Technische Universität München
TUM School of Education
Friedl Schöller-Stiftungslehrstuhl für Pädagogische Psychologie

Approximating Practice

Developing a Video-Based Simulation for Measuring Preservice Teachers' Diagnostic Skills

Elias Julian Codreanu

Vollständiger Abdruck der von der TUM School of Education der
Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Philosophie (Dr. phil.)

genehmigten Dissertation.

Vorsitzende: Prof. Dr. Doris Holzberger

Prüferin und Prüfer der Dissertation:

1. Prof. Dr. Christina Seidel
2. Prof. Dr. Andreas Obersteiner

Die Dissertation wurde am 27.04.2021 bei der Technischen Universität München eingereicht und durch die TUM School of Education am 05.07.2021 angenommen.

Acknowledgments

I would like to express deep appreciation to my primary supervisor, Professor Dr Tina Seidel, for her remarkable support, guidance, and encouragement. With her great expertise and experience, she guided me to develop my own research and made this dissertation possible. My great thanks also go to my mentor and co-author, Dr Daniel Sommerhoff, who invested so much time and effort to support me on this journey. I am deeply grateful to him for his consistent support throughout my research. His constructive feedback and helpful advice led me to improve myself immensely. I appreciate him for always being available for me and my questions. I would like to express my sincere gratitude to Professor Dr Stefan Ufer, who opened the doors to science for me and guided me through my doctoral time with sound feedback. I would also like to thank Professor Dr Andreas Obersteiner, who kindly agreed to review this dissertation and serve on the examination committee, as well as Professor Dr Doris Holzberger for chairing the committee.

I want to thank Sarah Reinhold and Dr Sina Huber for helping me to find my way in the world of science. The exchange with them and their experience oriented me and helped me know where start my research. A big thank you to my colleague and friend, Katharina. I am grateful that we experienced this exciting time together. She always had an ear for me and helped me with her way of thinking and her advice no matter what it was about. Your thoughts and ideas have often inspired me to think further. A big thanks also goes to all of my colleagues at the chair. It was a pleasure working with them, sharing an office, having interesting discussions, or just having a good time. Many thanks also to the whole research group. The interdisciplinary collaboration was very interesting and diversified. I want to thank all the doctoral students for this great community and team spirit.

A big thank you also goes to my family and friends who accompanied me and stood behind me throughout the entire time. I thank Michael, who always took my mind off work. At this point, I also want to thank Mone and Moritz for the endless hours full of music and joy. The moments of making music together have always given me great strength and positivity and carried me throughout my dissertation. My deepest appreciation belongs to my family for accompanying me on this adventure—and the many more that would not be possible without them.

Abstract

Diagnosing students' skills in classroom situations is an important element of the teaching profession. Teachers' diagnostic skills are the basis for adaptive teaching by considering students' learning requirements. One of the central tasks in classrooms is to diagnose learning-relevant student characteristics in "on-the-fly" situations to support the students individually. However, it is mostly unclear, which diagnostic activities teachers pursue in such situations. According to the high relevance for practice, there have been several calls for the development of methods for measuring (and subsequently supporting) preservice teachers' diagnostic skills. This dissertation answers these calls by developing a digital environment in form of a simulation utilized in two separate, but related, studies. A newly developed video-based simulation for measuring preservice mathematics teachers' diagnostic skills is presented here. In the simulation, participants are supposed to diagnose students' proving skills in geometry by watching video clips. The simulation was developed as a tool to investigate diagnostic skills in a standardized and ecologically valid way. In the first presented study (Article A), the video-based simulation was validated regarding two central aspects: (i) the implementation of an authentic environment with a close real-world relationship and (ii) the cognitive demand of the situation and its adaptation to the preservice teachers' skills by reducing the complexity of the simulated situation. The second presented study (Article B) focused on preservice teachers' diagnostic activities and examined, which information preservice teachers noticed when observing students. The analysis included observing what diagnostic activities the students demonstrated in the process of knowledge-based reasoning. The results of the validation revealed that the simulation was perceived as an authentic representation of a real diagnostic situation by preservice teachers and expert teachers. An analysis of the preservice teachers' performance considering relative and absolute accuracy suggests that the simulated diagnostic situation has an appropriate level of cognitive demand. Analyses focusing on diagnostic activities found that the preservice teachers noticed and described relevant information in the diagnostic situation, but linking this information to professional concepts and knowledge as well as drawing conclusions were mostly missing. Overall, the results illustrate the potential of the video-based simulation for measuring and exploring diagnostic skills in a contextualized and valid way. Results showed room for improvement of preservice teachers' diagnostic skills so that the simulation can serve as a basis for the development of a learning environment.

Zusammenfassung

Das Diagnostizieren von Schülerfähigkeiten im Unterricht ist ein wichtiges Element des professionellen Handelns von Lehrkräften. Die Fähigkeit zu diagnostizieren gilt als Grundvoraussetzung, um den Unterricht adaptiv zu gestalten und an die individuellen Lernvoraussetzungen von Schülerinnen und Schülern anzupassen. Eine der zentralen Aufgaben im Schulkontext besteht darin, lernrelevante Schülermerkmale während des Unterrichtsgeschehens zur individuellen Förderung der Schülerinnen und Schüler zu erkennen und einzuschätzen. Dabei ist bisher aber noch weitgehend offen, wie angehende Lehrkräfte in solchen Diagnosesituationen agieren. Angesichts der hohen praktischen Relevanz des Themas wird des Öfteren gefordert, die Methoden zur Erfassung und Erforschung der Diagnosefähigkeiten von Lehramtsstudierenden weiterzuentwickeln. Die vorliegende Dissertation leistet durch die Entwicklung einer digitalen Umgebung in Form einer Simulation einen Beitrag dazu. Es wird eine neuentwickelte videobasierte Simulation zur Messung der Diagnosefähigkeiten von angehenden Mathematiklehrkräften vorgestellt. Die Lehramtsstudierenden sollen darin anhand von Videos diagnostizieren, welche Fähigkeiten Schülerinnen und Schüler im mathematischen Beweisen haben. Ziel war es, ein Tool zu entwickeln, bei dem praxisnah die Diagnosefähigkeiten standardisiert gemessen und untersucht werden können. In der ersten Studie (Artikel A) wurde die videobasierte Simulation anhand zweier Aspekte validiert. Dabei wurde zum einen auf die Umsetzung einer authentischen Umgebung mit einem hohen Praxisbezug geachtet. Zum anderen wurde untersucht, inwieweit die Anforderungen der simulierten Diagnosesituation durch eine Reduzierung der Komplexität im Vergleich zu einer realen Diagnosesituation an die Fähigkeiten der Lehramtsstudierenden angepasst wurden. Die zweite Studie (Artikel B) beschäftigt sich mit den diagnostischen Aktivitäten von angehenden Mathematiklehrkräften beim Diagnostizieren. Es wurde untersucht, auf welche Inhaltsbereiche die Lehramtsstudierenden sich fokussierten um Informationen über Schülerinnen und Schüler zu sammeln. Darüber hinaus wurde analysiert, welche diagnostischen Aktivitäten die Lehramtsstudierenden im Laufe des Schlussfolgerns hin zu einer Diagnose zeigten. Die Ergebnisse der Validierung ergaben, dass die Simulation sowohl von Studierenden als auch von Experten als eine authentische Abbildung einer realen Diagnosesituation wahrgenommen wurde. Eine Analyse der Leistungen, die die Studierenden in der Simulation zeigten, lässt den Schluss zu, dass die simulierte Diagnosesituation ein angemessenes

Anforderungsniveau für angehende Mathematiklehrkräfte hat. Bei der Untersuchung der diagnostischen Aktivitäten zeigte sich, dass die Lehramtsstudierenden zwar relevante Inhalte wahrnahmen und diese beschrieben, Verknüpfungen zu Konzepten und Wissen aus dem Studium sowie weiterführende Schlussfolgerungen meist fehlten. Zusammengenommen unterstreichen die Ergebnisse der Studien das Potential der Simulation für die Messung und Erforschung von Diagnosefähigkeiten. Die Ergebnisse zeigten Stärken und Schwächen in den Diagnosefähigkeiten angehender Mathematiklehrkräfte, sodass die Simulation als Grundlage zur Entwicklung einer Lernumgebung dienen kann.

Included Publications

The present dissertation is written cumulatively and consists of two articles published in international peer-reviewed journals. The author of this dissertation is the first author of both publications and was primarily responsible for conceptualization, data analyses, writing the original and the revised manuscripts, and pursuing the publication process at the respective journals.

Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, 95, 103146. doi:10.1016/j.tate.2020.103146
(Article **A**)

The first author was primarily responsible for the publication of this article (70%). The co-authors, Dr. Daniel Sommerhoff, Dr. Sina Huber, Prof. Dr. Stefan Ufer, and Prof. Dr. Tina Seidel, contributed to the development and publication of the article with critical reviews.

Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2021). Exploring the process of preservice teachers' diagnostic activities in a video-based simulation. *Frontiers in Education*, 6, 626666. doi:10.3389/feduc.2021.626666
(Article **B**)

The first author was primarily responsible for the publication of this article (80%). The co-authors, Dr. Daniel Sommerhoff, Dr. Sina Huber, Prof. Dr. Stefan Ufer, and Prof. Dr. Tina Seidel, contributed to the development and publication of the article with critical reviews.

Contents

1 Introduction

Every day, when engaging in the practice of teaching, teachers have to make meaningful pedagogical decisions; whether it is about lesson planning, adapting the plan while teaching in class, or while reflecting on one's own teaching and students' learning outcomes (Carlson et al., 2019; Furtak et al., 2016; Shavelson et al., 2008a). All these pedagogical decisions are contingent on the teachers' information about their students' learning prerequisites, processes, and outcomes (Beck et al., 2008; Ingenkamp & Lissmann, 2008; Schrader, 2017; Vogt & Rogalla, 2009). Therefore, teachers' competencies to diagnose students' skills, abilities, or learning outcomes are essential for mastering the professional practice of teaching. In the discussion about teacher competencies, this is referred to as teachers' diagnostic competencies (Herppich et al., 2018; Schrader, 2011; Südkamp & Praetorius, 2017).

The mission of the educational researcher is to enhance the professionalization of the teaching profession to better prepare teachers for their future everyday practice (Baumert & Kunter, 2006; Cochran-Smith & Zeichner, 2005; Darling-Hammond & Bransford, 2007). The description of teachers' specific professional competencies is part of the discussion regarding the search for standards and guidelines in higher education that guarantee the quality of teacher training programs (Blömeke et al., 2013; Oser & Renold, 2005). The core question is what competencies and skills teachers need to fulfill their main task (Baumert & Kunter, 2006), that is, providing opportunities for student learning by effective teaching (Helmke, 2007; Seidel & Krapp, 2014). To enable the effectiveness of teacher educators at university, they should be guided by the models of teachers' professional competencies, knowledge, and skills derived from educational research in their teacher training programs (Baumert & Kunter, 2006; Cochran-Smith & Zeichner, 2005; Darling-Hammond & Bransford, 2007). In the ongoing debate regarding ways of achieving this successfully, two aspects are repeatedly emphasized: teacher training programs should be (i) evidence-based and (ii) practice-oriented (J. Bauer & Prenzel, 2012).

Evidence-based refers to teaching the best available knowledge of a domain in the best possible way according to findings of educational research (Nutley et al., 2003). The Refined Consensus Model of Pedagogical Content Knowledge in Science Education

(Carlson et al., 2019) describes this knowledge in the sector of teachers' pedagogical content knowledge as the so-called "collective knowledge facet". It refers to the currently available knowledge held by the community of educators and teachers of a domain (Carlson et al., 2019). The idea of evidence-based teaching thus implies that this collective knowledge should be integrated into teacher training programs. Teacher training programs should further be **practice-oriented**, which means that the collective knowledge addressed in teaching should be linked to professional practice (Grossman & McDonald, 2008; McDonald et al., 2013). Increasing experiences with practical situations requiring the execution of professional practices, or at least their observation and reflection, is a way to transform conceptual knowledge into practice-oriented enacted knowledge usable in teaching practice (Carlson et al., 2019; Gruber et al., 2000). This can prevent (preservice) teachers from having conceptual knowledge, but without being able to apply it in practical situations. In other domains, such as medical education, great attention has already been paid to practice orientation in training programs (Schmidt & Boshuizen, 1993). However, in the domain of teacher education, closer links to practice can, and should, be established.

Teacher training programs include various methods to realize a close relationship to real-world situations to facilitate preservice teachers in acquiring enacted knowledge and practice-oriented skills. These methods range from case vignettes to the use of video clips or internships in schools (Blomberg et al., 2013; Friesen et al., 2018; Seidel et al., 2011; Stürmer, 2011). Nevertheless, educational research shows that many novice teachers struggle when entering the teaching profession (Correa et al., 2015; Dicke et al., 2015; Stokking et al., 2003). This research also shows that novice teachers rarely rely on knowledge acquired in higher education when it comes to challenging practice situations (Fraser et al., 2019). Thus, new approaches are required that combine evidence-based teaching with practice orientation.

Regarding the acquisition of diagnostic competencies, the research group COSIMA developed a conceptual framework (see Figure 1) from which a research agenda can be derived regarding how to foster these competencies in higher education (Heitzmann et al., 2019). The framework guides the research group in developing simulation-based learning environments as representations of practice to further understand diagnostic competencies and processes in practical situations, and to develop suitable instructional support for the simulations to foster the acquisition of the competencies (Chernikova , Heitzmann, Fink, et al., 2020). The research group's conceptualization of diagnostic competencies is based on the prior work of Blömeke et al. (2015) regarding the concept of professional competencies (see Figure 1: right box). In the acquisition of competencies, the processes and activities carried out in the learning environments are crucial for learning success. Thus, the framework emphasizes the importance of diagnostic activities for the learning process in simulation-based learning environments on diagnostic competencies (see Figure 1: middle box). The learning process can be further enhanced by additional instructional support (see Figure 1: left box). Research suggests various ways of support among which "scaffolding" is the one the research group focuses on (Chernikova, Heitzmann, Fink, et al., 2020). In simulation-based environments, scaffolds can be implemented in a standardized way at different points of the diagnostic process. The effects on the acquisition of diagnostic competencies, which arise from the processes in a simulation-based learning environment, are influenced, on the one hand, by the prerequisites of the learners and, on the other hand, by contextual factors of the simulation itself (see Figure 1: upper and lower boxes). Both factors, the context and the individual learning prerequisites, have to be considered carefully when developing a simulation-based learning environment.

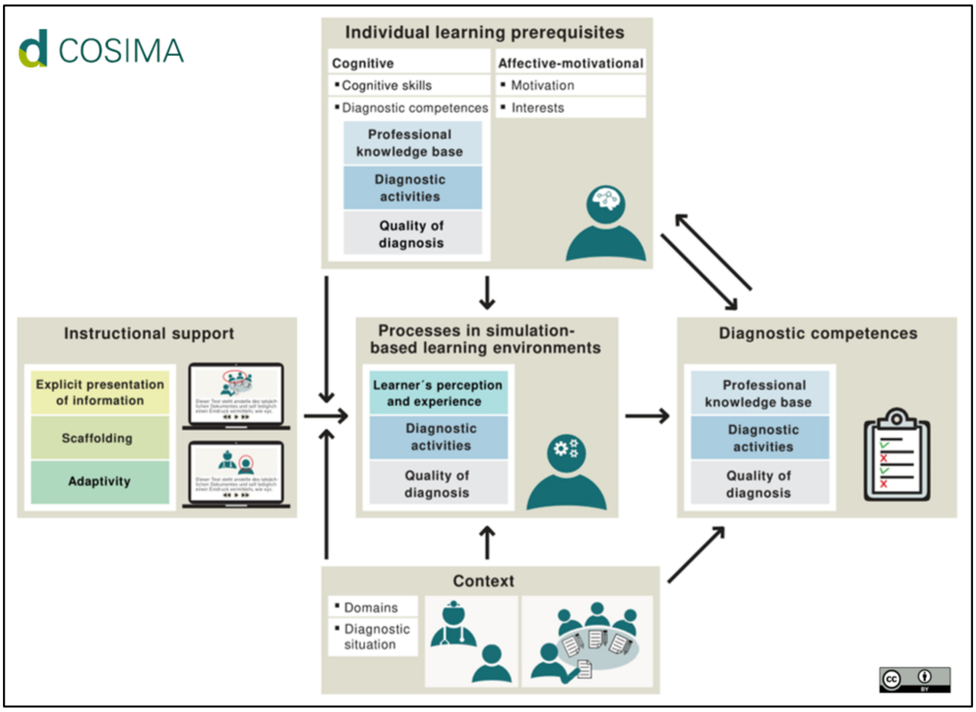

Figure 1: Framework of the research group COSIMA.[1]

[1] COSIMA framework model (www.for2385.uni-muenchen.de/aktuelles/rahmenmodellccby/cosima-frame-model_eng.pdf), COSIMA research unit, CC BY 4.0

This dissertation responds to the call for research on evidence-based and practice-oriented teaching in higher education. In the context of the dissertation, a video-based simulation for mathematics preservice teachers was developed. The video-based simulation is representative of a typical on-the-fly formative diagnostic situation in mathematics classrooms regarding the assessment of students' proving skills in geometry. In the development of this video-based simulation, different individual prerequisites of preservice teachers (see upper box in Figure 1), as well as specific context factors of a video-based simulation (see lower box in Figure 1), were taken into account. By the development and exploration of the video-based simulation, the dissertation achieved the following: first, it contributes new knowledge to the development of evidence-based representations of practice for higher education programs; second, it provides insights into preservice teachers' diagnostic processes and investigates diagnostic activities.

