79	105.89	105.18	**103.93**	105.87	108.02	107.63
Adjusted R^2	0.269	0.269	0.268	0.263	0.257	0.258	0.249	0.239	0.252	0.239	0.244	**0.251**	0.240	0.227	0.217

x denotes the predictors included in the model

*p<0.10, **p<0.05, ***p<0.001

Table 5 Regression coefficients of the selected model

Independent variables	B	SE B	β	p
Constant	1.499	0.333	–	0.000
Beliefs—Nature of mathematics	0.392	0.278	0.136	0.162
Beliefs—Learning of mathematics	0.754	0.266	0.280	0.005
Teaching experience—Generalist	0.817	0.301	0.223	0.007

B unstandardized regression coefficients, *SE B* standard error of B, *β* standardized regression coefficients

ematics learning as a student-centered process significantly added to the prediction, $\beta = 0.24$, $p = 0.027$.

To find a more parsimonious model, we followed the approach suggested by Fahrmeir et al. (2013). We examined 15 regression models, summarized in Table 4. All the models included the variable of beliefs about mathematics learning, as it significantly contributed to the regression. The AIC was used as a criterion to select model number 12, which significantly predicted the competence under study, $F(3, 127) = 14.173$, $p < 0.001$ and explained 25.1% of the variance.

The estimated coefficients of the selected model, displayed in Table 5, indicate that the selected model contained only two significant regression coefficients that differ from zero at the 5% level: student-centered beliefs about mathematics learning ($\beta = 0.28$, $p = 0.005$) and generalist teaching experience ($\beta = 0.22$, $p = 0.007$). Beliefs about the nature of mathematics as an active process had a non-significant regression coefficient of small effect size, and it explained some of the variance in competence. Thus, if both belief-related variables remained unchanged, PSTs with more experience teaching primary school students were able to provide more hypotheses about the causes of students' errors. Similarly, if their beliefs about the nature of mathematics and their experience teaching as generalists are controlled for, PSTs who held stronger constructivist beliefs about learning mathematics performed better on the test.

5.3 Development of PSTs' Competence to Hypothesize About the Causes of Students' Errors

Several multiple regression analyses were conducted to answer research question 3 and further investigate the changes in competence described in Sect. 5.1 and their association with PSTs' characteristics. Each analysis considered the number of hypothesized causes at the post-test as the dependent variable and included two predictors: the number of hypotheses at the pre-test and the value of the trait under investigation. As expected, the number of hypothesized causes at the pre-test was a significant predictor in all models. In the model that included PSTs' beliefs about the nature of mathematics as an inquiry process as a second predictor, this was a significant predictor ($\beta = 0.200$, $p = 0.019$). Together, beliefs and the pre-test value, explained 20.6% of the variance. That is, when controlling for pre-test competence, PSTs who held stronger beliefs about mathematics as an inquiry process showed greater improvements in their competence to hypothesize about the causes of students' errors. Contrary to our expectations, student-centered beliefs about mathematics learning

did not significantly contribute to prediction in the corresponding model ($\beta = 0.114$, $p = 0.199$).

General academic abilities, as measured by university entrance test score, was a significant predictor ($\beta = 0.226$, $p = 0.007$). Along with pre-test competence, it explained 21.7% of the variance. In contrast, MKT was not a statistically significant predictor in the model in which it was included ($\beta = 0.067$, $p = 0.432$). Variables related to PSTs' opportunities to learn significantly contributed to the models in which they were included as a second predictor along with the pre-test number of hypotheses. PSTs' study progress significantly added to the model ($\beta = 0.351$, $p < 0.001$). In other words, participants who had completed more semesters of their studies showed a greater increase in the number of hypotheses they could produce when the number of pre-test hypotheses was held constant. Similarly, PSTs' finished coursework on mathematics or mathematics education significantly contributed to predicting the number of hypotheses produced at the post-test ($\beta = 0.270$, $p = 0.001$). Together with pre-test competence, finished coursework explained 23.7% of the variance. The same was true in the model that included one-semester school practicums as a predictor ($\beta = 0.273$, $p = 0.001$).