As a foundation for this dissertation, a short summarizing review of prior research is provided. First, the theoretical concept of diagnostic competencies and skills, which is the basis of the simulation, will be discussed. The conceptualization is derived from known frameworks and concretized for the operationalization in the simulation. Furthermore, the aspects in which this dissertation can contribute to research gaps in the conceptual understanding of diagnostic competencies and skills will be discussed. Subsequently, the design principles considered in the development of the simulation and their implementation are presented. Guided by Grossman et al.'s (2009) framework, the aim is to ensure that this representation of practice is suitable for the acquisition of diagnostic skills. Finally, considerations regarding the selection and construction of the simulated situation are presented. It will be shown why diagnosing students' geometrical proving skills is a relevant situation from the point of view of practical relevance, but also a suitable situation for a simulation from the point of view of teacher training.

2 Theoretical Background

2.1 Diagnostic Skills: A Professional Requirement

Teachers' diagnostic skills and competencies are a key component for the teaching profession and contribute to successful teaching practice (Baumert & Kunter, 2006; Darling-Hammond & Bransford, 2007). They are often mentioned in the same breath with the goal of teaching adaptively to the students' prerequisites as they serve as a basis for adaptive teaching (Beck et al., 2008; Brühwiler, 2017; Vogt & Rogalla, 2009). Due to this, among other reasons, diagnostic competencies take on an important role in many competence frameworks for teachers (European Commission, 2013; Kultusministerkonferenz, 2004). The concept of "diagnosing" is associated with medical contexts, but it is also common in the field of teacher education. Unfortunately, a variety of terms such as "teacher assessment" (e.g., Herppich et al., 2018) and "teacher judgment" (e.g., Urhahne & Wijnia, 2021) are used analogously in the field of teacher education. In this dissertation, the term "diagnosing" is used to align with the wording in the introduced framework of the research group COSIMA (see Figure 1). This wording emphasizes the similarities between diagnosing in the medical and teaching contexts, although other terms were used in the articles to adapt to the common wording.

2.1.1 Characteristics of diagnostic situations

Schrader (2011) connected the concept of diagnosing with the ability to successfully cope with different diagnostic situations in the teaching profession. Based on this point of view, a close look at diagnostic situations is pivotal for a detailed conceptualization of teachers' diagnostic competencies and skills. In their chapter on structuring diagnostic situations, Karst et al. (2017) described four main characteristics that can be used to classify sets of diagnostic situations: (i) the purpose of a diagnosis, (ii) the consequences of a diagnosis, (iii) the focus of a diagnosis, and (iv) the plannability of a situation (see Figure 2).

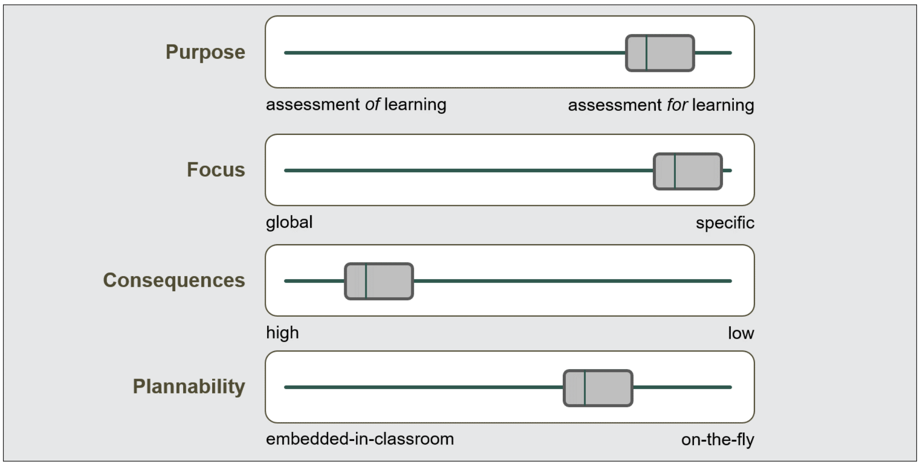

Figure 2. Characteristics to classify diagnostic situations in the teaching profession (adapted from Karst et al., 2017).

First, the handling of diagnostic situations can vary significantly depending on the **purpose** the diagnosis is intended to serve. Karst et al. (2017) differentiate between assessment *for* learning and assessment *of* learning (c.f. Ingenkamp & Lissmann, 2008). Assessment for learning focuses on students' actual learning processes and serves the purpose of improving the students' learning formatively (Alonzo, 2011; Black & Wiliam, 2009; Furtak et al., 2016). Assessment of learning targets final learning outcomes and is used for issuing summative qualifications (Ingenkamp & Lissmann, 2008; Karst & Förster, 2017). Especially for the assessment for learning, diagnoses serve the purpose of adapting classroom teaching constantly to students' individual learning needs (Vogt & Rogalla, 2009). Second, diagnostic situations show a wide range regarding the **focus** of the diagnoses. A diagnosis can have a global focus (Huber & Seidel, 2018; Karst & Dickhäuser, 2016; Kosel et al., 2021), such as, for example, students' overall skills and abilities in science, or a specific focus on individual students' skills, such as the students' mathematical proving skills in geometry. Third, the **consequences** of a diagnosis are characterized for diagnostic situations (Behrmann & van Ophuysen, 2017). Differentiation can be made between whether a diagnosis directly informs a specific educational decision or action (high), or whether a diagnosis is used as an orientation but no educational decision or action results directly from it (low). Fourth, diagnostic situations differ in their **plannability** for the diagnostician. This ranges from diagnostic situations that teachers can prepare well in advance, such as a written exam to test the students' mathematical skills after completion of a topic (embedded-in-classroom), to diagnostic situations that spontaneously arise while teaching in classroom (on-the-fly) (Shavelson et al., 2008b).

The assessment *of* learning can serve as a basis for a variety of pedagogical decisions related to adaptive teaching (DeLuca et al., 2016; Vogt & Rogalla, 2009). For many short-term pedagogical decisions, the assessment of specific student skills is primarily interesting to suggest subsequent pedagogical actions in upcoming lessons. In their professional practice, teachers face a multitude of on-the-fly diagnostic situations that can inform the diagnosis of students' specific skills (Praetorius et al., 2012; Ruiz-Primo & Furtak, 2007). Prior research has shown that many novice teachers struggle with the complexity of this kind of on-the-fly diagnostic situation (Correa et al., 2015; Dicke et al., 2015; Levin et al., 2009). One goal of the research is to gain more precise information about the diagnostic process in such situations (Schrader, 2017). This includes in particular the difficult-to-measure internal diagnostic activities of noticing and processing existing information from which the diagnosis results (Altmann & Nückles, 2017). Research in this area could be very helpful for the further understanding of diagnostic activities and their relationship to performance (Behrmann & Glogger-Frey, 2017). Therefore, this dissertation conducts research in the area of on-the-fly formative diagnostic situations regarding the assessment of specific students' skills (geometrical proving skills) suggesting relatively specific pedagogical consequences.

2.1.2 Diagnostic competencies and skills

Currently, many different perspectives on diagnostic competencies exist in research on teachers' competencies along with different research interests, each emphasizing different aspects that are embraced by the term "diagnostic competencies" (Heitzmann et al., 2019; Herppich et al., 2018; Koeppen et al., 2008; Leuders & Loibl, 2020; Loibl et al., 2020). Looking at teachers' competencies in general, one commonly used conception of competencies is suggested by Blömeke et al. (2015), that is, which combines a behavioral and analytical perspective on professional competencies. This conceptualization understands competencies as "the latent cognitive and affective-motivational underpinning of domain-specific performance in varying situations" (Blömeke et al., 2013, p. 202).

The framework for diagnostic competencies for this dissertation is significantly influenced by the framework of Blömeke et al. (2015) on competencies and incorporates three aspects: dispositions, activities, and performance (see *Figure 3*). This perspective on competencies involves dispositions that underly a person's behavior that results in actual performance (Koeppen et al., 2008). These dispositions, including professional knowledge base and performance (the diagnosis), are connected by diagnostic activities such as noticing and knowledge-based reasoning of relevant information in the diagnostic situation (Goodwin, 1994; Seidel & Stürmer, 2014; van Es & Sherin, 2002). The diagnostic activities mediate the relationship between dispositions and performance (G. Kaiser et al., 2017; Koeppen et al., 2008). Blömeke et al. (2015) thus describe competencies as a continuum between dispositions and performance.

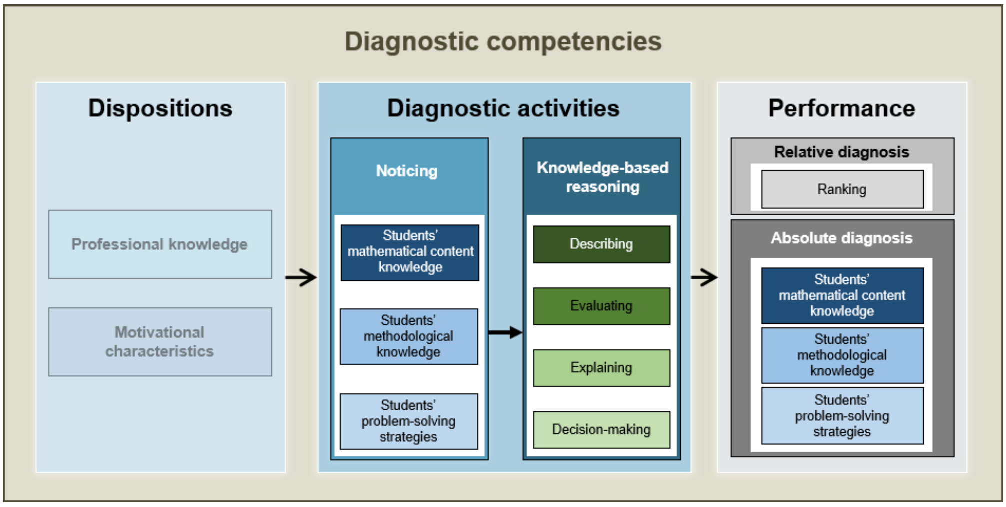

Figure 3. The framework for diagnostic competencies and the operationalization of diagnostic activities and performance for this dissertation.