To determine the influence of teaching experience on the changes on PSTs' competence to hypothesize about the causes of students' errors, differences between groups with and without experience in teaching in primary schools were explored. ANCOVAs were conducted to control for competence at the pre-test. After adjusting for the pre-test measure, statistically significant differences were found between the groups with and without experience teaching as generalists in primary schools, $F(1,128) = 11.105$, $p = 0.001$, partial $\eta^2 = 0.08$. PSTs who had taught in primary school classrooms provided, on average, a higher number of valid hypotheses than inexperienced PSTs, after adjusting for the pre-test value. A similar difference was found between the groups with and without experience teaching mathematics in primary classrooms, $F(1,128) = 12.810$, $p < 0.001$, partial $\eta^2 = 0.091$. Experience giving private lessons to primary school students produced a smaller statistically significant difference at the 5% level, $F(1,128) = 5.161$, $p = 0.025$, partial $\eta^2 = 0.039$. This suggests that PSTs with more practical experience achieved greater changes in their competence to hypothesize about the causes of students' errors than PSTs without teaching or private tutoring experience.

Several regression analyses were performed to investigate the influence of all independent variables on changes on PSTs' competence. The number of hypotheses at the post-test was taken as the criterion variable. In the first model, the pre-test number of hypotheses and all the independent variables with statistically significant results in the previous analyses were entered into the model (see Model 1 in Table 6). This first model significantly predicted the post-test measure, $F(8, 122) = 6.990$, $p < 0.001$, adj. $R^2 = 0.269$. However, most predictors did not significantly contribute to the model, so further analyses were conducted to find a parsimonious model. To select the simplest model, several models were first estimated by including or excluding predictors according to their relevance to the model and then compared using the AIC (Fahrmeir et al. 2013). All models, displayed in Table 6, included the pre-test measure of the number of hypotheses as a predictor. The selected final model (Model 7 in Table 6) includes, in addition to the pre-test measure, beliefs about the

Table 6 Regression models of the development of competence to hypothesize about the causes of students' errors

Model Predictors	1	2	3	4	5	6	7	8	9	10	11	12	13
No. hypotheses pre-test	x**	x**	x**	x**	x**	x**	**x****	x**	x**	x**	x**	x**	x***
Beliefs—Nature of mathematics	x	x	x	x	x	x	x	–	x	x	x	x	–
University entrance test score	x	x	x	x	x	x	–	x	–	–	–	–	x
Study progress	x	–	x**	x**	x**	x**	**x*****	x***	x**	x**	x**	x**	x***
Coursework mathematics	x	x	–	–	–	–	**–**	–	–	x	–	–	–
School practicums	x	x	–	–	–	–	**–**	–	x	–	–	–	–
Teaching experience—Generalist	x	x	x	–	x	–	**–**	–	–	–	x	–	–
Teaching experience—Mathematics	x	x	x	x	–	–	**–**	–	–	–	–	x	–
AIC	69.583	69.622	65.954	63.961	64.229	62.579	**61.870**	62.121	63.143	63.870	63.171	62.815	61.891
Adjusted R^2	0.269	0.264	0.279	0.285	0.283	0.287	**0.286**	0.284	0.284	0.280	0.284	0.286	0.280

x denotes the predictors included in the model

*p<0.10, **p<0.05, ***p<0.001

Table 7 Regression coefficients of the selected model for the development of competence to hypothesize about the causes of students' errors

Independent variables	B	*SE* B	β	*p*
Constant	1.946	0.303	–	0.000
Nr. Hypotheses pre-test	0.242	0.073	0.274	0.001
Beliefs—Nature of mathematics	0.294	0.209	0.116	0.162
Study progress	0.187	0.048	0.321	0.000

B unstandardized regression coefficients, *SE B* standard error of B, β standardized regression coefficients

nature of mathematics and study progress as predictors. The model significantly predicts post-test competence to hypothesize about the causes of students' errors, $F(3, 127) = 18.336$, $p < 0.001$, adj. $R^2 = 0.286$.

The estimated coefficients for the selected model are displayed in Table 7. Both the pre-test measure of competence ($\beta = 0.274$, $p = 0.001$) and study progress ($\beta = 0.321$, $p < 0.001$) contributed significantly to the prediction. Beliefs about the nature of mathematics as a process of inquiry had a statistically non-significant regression coefficient of small effect size, but explained some of the variance and contributed to the model in the expected direction ($\beta = 0.116$, $p = 0.162$). These results indicate that PSTs who provided a higher number of hypotheses in the pre-test were able to make more hypotheses about the causes of students' errors in the post-test when controlling for their beliefs and study progress. In addition, the results indicate that PSTs who were more advanced in their studies provided more hypotheses about the causes of students' errors when controlling for their pre-test performance and beliefs. Similarly, if pre-test performance and study progress are controlled for, PSTs holding stronger beliefs about the nature of mathematics as an inquiry process showed better competence in the post-test.

6 Summary, Discussion, and Limitations of the Study

As classrooms become increasingly heterogeneous, teachers are required to individualize teaching strategies and provide targeted support to students to promote mathematical learning. Teachers' diagnostic competence is crucial to this process, as it allows teachers to understand students' thinking and tailor their pedagogical decisions (Artelt and Gräsel 2009; Brandt et al. 2017). In particular, by analyzing students' errors, teachers can gain valuable insight into students' mathematical understanding and areas in which students need further support. A crucial step in the diagnostic process that teachers follow when analyzing students' errors is interpreting students' thinking by making hypotheses about the causes of an error. The quality of these hypotheses has a strong impact on the quality of the teaching strategies that can be implemented. Providing learning opportunities that promote the development of this competence during initial teacher education has proven to be a challenging endeavor for teacher educators (Cooper 2009; Heinrichs 2015; Son 2013; Türling et al. 2012). Therefore, the study presented in this paper aimed to unpack the characteristics and development of pre-service primary school teachers' competence to hypothesize about the causes of students' errors. To support the de-

velopment of this competence, a university seminar sequence was implemented in several Chilean universities.

In relation to our first research question, we found that primary school PSTs faced difficulties when hypothesizing about the causes of students' errors. They were often unable to provide valid hypotheses and instead gave incorrect, incomplete, or contradictory suggestions; very general causes that were not specific to the error; or a description of the error that did not refer to its cause. This suggests that PSTs were largely unable to interpret students' thinking, making it difficult to provide appropriate pedagogical responses. These results align with findings from other studies involving teachers (Seifried and Wuttke 2010) and PSTs (Türling et al. 2012; Wuttke and Seifried 2013). Most PSTs with valid hypotheses attributed the error to a lack of conceptual understanding of a mathematical issue. To a lesser extent, they ascribed errors to students' lack of procedural knowledge or inappropriate instructional decisions.

Regarding our second research question, all variables related to PSTs' beliefs, knowledge, and teaching experience were significantly associated with their competence to hypothesize about the causes of students' errors. PSTs' beliefs about learning mathematics as a student-centered, active process showed the strongest correlation with their competence to make hypotheses. This result was confirmed in the multiple regression analysis that included PSTs' beliefs about the learning of mathematics and their teaching experience as generalists in primary schools as significant predictors. Beliefs about mathematics learning as an active process are closely related to a constructivist paradigm, in which efforts to understand students' mathematical thinking and interpret their errors play a particularly relevant role. Thus, PSTs who hold such beliefs tend to be more willing to seek out causes of students' errors. The role of teaching experience as a predictor supports findings from other studies that highlighted the relevance of teaching experience for the development of diagnostic competence (Heinrichs 2015; Türling et al. 2012). It is possible that school experiences during initial teacher education may have contributed to a better understanding of young students' thinking processes and the need for a flexible perspective to comprehend their ideas. We also expected that professional knowledge would play an important role; more advanced knowledge of mathematics and mathematics teaching and learning would provide PSTs with more resources to interpret students' thinking. Other studies have shown the relevance of professional knowledge for diagnosing students' mathematical thinking (i.e., Bartel and Roth 2017; Kron et al. 2021; Ostermann et al. 2018; Streit et al. 2019). Although no variable related to professional knowledge was included as a predictor in the final regression model, PSTs' MKT showed a significant correlation with competence and had a positive effect as a predictor in preliminary analyses. We interpret this phenomenon as an interesting indicator for the importance of knowledge for the interpretation of students' errors. However, our final model highlights more the effect of beliefs and experience for the formulation of hypotheses about the causes of errors. This raises some questions for further research, such as whether it is possible to compensate lacking knowledge for error interpretation by experience or error-supporting claims. It is also worth asking whether the ability to take a flexible perspective for interpreting students' thinking plays an important role in

the formulation of hypotheses about the causes of errors and, if so, whether beliefs or teaching experience allow this flexibility to be developed.