Competencies are generally understood to have a certain consistency and stability (Funder, 2012; Ufer & Neumann, 2018). It therefore can be assumed that a teacher with specific diagnostic competencies can successfully cope with a specific set of diagnostic situations (consistency) (Schrader, 2013). This should be possible over a certain period of time (stability). Thus, the specific diagnostic competencies of a person can only be measured validly in specific diagnostic situations and over a given time. With the measurement of a persons' behavior, including the diagnostic activities and performance in a single diagnostic situation at a certain point of time, we should rather speak of a person's current diagnostic skills. The improvement of diagnostic skills in particular diagnostic situations can serve as a basis for the development of the person's diagnostic competencies. In this dissertation, the preservice teachers' diagnostic skills with their performance and diagnostic activities were examined in one diagnostic situation.

2.1.3 Diagnostic activities and performance

Research on teachers' diagnostic skills has generated a substantial number of studies on the accuracy of teacher diagnoses (Helmke & Schrader, 1987; Hoge & Coladarci, 1989; McElvany et al., 2009; Urhahne & Wijnia, 2021). The accuracy can be used to operationalize and determine (preservice) teachers' performance in a diagnostic situation. Artelt and Gräsel (2009) described accuracy as the ability to adequately assess student characteristics compared to either expert ratings regarding these characteristics, results of specific diagnostic tests, or students' performance measures regarding a specific topic. Research on teachers' accuracy is conducted for a wide variety of students' characteristics such as, for example, academic achievement (Südkamp et al., 2012; Thiede et al., 2015), motivation (Praetorius et al., 2017), or student profiles regarding cognitive and motivational-affective characteristics (Kosel et al., 2021; Seidel, 2007). In this line of research, a distinction can be made between the relative and absolute accuracy of a diagnosis (Leuders et al., 2018; Urhahne & Wijnia, 2021) (see Figure 4).

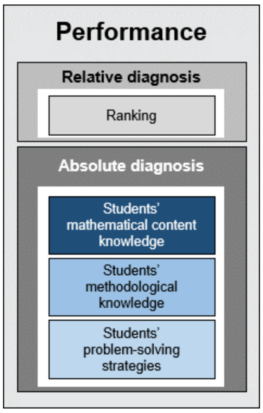

Figure 4. The preservice teachers' performance in diagnostic situations measured by the relative and absolute accuracy of the diagnosis.

Relative accuracy refers to situations wherein teachers must distinguish students based on multiple relative levels of their characteristics or skills. This is also called the rank component of the accuracy (Helmke & Schrader, 1987; McElvany et al., 2009), with which teachers are asked to rank the students based on a comparison of their characteristics or skills. This requires that students with different skill levels can be sequenced. Another way to estimate the accuracy of a diagnosis is the **absolute accuracy**. Teachers must diagnose the actual level of students' characteristics or skills (Hosenfeld et al., 2002; Thiede et al., 2015). Based on these ratings, the level component of the accuracy is measured (e.g., Südkamp et al., 2008). In doing so, the teacher's diagnosis can be compared to actual student skills. In simulated situations represented by scripted classroom videos, the students' skill levels and how they demonstrate them can be planned precisely. The actual level of student skills can then be determined by a validation of the video clips.

During a diagnostic process of an on-the-fly situation, teachers must pay attention to relevant information and apply their knowledge to reason out the noticed information resulting in a concluding diagnosis (Herppich et al., 2018; Schrader, 2017). Current research is engaged in further understanding these diagnostic processes and the occurring diagnostic activities (Altmann & Nückles, 2017; Kramer et al., 2021; Sherin et al., 2011). Professional vision frameworks are commonly used in research on teacher education to describe such activities (Goodwin, 1994; Seidel & Stürmer, 2014; van Es & Sherin, 2002). Noticing and knowledge-based reasoning can be seen as important diagnostic activities that occur in the diagnostic process. The diagnostic activity of noticing refers to paying attention to relevant information in the classroom (Schack et al., 2017; van Es & Sherin, 2008). With the abundance of information in on-the-fly situations, not every event can be noticed, so the challenge is to focus on the most relevant information.

Noticing is the foundation of knowledge-based reasoning (Seidel & Stürmer, 2014). Reasoning is the act of interpreting the noticed information based on one's personal knowledge (Gegenfurtner et al., 2020). For a subdivision of knowledge-based reasoning into different diagnostic activities, various prior works can be considered. Seidel and Stürmer (2014) included the activities of describing, explaining, and predicting. Fischer et al.'s (2014) framework for activities in the context of scientific reasoning and argumentation was not developed with direct reference to diagnostic situations. Nevertheless, diagnosing can be interpreted as a specific kind of scientific reasoning and multiple similarities can be found in the works of Seidel and Stürmer (2014) and Fischer et al. (2014) (c.f. Kramer et al., 2021).

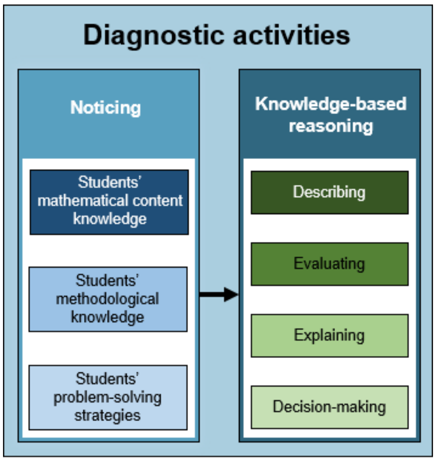

Figure 5. The operationalization of the preservice teachers' diagnostic process with the therein diagnostic activities.

A framework for the description of knowledge-based reasoning activities was developed in collaboration with the authors from Kramer et al. (2021). It is mainly based on the works of Seidel and Stürmer (2014) and Fischer et al. (2014) and adapted for diagnostic situations in classrooms. Here, knowledge-based reasoning is divided into four diagnostic activities: describing, evaluating, explaining, and decision-making (see Figure 5). Describing refers to verbalizing noticed information without making any further judgments. Evaluating refers to generalizing a student's pattern of behavior or level of knowledge. Explaining refers to linking these considerations to professional concepts and knowledge and decision-making refers to the activity of drawing conclusions based on the other activities. By operationalizing these diagnostic activities, the preservice teachers' diagnostic process in a diagnostic situation can be described.

2.2 Representations of Practice

A main challenge in teacher education is to provide learning environments for preservice teachers that support the development of complex skills, such as diagnostic skills (Chernikova, Heitzmann, Stadler, et al., 2020; McDonald et al., 2013). Researchers have called for practice-oriented environments in teacher learning where knowledge application in authentic situations is possible (Grossman & McDonald, 2008; Kolodner, 1992). Since a close relation with real-world professional situations is required, internships in schools have traditionally served as the main learning environments with authentic situations in common teacher training programs (Bennack & Jürgens, 2002). However, it appears that real classroom situations are overwhelming and too demanding for preservice teachers even after their university studies (Dicke et al., 2015; Stokking et al., 2003). A particular challenge is the high density of interactions and information in classrooms (Grossman & McDonald, 2008). This can hinder preservice teachers in performing diagnostic activities, such as the noticing of relevant classroom situations (Schoenfeld, 2011; van Es, 2011). Consequently, this can avoid knowledge-based reasoning about these relevant classroom situations and thus the linking of knowledge to practice examples (Alles et al., 2019; Förtsch et al., 2018). Therefore, researchers in teacher education must find new options to provide learning environments in which preservice teachers can practice professional skills.

2.2.1 Design principles

Grossman et al. (2009) developed a framework to design representations of practice for university teaching that highlights two principles that need to be balanced: authenticity regarding real-world practice (approximation of practice) and adequate complexity regarding the required cognitive demand (decomposition of practice). The principle of **approximation of practice** emphasizes the importance of a close relationship between the learning environment and real-world situations and suggests that simulated situations should imitate real-world situations as closely as possible. This implies not only an authentic representation of the situation, but, in particular, it proposes that typical activities of practitioners, which are performed in real-world situations, should also be accessible to learners in simulated situations (Hartmann, 1996). The approximation of practice suggests that a trained behavior in a simulated situation can be transferred to real-world situations. Indicators for successful approximations of practice are, for example, the perceived authenticity of both the target group and an expert group for such situations (Seidel et al., 2010). Furthermore, it is crucial that learners feel sufficiently present to become involved in the simulated situation (Schubert et al., 2001). In order to act comparably to a real-world situation and to perform similar activities in the simulated situation, preservice teachers need to be cognitively involved (Dankbaar et al., 2016) and motivated (Witmer & Singer, 1998).