The results for the third research question suggest that PSTs' competence to hypothesize about the causes of students' errors can be developed within initial teacher education. There were significant improvements between the pre- and post-tests, not only quantitatively (i.e., an increase in the number of valid hypotheses PSTs were able to formulate at the second testing time) but also qualitatively (i.e., the types of hypotheses they formulated and their quality). At the second testing time, PSTs attributed student errors less frequently to inappropriate instructional decisions and provided a greater proportion of hypotheses that attributed student errors to a lack of conceptual or procedural understanding. This suggests that prospective teachers improved their ability to focus on student thinking. Although the gains cannot be directly attributed to participation in the seminar sequence because the study did not have an experimental design (i.e., no control group) and other factors may have contributed to the development of competence, the results are promising and indicate that PSTs can develop this competence during initial teacher education. This supports evidence from other research studies, which show that interventions during initial teacher education can significantly improve PSTs' ability to diagnose students' thinking (e.g., Phelps-Gregory and Spitzer 2018; Santagata et al. 2021).

The individual analyses indicated multiple associations between PSTs' beliefs, knowledge, and teaching experience and greater changes in competence to make hypotheses about the causes of student errors. The final regression model includes study progress as a significant predictor in addition to baseline competence. PSTs who have progressed further in their studies have had more learning opportunities, both in practical experience and professional knowledge (not only in mathematics and mathematics teaching but also in other areas such as curriculum, assessment, or educational psychology). It is possible to envisage that these opportunities may have provided PSTs with knowledge and skills with which to connect the newly learned content from the seminar sequence.

Although the results provide answers to our research questions, the present study has limitations, and the findings should be interpreted cautiously. The implementation of the seminar sequence varied at the participating universities; at some universities, it was part of the mandatory activities, while at others it was offered as a voluntary extracurricular opportunity. For this reason, and because this was a field study in which the intervention took place in natural settings, it cannot be assumed that the conditions were identical in all groups. In addition, probability sampling techniques could not be used, and we obtained a convenience sample with no control group. Consequently, the study is not representative of primary school PSTs, and the results cannot be generalized. This is sufficient for our exploratory goals, but the results need to be replicated in future studies. Further, the error tasks used in the seminar sequence and those selected for the error analysis tasks in the instrument were all related to elementary-level numbers and operations content. This limits the significance of the results to error diagnosis in this domain of mathematics; the findings cannot be easily transferred to other areas, such as geometry, early algebra, problem solving, or modeling. Finally, because the study has a pre- and post-test

design and no third testing time, the stability of the development of competence was not examined.

Despite these limitations, the results of the present study suggest that the diagnostic competence of primary school PSTs' can be developed within initial teacher education. Furthermore, our findings help to unpack the complexity of PSTs' competence to interpret students' errors by identifying associations between this competence and several individual features. The results provide evidence that competence models should incorporate beliefs and knowledge as aspects of observable performance (Blömeke et al. 2015), and they suggest that practical experience plays an important role in competence. Although further studies are needed to replicate the findings and better understand the process by which PSTs interpret students' errors, how they develop this competence, and how it is affected by various factors, the present study provides valuable insights. It suggests that implementing learning opportunities to develop PSTs' diagnostic competence is a promising endeavor for initial teacher educators. This is especially true for PSTs in the middle or advanced stages of their programs, when they have a base of professional knowledge and have had practical experiences with teaching. The findings of this study also indicate that beliefs play an important role in competence and its development. For this reason, development of an awareness of the learning potential of errors and the importance of focusing on students' thinking for successful teaching appear to be particularly relevant.

References

Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, *23*(34), 157–160.

Ashlock, R. B. (2010). *Error patterns in computation: using error patterns to help each student learn* (10th edn.). Boston: Allyn & Bacon.

Ávalos, B., & Matus, C. (2010). *La formación inicial docente en Chile desde una óptica internacional: informe nacional del estudio internacional IEA TEDS-M*. Santiago de Chile: Ministerio de Educación.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: what makes it special? *Journal of Teacher Education*, *59*, 389–407.

Bartel, M. E., & Roth, J. (2017). Diagnostische Kompetenz von Lehramtsstudierenden fördern. In J. Leuders, T. Leuders, S. Prediger & S. Ruwisch (Eds.), *Mit Heterogenität im Mathematikunterricht umgehen lernen* (pp. 43–52). Wiesbaden: Springer Spektrum.

Barth, C., & Henninger, M. (2012). Fostering the diagnostic competence of teachers with multimedia training—A promising approach? In I. Deliyannis (Ed.), *Interactive multimedia*. Rijeka: InTech.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, *35*(5), 463–482.

Blömeke, S., Gustafsson, J. E., & Shavelson, R. (2015). Beyond dichotomies: competence viewed as a continuum. *Zeitschrift für Psychologie*, *223*(1), 3–13.

Brandt, J., Ocken, A., & Selter, C. (2017). Diagnose und Förderung erleben und erlernen im Rahmen einer Großveranstaltung für Primarstufenstudierende. In J. Leuders, T. Leuders, S. Prediger & S. Ruwisch (Eds.), *Mit Heterogenität im Mathematikunterricht umgehen lernen* (pp. 53–64). Wiesbaden: Springer Spektrum.

Brodie, K. (2014). Learning about learner errors in professional learning communities. *Educational studies in mathematics*, *85*(2), 221–239.

Bromme, R. (2005). Thinking and knowing about knowledge. In M. H. Hoffmann, J. Lenhard & F. Seeger (Eds.), *Activity and sign* (pp. 191–201). Boston: Springer.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive science*, *2*(2), 155–192.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th edn.). Oxon: Routledge.

Cooper, S. (2009). Preservice teachers' analysis of children's work to make instructional decisions. *School Science and Mathematics*, *109*(6), 355–362.

Cox, L. S. (1975). Systematic errors in the four vertical algorithms in normal and handicapped populations. *Journal for Research in Mathematics Education*, *6*(4), 202–220.

van Es, E. A., Cashen, M., Barnhart, T., & Auger, A. (2017). Learning to notice mathematics instruction: Using video to develop preservice teachers' vision of ambitious pedagogy. *Cognition and Instruction*, *35*(3), 165–187.

Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression: models, methods and applications*. Berlin, Heidelberg: Springer.

Felbrich, A., Kaiser, G., & Schmotz, C. (2012). The cultural dimension of beliefs: an investigation of future primary teachers' epistemological beliefs concerning the nature of mathematics in 15 countries. *ZDM—Mathematics Education*, *44*(3), 355–366.

Götze, D., Selter, C., & Zannetin, E. (2019). *Das KIRA-Buch: Kinder rechnen anders. Verstehen und fördern im Mathematikunterricht*. Hannover: Klett.

Hatcher, L. (2013). *Advanced statistics in research: Reading, understanding, and writing up data analysis results*. : Shadow Finch Media, LLC.

Heinrichs, H. (2015). *Diagnostische Kompetenz von Mathematik-Lehramtsstudierenden: Messung und Förderung*. Wiesbaden: Springer.

Heinrichs, H., & Kaiser, G. (2018). Diagnostic competence for dealing with students' errors: Fostering diagnostic competence in error situations. In T. Leuders, K. Philipp & J. Leuders (Eds.), *Diagnostic competence of mathematics teachers* (pp. 79–94). Cham: Springer.

Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M. R., et al. (2019). Facilitating diagnostic competences in simulations: a conceptual framework and a research agenda for medical and teacher education. *Frontline Learning Research*, *7*(4), 1–24.

Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. Bobingen: Klett.

Helmke, A., & Schrader, F. W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching and Teacher Education*, *3*(2), 91–98.

Herppich, S., Praetorius, A. K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Behrmann, L., Böhmer, M., Ufer, S., Klug, J., Hetmanek, A., Ohle, A. Böhmer, I., Karing, C., Kaiser, J., & Südkamp, A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher education*, *76*, 181–193.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: an exploratory study. *Cognition and instruction*, *26*(4), 430–511.

Hoth, J., Döhrmann, M., Kaiser, G., Busse, A., König, J., & Blömeke, S. (2016). Diagnostic competence of primary school mathematics teachers during classroom situations. *ZDM—Mathematics Education*, *48*(1–2), 41–53.