The principle of approximation of practice is contrasted by the ***decomposition of practice*** referring to the difficulty of a simulated situation. The idea is to deconstruct certain facets of the practice situation such that irrelevant aspects that potentially hinder the learning process are reduced. The goal is to adjust the cognitive demand of the simulated situation in such a way that learners are challenged in a proximal zone of development but not overburdened (Vygotsky, 1978). This can be done, for example, by breaking down practice and highlighting relevant parts for the purpose of learning (Grossman & McDonald, 2008).

2.2.2 The use of staged videos

Designing a representation of practice according to Grossman et al.'s (2009) postulated principles is increasingly implemented in teacher education based on the use of video clips (Gaudin & Chaliès, 2015). Nowadays, a common approach in research to capture and present classroom situations is the use of video clips (e.g., Kramer et al., 2020). Videos can capture the authenticity of classroom practice (Brophy, 2004), and they can allow for careful observations with noticing and knowledge-based reasoning activities about teaching and learning situations. Videos have some advantages compared to common methods such as texts or case vignettes in realizing the principle of the approximation of practice (Friesen et al., 2018; van Es et al., 2020). Previous research has repeatedly shown the success of using videos in teacher education for different domains and purposes (Barnhart & van Es, 2015; Derry et al., 2014; Lindmeier, 2013; van Es et al., 2017). In particular, research has shown that preservice teachers perceived video clips as authentic approximations of practice (Seidel et al., 2010). However, videos filmed in classrooms can often contain a large amount of information about the situation. Although a video clip shows only one particular perspective and captures only a piece of the classroom actions, the decomposition of practice is often not sufficiently implemented. Even though video clips can capture a great deal of information about the situation, relevant information can sometimes be missing, which can impair the sense-making of the situation. One way to control a sufficient decomposition of practice in videos is the so-called staged video clip, which shows scripted simulated situations. For staged videos, a scenario is selected based on evidence-driven considerations, the content is carefully planned and scripted, and finally videotaped with the help of actors (Dieker et al., 2009). Following the idea of decomposing a situation, simulated situations can have a reduced and adapted complexity to the preservice teachers' skills. They allow preservice teachers to focus on relevant events and blend out other possibly distracting factors such as background noise or peripheral actions. This type of video clip compromises in the approximation of practice, but previous research has shown that preservice teachers also perceive staged videos as authentic approximations of classroom situations (Böttcher & Thiel, 2018; Piwowar et al., 2017).

2.2.3 Simulations

Another approach, besides the use of staged videos, to realize representations of practice in higher education is the development of simulations (Heitzmann et al., 2019). A simulation is a model of a real-world situation that imitates this situation (Frasson & Blanchard, 2012; Shannon, 1975; Wissenschaftsrat, 2014). Participants should have the option of influencing the course of action in a simulation by making decisions (Fischer & Opitz, in press). This makes it possible for participants to get involved in the simulation and perform activities similar to those of practitioners in real-world situations. With this, not only can diagnostic situations be presented and observed (as with videos), but also diagnostic activity. For example, a simulation not only allows the gathering of data (as compared to video clips) but also allows for a decision to be made regarding further actions within the simulation. By using a simulation as a standardized model of a real-world diagnostic situation, particular diagnostic activities can be focused on; thus, further insights into the relationship between diagnostic activities and performance can be gained (de Jong, 2006). The influence on the course of action by the interaction with the simulation can positively affect the perceived presence of the simulation's participants (Slater & Wilbur, 1997; Vorderer et al., 2004). Moreover, prior research has shown that the opportunity to influence the course of action supports the learners' cognitive involvement (Paas et al., 2003).

Simulations are already frequently used in higher education in other domains such as medicine (Cook et al., 2011; Fink et al., 2021; Radkowitsch et al., 2020). Simulated virtual patients, for example, are widely used as a type of computer simulation in medical training programs (Villaume et al., 2006; Vu & Barrows, 1994). Whereas in teacher education at the university level, there are only a few simulations, such as the "simulated classroom," to assess diagnostic skills (Südkamp et al., 2008), these simulations in higher education suggest that the use of simulations can be successful (Hertel & Millis, 2002).

2.3 Diagnosing Geometrical Proving Skills

In the academic discussion about teaching mathematics, it repeatedly becomes apparent that acquiring mathematical argumentation and proving skills is a challenging issue for students and teachers (Harel & Sowder, 1998; Healy & Hoyles, 2000; Stylianides, 2007). Since mathematics is a proving science and proving is an essential part of mathematic activities (Boero, 2007), argumentation and proof are also part of mathematics classrooms. The National Council of Teachers of Mathematics (NCTM) highlights the importance of argumentation skills for mathematics classrooms as one of the ten standards in the NCTM Principles and Standards (NCTM, 2000). Similarly, the Standing Conference of the Ministers of Education and Cultural Affairs of Germany identified mathematical argumentation skills as one of the six central mathematical skills to be taught (Kultusministerkonferenz, 2012). Mathematical argumentation and proving skills are closely related since a mathematical proof can be interpreted as a specific kind of mathematical argument. Compared to the broad term of mathematical argumentation, mathematical proving is seen as a more formal form of argumentation that follows specific norms of the mathematical community to show the validity of an assumption (Stylianides, 2007). For example, only deductive conclusions are accepted in a mathematical proof (Heinze & Reiss, 2003; Pedemonte, 2007).

2.3.1 Students' geometrical proving skills

In school, students usually focus on proof and proving for the first time in the context of plane Euclidean geometry (ISB, 2021; Stylianides, 2007). In geometry, concepts and terms are initially introduced through visualizations and based on examples. Many concepts are already established in primary school and generalized and further formalized in secondary school. Deductive thinking and reasoning regarding such concepts often cause students problems and difficulties (Heinze & Kwak, 2002; Reiss et al., 2002). Research has repeatedly shown that students have substantial problems with handling mathematical proofs (Harel & Sowder, 1998; Healy & Hoyles, 2000) because it depends on different student prerequisites (Reiss & Ufer, 2009; Sommerhoff, 2017; Ufer et al., 2008). Deficits in these prerequisites can cause problems in the handling of mathematical proofs. Although several prerequisites have been studied, including the students' beliefs about the function of proofs

(Sommerhoff, 2017; Sommerhoff et al., 2020), it has been shown that the following three can be considered among the most important: the students' mathematical content knowledge, methodological knowledge, and problem-solving strategies.

First, **mathematical content knowledge** in the area of the proof task is indispensable. A students' mathematical content knowledge refers to knowledge of definitions and propositions from the field of mathematics in which the proof has to be constructed (Chinnappan et al., 2011; Sommerhoff, 2017; Ufer et al., 2008; Weigand et al., 2014). For example, the geometrical proof task shown in Figure 6 requires mathematical content knowledge regarding congruence theorems and theorems about angles at line intersections with parallel lines.

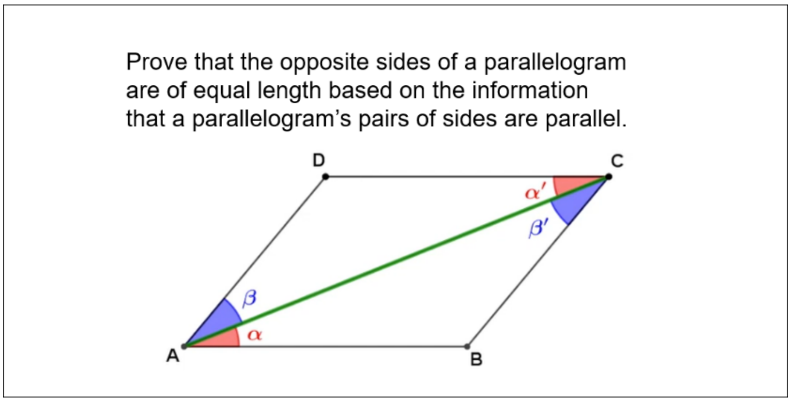

Figure 6. The proof task from the field of plane Euclidean geometry that is used in the video-based simulation.

Second, knowledge about norms that define an acceptable proof in the respective community (**methodological knowledge**) is necessary for students to handle proof tasks (Heinze & Reiss, 2003; Sommerhoff & Ufer, 2019). It is not only crucial to know the kind of arguments that can be used but also the proof structure. A proof starts at a given premise and ends at a specific assertion without gaps in the structure of argumentation in between (Heinze & Reiss, 2003). Research regarding the proof structure has shown that assessing whether a proof with its structure is correct or incorrect challenges students (Lin, 2005; Reiss et al., 2002). Research regarding the use of valid arguments in proofs showed that often student solutions to proof tasks are based on empirical arguments and lack exposure to deductive reasoning (Healy & Hoyles, 2000). An example of an invalid argument for the proof task above (see Figure 6) is that the angles α and α' are assumed to be equal because they were measured in this particular parallelogram.