Jacobs, V. R., & Philipp, R. A. (2004). Mathematical thinking: helping prospective and practicing teachers focus. *Teaching Children Mathematics*, *11*(4), 194–201.

Kaiser, G., Busse, A., Hoth, J., König, J., & Blömeke, S. (2015). About the complexities of video-based assessments: theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *International Journal of Science and Mathematics Education*, *13*(2), 369–387.

Kaiser, J., Südkamp, A., & Möller, J. (2017). The effects of student characteristics on teachers' judgment accuracy. *Journal of Educational Psychology*, *109*(6), 871–888.

Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of teachers: a process model that accounts for diagnosing learning behavior tested by means of a case scenario. *Teaching and Teacher Education*, *30*, 38–46.

Kron, S., Sommerhoff, D., Achtner, M., & Ufer, S. (2021). Selecting mathematical tasks for assessing students understanding: Pre-service teachers' sensitivity to and adaptive use of diagnostic task potential in simulated diagnostic one-to-one interviews. *Frontiers in Education*, *6*, 604568.

Kuckartz, U. (2014). *Qualitative text analysis: a guide to methods, practice and using software*. London: SAGE.

Kuckartz, U. (2019). Qualitative text analysis: a systematic approach. In G. Kaiser & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education* (pp. 181–197). Cham: Springer.

Larrain, M., & Kaiser, G. (2019). Analysis of students' mathematical errors as a means to promote future primary school teachers' diagnostic competence. *Uni-pluriversidad*, *19*(2), 17–39.

Leder, C., Pehkonen, E., & Törner, G. (Eds.). (2002). *Beliefs: a hidden variable in mathematics education?* Dordrecht: Kluwer Academic Publishers.

Lenz, K., Dreher, A., Holzäpfel, L., & Wittmann, G. (2020). Are conceptual knowledge and procedural knowledge empirically separable? The case of fractions. *British Journal of Educational Psychology*, *90*(3), 809–829.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean. *Rasch Measurement Transactions*, *16*(2), 878.

Loibl, K., Leuders, T., & Dörfler, T. (2020). A framework for explaining teachers' diagnostic judgements by cognitive modeling (DiaCoM). *Teaching and Teacher Education*, *91*, 103059.

Martínez, F., Martínez, S., Ramírez, H., & Varas, M. L. (2014). *Refip Matemática. Recursos para la formación inicial de profesores de educación básica*. Proyecto FONDEF D09I-1023, Santiago. http://refip.cmm.uchile.cl/files/memoria.pdf

Mayring, P. (2000). Qualitative content analysis. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *1*(2). https://doi.org/10.17169/fqs-1.2.1089

McDuffie, A. R., Foote, M. Q., Bolson, C., Turner, E. E., Aguirre, J. M., Bartell, T. G., Drake, C., & Land, T. (2014). Using video analysis to support prospective K-8 teachers' noticing of students' multiple mathematical knowledge bases. *Journal of Mathematics Teacher Education*, *17*(3), 245–270.

McGuire, P. (2013). Using online error analysis items to support preservice teachers' pedagogical content knowledge in mathematics. *Contemporary Issues in Technology and Teacher Education*, *13*(3), 207–218.

Neumann, I., Neumann, K., & Nehm, R. (2011). Evaluating instrument quality in science education: Rasch-based analyses of a nature of science test. *International Journal of Science Education*, *33*(10), 1373–1405.

O'Connell, N. S., Dai, L., Jiang, Y., Speiser, J. L., Ward, R., Wei, W., Carroll, R., & Gebregziabher, M. (2017). Methods for analysis of pre-post data in clinical research: a comparison of five common methods. *Journal of biometrics & biostatistics*, *8*(1), 1–8.

Ostermann, A., Leuders, T., & Nückles, M. (2018). Improving the judgment of task difficulties: prospective teachers' diagnostic competence in the area of functions and graphs. *Journal of Mathematics Teacher Education*, *21*(6), 579–605.

Phelps-Gregory, C. M., & Spitzer, S. M. (2018). Developing prospective teachers' ability to diagnose evidence of student thinking: Replicating a classroom intervention. In T. Leuders, K. Philipp & J. Leuders (Eds.), *Diagnostic competence of mathematics teachers* (pp. 223–240). Cham: Springer.