Third, based on the definition of problems (Reiss & Törner, 2007; Schoenfeld, 1985), mathematical proof tasks can be seen as specific types of problems. Accordingly, **problem-solving strategies** are an important prerequisite for handling mathematical proofs (Ufer et al., 2008). This refers to strategies such as decomposing a task to cope with smaller parts of the task (Pólya, 1973) and to meta-cognitive abilities to control and monitor the problem-solving process (Schoenfeld, 1992). Studies show, for example, that inadequate strategies, such as a lack of exploration of the conjecture, can have negative effects on the construction of a proof (Lin, 2005).

2.3.2 Diagnosing students' geometrical proving skills

In diagnosing students' geometrical proving skills, the above-mentioned prerequisites play an important role. If preservice teachers can diagnose the students' availability of these prerequisites, they get important information on central aspects of students' overall geometrical proving skills, which helps them to create a general diagnosis. To diagnose different levels of students' geometrical proving skills the van Hiele model of geometrical thinking can be useful (Usiskin, 1982). This model describes the geometrical thinking of students with a broad and global perspective. The model and further updates of the model (Kunimune et al., 2009) essentially distinguish four different levels of geometrical thinking in secondary school. Students start with a level on which concepts are formed visually and visual arguments are used in their mathematical argumentations and proofs, up to a high level, where generalized and valid arguments can be understood and used (Kunimune et al., 2009; Usiskin, 1982). These levels can also be applied in the context of geometrical proving. The van Hiele model can be used to diagnose students who show relative consistent cues for all three prerequisites. On the one hand, the model helps to create a global diagnosis of the students' geometrical proving skills, and on the other hand, it supports the rank component of a diagnosis by clustering the wide range of students with different proving skills occurring in school into four levels.

3 Present Research

In the background of the importance of diagnostic competencies in professional teacher training programs in higher education, this dissertation contributes to the lack of research on diagnostic skills in on-the-fly formative diagnostic situations. Currently, different frameworks related to diagnostic skills exist, but the understanding of activities involved in the diagnostic process and their relationship to the resulting diagnoses remains vague. Often, studies on diagnostic skills suffer from a lack of contextualized, authentic environments in which these skills can be examined with a close connection to real-world situations. The overarching aim of this dissertation is to contribute to the understanding of (preservice) teachers' diagnostic skills. Based on the call for research in representations of practice with a close real-world relation, a video-based simulation was developed as a standardized representation of a typical diagnostic situation in mathematics classrooms. The video-based simulation and its development process can be used as a best-practice example to illustrate how representations of practice can be designed to create authentic environments in which professional skills can be investigated in a standardized and valid way. The dissertation sheds further light on preservice teachers' diagnostic activities when handling on-the-fly formative diagnostic situations. The findings of the two studies presented in the two articles included in this dissertation contribute to creating suitable learning environments for professional competencies and skills.

Article A. The first article provides insights into the validity of the newly developed video-based simulation. The simulation was designed for use in teacher training programs at the university level to investigate preservice teachers' diagnostic skills throughout their higher education. The article addresses the question of whether the video-based simulation represents, on the one hand, an approximation of practice and, on the other hand, an adequate decomposition of practice for research on preservice teachers' diagnostic skills. Therefore, the balance of the authenticity of the simulation and its cognitive demand for preservice teachers were examined. The first goal of the study was to analyze whether preservice teachers perceive the video-based simulation as an authentic approximation of practice (RQ 1). To answer the research question, teachers' perceptions of the authenticity of the simulated diagnostic situation, in particular regarding the video clips and their embedding in the simulation, were surveyed. Moreover, participants' cognitive involvement and motivation were assessed as further indicators for an authentic approximation of practice. The second research question discussed in this article addresses the cognitive demand of the video-based simulation. For this purpose, we investigated the question of whether the decomposition of the diagnostic situation resulted in a diagnostic task with an adequate level of difficulty for preservice teachers (RQ 2). Both relative accuracy and absolute accuracy were analyzed. To inform the answer regarding the relative accuracy, the preservice teachers' performance on ranking the simulated students according to their geometrical proving skills was examined. For analyzing the absolute accuracy, the preservice teachers' detailed diagnoses of the three prerequisites for students' geometrical proving skills served as a basis.

Article B. The second article was situated within the context of research on the diagnostic process and the diagnostic activities therein. The main goal of the study was to provide insights into the preservice teachers' diagnostic process when handling an on-the-fly formative diagnostic task. Therefore, the article focuses on the diagnostic activity of noticing relevant information, and on participants' knowledge-based reasoning regarding the noticed information. As a first step, the diagnostic process was described by examining the preservice teachers' observations based on whether they focused on professionally relevant information (RQ 1). The relevant information in the diagnostic situation of the video-based simulation is represented by the simulated students' statements and actions, indicating their levels of prerequisites for handling geometrical proof tasks. Therefore, the extent to which preservice teachers noticed information regarding the prerequisites (mathematical content knowledge, methodological knowledge, and problem-solving strategies) was investigated. The second research question dealt with the diagnostic activity of knowledge-based reasoning. The article explores the question regarding the extent to which preservice teachers show different knowledge-based reasoning activities in the process of diagnosing students' geometrical proving skills (RQ 2). To answer this research question, the frequency of the activities describing, evaluating, explaining, and making decisions was analyzed. Of particular interest was the frequency of preservice teachers' activities where they use their knowledge, such as in evaluation, explanations, and decisions. For both questions, we also investigated whether there were differences in the distribution of diagnostic activities when diagnosing students with different levels of geometrical proving skills.

4 Methodology

4.1 Project Context

The present dissertation is embedded in the research project, Visit-Math (SE 1397/11-1), which is part of the research group COSIMA (Grant No. 2385)[2]. The research group aims to foster diagnostic competencies in simulation-based learning environments in higher education. This involves the analysis of the design and use of simulation-based learning environments in university, as well as the investigation of instructional support mechanisms for diagnostic competencies and skills in an early phase of competence acquisition. The research group was structured so that the topic of diagnostic competencies was considered in different domains (c.f. Figure 1). Researchers from the field of teacher education collaborate with researchers from the medical field in the research group COSIMA. Overall, researchers from medical education, mathematics education, biology education, and educational psychology are involved. Divided into seven research projects, they work on the objectives of the research group in different diagnostic situations (see Figure 7).

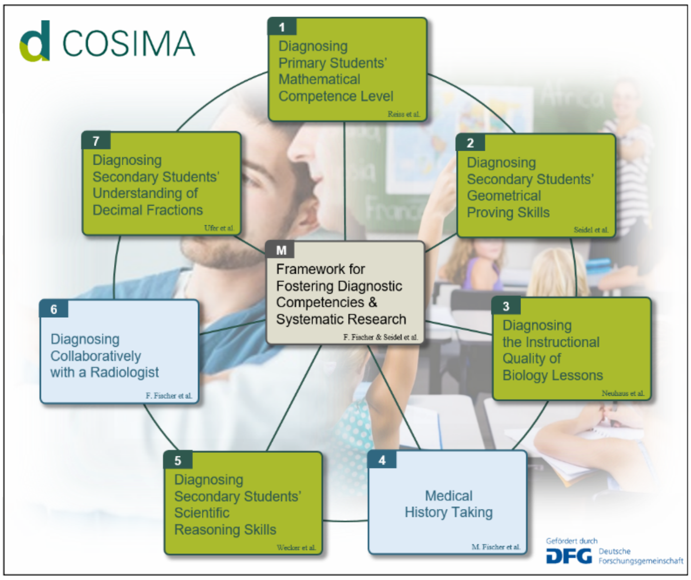

Figure 7. Structure of the research group COSIMA with its research projects conducting prior studies (1-7) in the teacher educational (green) and medical field (blue) and the overarching meta-analytical project M.

[2] For more details: www.for2385.uni-muenchen.de

The Visit-Math project is one of the COSIMA research projects (Project 2, see Figure 7) from the teacher education field with the goal of investigating and promoting diagnostic skills of mathematics preservice teachers for secondary school. It aims to simulate a typical on-the-fly diagnostic situation during mathematics class using staged video clips and embedding them in a mutable digital environment. Preservice teachers should use these video clips to diagnose the students' geometrical proving skills. The developed video-based simulation should, on the one hand, be capable of measuring the preservice teachers' diagnostic skills throughout their university education phase (Visit-Math Assess). Specifically, the simulation should provide insights into the rarely captured diagnostic process of preservice teachers in an on-the-fly diagnostic situation. The video-based simulation should in principle allow the integration of different scaffolds, such as different types of prompts (Visit-Math Learn).

This dissertation with its two studies was influenced by the guidelines and specifications of the Visit-Math project as well as by the structure of the research group COSIMA. For example, findings from the medical education field largely inspired the design of the video-based simulation. The coding scheme for categorizing different diagnostic activities in the preservice teachers' diagnostic process, which was used in the second study, was developed in strong cooperation with other school-based research projects of the research group. All studies presented in this dissertation were designed in such a way that data is suitable for the integration in future, research group-wide meta-analyses (performed by an additional project called Project M; see Figure 7). To enable comparability between the different research projects, there was a special focus on finding and creating similarities among the studies during their development.