Praetorius, A.-K., Lipowsky, F., & Karst, K. (2012). Diagnostische Kompetenz von Lehrkräften: Aktueller Forschungsstand, unterrichtspraktische Umsetzbarkeit und Bedeutung für den Unterricht. In R. Lazarides (Ed.), *Differenzierung im mathematisch-naturwissenschaftlichen Unterricht* (pp. 115–146). Bad Heilbrunn: Klinkhardt.

Prediger, S., & Wittmann, G. (2009). Aus Fehlern lernen – (wie) ist das möglich. *Praxis der Mathematik in der Schule*, *51*(3), 1–8.

Rach, S., Ufer, S., & Heinze, A. (2013). Learning from errors: effects of teachers training on students' attitudes towards and their individual use of errors. *PNA*, *8*(1), 21–30.

Radatz, H. (1980). *Fehleranalysen im Mathematikunterricht*. Braunschweig: Vieweg.

Radatz, H. (1979). Error analysis in mathematics education. *Journal for research in mathematics education*, *10*(3), 163–172.

Reisman, F. K. (1982). *A guide to the diagnostic of teaching of arithmetic* (3rd edn.). Columbus: C. E. Merrill.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: does one lead to the other? *Journal of educational psychology*, *91*(1), 175.

Rittle-Johnson, B., & Koedinger, K. (2009). Iterating between lessons on concepts and procedures can improve mathematics knowledge. *British Journal of Educational Psychology*, *79*(3), 483–500.

Santagata, R. (2005). Practices and beliefs in mistake-handling activities: A video study of Italian and US mathematics lessons. *Teaching and Teacher Education*, *21*(5), 491–508.

Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM—Mathematics Education*, *43*(1), 133–145.

Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A.-K., Yang, X., & Kaiser, G. (2021). Mathematics teacher learning to notice: a systematic review of studies of video-based programs. *ZDM—Mathematics Education*, *53*(1), 119–134.

Scherer, P., & Opitz, M. E. (2012). *Fördern im Mathematikunterricht der Primarstufe*. Berlin Heidelberg: Springer.

Seifried, J., & Wuttke, E. (2010). Student errors: How teachers diagnose them and how they respond to them. *Empirical research in vocational education and training*, *2*(2), 147–162.

Sherin, M., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of technology and teacher education*, *13*(3), 475–491.

Sherin, M., Jacobs, V., & Philipp, R. (Eds.). (2011). *Mathematics teacher noticing: Seeing through teachers' eyes*. New York: Routledge.

Smith, J. P., diSessa, A. A., & Roschelle, J. (1993). Misconceptions reconceived: a constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, *3*(2), 115–163.

Son, J. W. (2013). How preservice teachers interpret and respond to student errors: ratio and proportion in similar rectangles. *Educational studies in mathematics*, *84*(1), 49–70.

Star, J. R., & Strickland, S. K. (2008). Learning to observe: using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics teacher Education*, *11*(2), 107–125.

Streit, C., Rüede, C., Weber, C., & Graf, B. (2019). Zur Verknüpfung von Lernstandeinschätzung und Weiterarbeit im Arithmetikunterricht: Ein kontrastiver Vergleich zur Charakterisierung diagnostischer Expertise. *Journal für Mathematik Didaktik*, *40*(1), 37–62.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: a meta-analysis. *Journal of Educational Psychology*, *104*, 743–762. https://doi.org/10.1037/a0027627.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Türling, J. M., Seifried, J., & Wuttke, E. (2012). Teachers' knowledge about domain specific student errors. In E. Wuttke (Ed.), *Learning from errors at school and at work* (pp. 95–110). Opladen: Budrich.

Weimer, H. (1925). *Psychologie der Fehler*. Leipzig: Julius Klinkhardt.

Wildgans-Lang, A., Scheuerer, S., Obersteiner, A., Fischer, F., & Reiss, K. (2020). Analyzing prospective mathematics teachers' diagnostic processes in a simulated environment. *ZDM—Mathematics Education*, *52*(2), 241–254.

Wuttke, E., & Seifried, J. (2013). Diagnostic competence of (prospective) teachers in vocational education: An analysis of error identification in accounting lessons. In K. Beck & O. Zlatkin-Troitschanskaia (Eds.), *From diagnostics to learning success: proceedings in vocational education and training* (pp. 225–240). Rotterdam: Sense.