4.2 Design of the Video-based Simulation

The video-based simulation Visit-Math Assess is a digital environment that was programmed and designed based on the software Unipark[3]. The simulation is a representation of an on-the-fly formative diagnostic situation with one-on-one teacher-

[3] EFS Survey, Version Fall 2020.
Questback GmbH, Cologne.
www.unipark.com

student interactions. In the simulated situation, four simulated seventh-graders work on a geometry proof task; they must prove that opposite sides of a parallelogram are of equal length based on the information that a parallelogram's pairs of sides are parallel. The topic of the situation was chosen based on a literature review of students' mathematical skills and their assessment in school. Resulting from this review, geometrical proving skills were identified as highly relevant and applicable content for the diagnostic situation.

To represent the diagnostic situation in a video-based environment, staged video clips were produced showing a simulated teacher and one of four simulated students. In the video clips, a simulated teacher and the simulated students discuss the students' progress and argumentation in the context of the geometry proof task. Before the actual conversations were filmed with students as actors, a contextual frame and detailed scripts for the staged video clips were developed as part of this dissertation. Each video clip was constructed to contain explicit cues about the students' levels of prerequisites for their geometrical proving skills. Therefore, cues for the different prerequisites were included in the videos based on the spoken and written words.

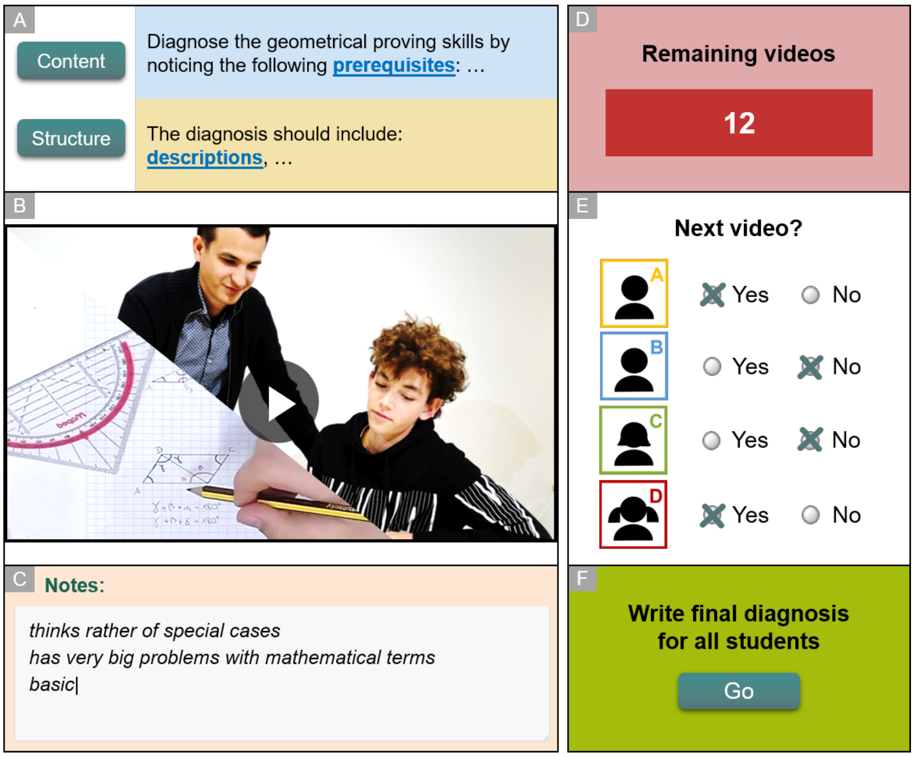

Figure 8. Illustration of the structure of the video-based simulation Visit-Math Assess that indicates different steps during working in the simulation (adapted from Codreanu et al., 2021).

At the beginning of the video-based simulation, the participants were familiarized with the diagnostic situation. They were introduced to their role of a preservice intern accompanying the simulated teacher in his class, a situation familiar and typical for students enrolled in a teacher training program. Figure 8 illustrates the subsequent parts of the simulation, Visit-Math Assess, until participants finished their final diagnoses. After familiarization, participants received information on students' prerequisites of geometrical proving skills on which they were asked to focus, as well as on the components that a diagnosis should ideally contain (see Figure 8, area A). During the actual diagnostic process, participants could work independently in the simulated classroom and gather information about the simulated students by watching the video clips and taking notes (see Figure 8, areas B and C). After an initial round of watching video clips of all simulated students, participants were able to choose to either further observe a student or to decide that they had gathered sufficient information for the diagnosis of this particular simulated student (see Figure 8, area E). In total, a maximum of 20 video clips could be observed (see Figure 8, area D), and participants were free to decide on how to distribute these observations among the simulated students. Afterward, participants were asked to use their notes to create a written diagnosis of each simulated student (see Figure 8, area F). Then, participants ranked the simulated students according to their overall mathematical argumentation skills from weakest to strongest. Subsequently, participants rated each student's level of prerequisites using 4-point Likert response items from "mostly not available" to "available." This describes the structure of the core of the video-based simulation. In the two studies, before and after the simulated diagnostic situation, further instruments were integrated into the digital environment to survey variables such as the participants' professional knowledge or their interest.

4.3 Measures and Instruments

Both articles analyzed data collected with the video-based simulation, Visit-Math Assess. The emphasis in Article A was on the approximation and cognitive demand of the simulation. As an indicator for an adequate cognitive demand of the simulation for preservice teachers, participants' performances on the diagnostic task and the thereby estimated difficulty of the simulated diagnostic situation was used. In Article B, participants' diagnostic processes and the more specific diagnostic activities were investigated. For this purpose, the free text final written diagnoses were coded and analyzed.

4.3.1 Approximation of simulation and accuracy in performance

To measure the extent to which the video-based simulation was perceived as an authentic approximation of practice, preservice teachers and experts rated their perceptions on four different scales. One scale addressed the authenticity of the video clips themselves, and another focused on the representation of the diagnostic situation in the simulation. Additionally, two scales dealt with the cognitive involvement and motivation of the participants.

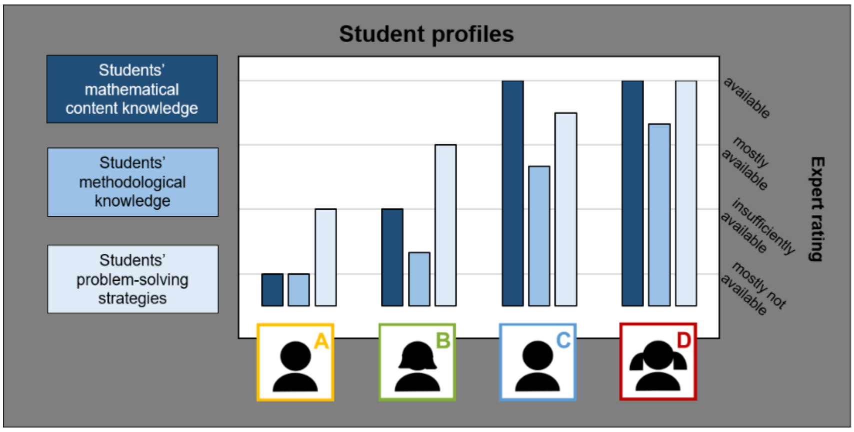

Figure 9. Student profiles designed for the video-based simulation Visit-Math Assess (adapted from Codreanu et al., 2021).

For measuring the preservice teachers' performances in Article A, the following two diagnostic tasks were analyzed: (i) ranking the simulated students according to their overall geometrical proving skills, and (ii) the preservice teachers' detailed diagnosis of each simulated student. The first task was used to determine the preservice teachers' relative accuracy (rank component). The simulated students were designed in a way that they could be ranked according to their overall geometrical proving skills. The simulated students exhibited coherent cues related to their geometrical proving skills in the video clips, which highlighted their different skill levels (see Figure 9). In the second task, the participants rated the specific prerequisites of the simulated students on eight Likert response items. This was used to determine the preservice teachers' absolute accuracy (level component). The simulated students could be diagnosed according to the three prerequisites for geometrical proving skills: mathematical content knowledge, methodological knowledge, and problem-solving strategies. Before the study, two researchers from the field of mathematics education independently rated the four simulated students regarding each of the above-mentioned prerequisites on the later-used 4-point Likert scale. Figure 9 shows the assessments of the two experts for the respective prerequisites, which were calculated from the specifications on the individual items of the scale. The absolute accuracy of the participants' ratings was based on the agreement with these expert ratings.

4.3.2 Diagnostic activities

To examine the preservice teachers' diagnostic processes in more detail in the context of Article B, the diagnostic process is described as consisting of two different types of diagnostic activities. First, the diagnostic process is described by examining the noticing based on whether preservice teachers focused on professionally relevant information. Second, the diagnostic activities of knowledge-based reasoning were analyzed. To detect these diagnostic activities, the free text diagnoses were divided into distinct coding units and then coded by two members of the research team. The specific codes for the diagnostic activities were identified using both deductive and inductive methods.

For the activity of noticing, we examined the content the preservice teachers mentioned within their free text diagnoses. Codes were assigned for the three different prerequisites: the students' mathematical content knowledge in geometry, the students' methodological knowledge about mathematical proofs, and the students' problem-solving strategies. For coding knowledge-based reasoning, four different codes were applied, following the professional vision literature (Seidel & Stürmer, 2014) and the literature for scientific reasoning and argumentation (Fischer et al., 2014): describing the concrete course of action, evaluating the students' behavior, explaining the students' level of knowledge, and making decisions about the need for individual learning support (c.f. Kramer et al., 2021).

4.4 Data Collection

In the two articles of this dissertation, I analyzed data from two different studies, which were collected in the context of the research project, Visit-Math. In both studies, the video-based simulation Visit-Math Assess was used. Data for the validation study (1st study) were collected in the summer semester of 2018 at the Technical University of Munich. The sample consisted of $N = 28$ preservice high school teachers. In advance, 10 external mathematics education experts (criteria: more than five years of secondary school experience and currently active in teacher education) provided formative feedback as part of an external validation of the video-based simulation. Data collection for the second study took place one year later in the summer semester of 2019 at the Ludwig-Maximilians-University of Munich. It was possible to collect data from $N = 51$ preservice teachers who worked on the video-based simulation for this study.

4.5 Data Analyses

The original data sets were provided by the software Unipark, with which the video-based simulation was programmed. The coding of the free text diagnoses was done with Microsoft Excel. R was used for data wrangling as well as for all statistical analyses (R Core Team, 2020). For both studies, the data were first analyzed using descriptive statistics. To show differences between groups or tasks, for example, diagnosing the four different simulated students, mostly repeated measures ANOVAs were calculated. The repeated measures ANOVAs were computed with contrasts and bonferroni-adjusted post-hoc analyses for multiple comparisons to analyze differences. The following significance levels were used throughout the data analyses: $^{***}p < .001$, $^{**}p < .010$, $^{*}p < .050$.

5 Summary of Publications

5.1 Article A: “Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education”

Article A addresses the actual design of the video-based simulation, Visit-Math Assess, and the underlying considerations in the developing process. It aimed to investigate the question of whether the video-based simulation is appropriate for use in university teacher education and thereby useable for measuring preservice teachers’ diagnostic skills. The main focus of this study was the perceived authenticity of the simulation and the preservice teachers’ cognitive demand. The principal considerations in the design of the video-based simulation were based on the framework of Grossman et al. (2009) regarding teaching practice in professional education. The three mentioned key concepts (representations, decomposition, and approximations of practice) for understanding the pedagogies of practice in professional education guided the process of developing the simulation and inspired the design of the study presented in Article A. Careful consideration about different forms of representations for a diagnostic situation resulted in the use of staged videos embedded in an interactive simulation.

The main goal of the study presented in Article A was to internally validate the developed video-based simulation by balancing authenticity and cognitive demand. The first research question asked how preservice teachers experienced the video-based simulation regarding an authentic approximation of practice. The second research question addressed the cognitive demand by asking: How do the diagnostic tasks in the video-based simulation indicate adequate levels of difficulty and distributions of student performances in mastering the diagnostic situation? The preservice teachers’ statements about the authenticity of the simulation and their performance in the simulation were included in the analysis. To assess the extent to which the video-based simulation was perceived as an authentic approximation of practice, participants rated their perceptions of authenticity on four different scales. Preservice teachers’ performances on two diagnostic tasks served as indicators for their cognitive demand. On the on hand, participants were asked to provide a correct ranking of the four represented students according to their overall mathematical argumentation skills. On the other hand, they were instructed to provide a detailed diagnosis of the specific skills and characteristics of mathematical argumentation skills as exhibited by the students in the video-based simulation. In line with the goal of the article, the preservice teachers confirmed the simulation's high degree of authenticity. This was also acknowledged by a group of external mathematics education experts in a previous pilot study. The preservice teachers' performance on the diagnostic tasks was, overall, at the medium level.

The results showed that video clips can capture the authenticity of classroom situations, matching prior research on the use of videos in teacher education. Furthermore, the possibility to influence the course of action in a simulation was perceived positively. These findings indicated that a valid approximation of practice was created. Although it can be concluded from the results that the first diagnostic task was quite easy to master, the actual task of writing a detailed diagnosis showed an adequate level of cognitive demand for preservice teachers and, thus, an adequate decomposition of the situation. The study represents an exemplary application of concepts of the framework for teaching practice in the design of a video-based simulation. Based on the findings of the study, the video-based simulation seems useful for teacher education programs as well as for measuring diagnostic skills. Thus, the video-based simulation provides a proper basis for the study in Article B.

The video-based simulation and the study design were developed by all authors. I supervised the programming, created the visualization, and collected data. I handled and analyzed the data, wrote the manuscript, and implemented the suggestions and remarks by co-authors and reviewers. I supervised the submission process and serve as a corresponding author for this publication.

5.2 Article B: "Exploring the process of preservice teachers' diagnostic activities in a video-based simulation"

Article B addresses the preservice teachers' diagnostic process with their context-specific skills in a diagnostic situation. It aimed to provide insights into preservice teachers' handling of situation-specific professional tasks such as on-the-fly student assessment. The main focus of this study was to examine the preservice teachers' noticing and knowledge-based reasoning in the diagnostic process. The study was guided by findings from research on teachers' professional vision (Goodwin, 1994; Seidel & Stürmer, 2014). First, the noticing of relevant information by examining important mathematical content mentioned by the preservice teachers was considered. Second, the subsequent knowledge-based reasoning, which is described by activities that are derived from several frameworks was examined. The article goes into detail about the prerequisites students' need to handle mathematical proofs and the importance of preservice teachers' ability to notice the students' availability of these prerequisites as a basis for the later diagnosis. Further, the specific diagnostic activities of knowledge-based reasoning are discussed by extracting the diagnostic activities relevant for the study from the comparison of different frameworks.

The main aim of the study presented in Article B was to investigate preservice teachers' diagnostic activities. First, the diagnostic process was described by examining the preservice teachers' noticing based on whether they could focus on relevant information such as facets of the students' mathematical argumentation skills (RQ 1). Second, the preservice teachers' diagnostic activities of knowledge-based reasoning (describing, evaluating, explaining, and decision-making) were explored (RQ 2). Beyond those, both research questions were examined concerning differences in the diagnostic process for students with different levels of mathematical argumentation skills. The free texts written by the participants in the simulation served as the database for the analysis in the article. For the assessment of the texts, we developed a coding scheme using both deductive and inductive methods.

Results show that in the video-based simulation, preservice teachers were largely able to notice the relevant information, after a precise introduction on what to focus on and efficiently proceeded to get an impression of the unknown students and their skills. Preservice teachers not only saw a student on his or her level of thinking holistically but also shifted their focus to certain prerequisites that are especially helpful in diagnosing a particular student's mathematical argumentation skills. During knowledge-based reasoning, it happened that the main activities of the preservice teachers were describing and evaluating. The pattern of the knowledge-based reasoning activities shown by the preservice teachers was shown to be independent of the students who were diagnosed. The study demonstrates that the use of different perspectives in examining the diagnostic process can highlight supplementary aspects that can be combined in meaningful and powerful ways. With the video-based simulation, preservice teachers' diagnostic skills, and especially the diagnostic process, can be examined in authentic classroom situations. Furthermore, the study shows the preservice teachers' difficulties in the diagnostic process and where scaffolds should ideally be implemented.

The study design was developed by all authors. I developed the coding scheme; I handled and analyzed the data, wrote the manuscript, and implemented the suggestions and remarks by co-authors and reviewers. I supervised the submission process and served as a corresponding author for this publication.

6 Discussion

The overall goal of this dissertation is to (i) develop and validate a video-based simulation as a representation of practice and (ii) provide insights into preservice teachers' diagnostic skills in an on-the-fly diagnostic situation from the mathematics classroom represented by the video-based simulation. These questions are empirically addressed and discussed in the two enclosed papers. The present overall discussion thus focuses on the joint discussion of the central findings and their implications for research and practice.

6.1 Interpretation of Central Findings

6.1.1 Validating a representation of practice

In order to gain knowledge about preservice teachers' diagnostic skills, the first goal of this dissertation was to develop and validate a representation of practice for a diagnostic situation. The video-based simulation was intended to be used for research, in particular, to provide insights into the diagnostic process. In addition, it was intended to be used in teacher training programs to facilitate the development of diagnostic skills. The use of video clips (Gaudin & Chaliès, 2015; Kang & van Es, 2018; Tekkumru-Kisa & Stein, 2017), as well as their embedding in a simulation (Kramer et al., 2020; Pickal et al., 2020), is currently discussed in teacher education as a promising method to measure complex skills. However, few such representations of practice have been developed, validated, reported, and discussed in a research context. Based on Grossman et al. (2009), two principles need to be balanced in a representation of practice: the approximation of practice and the decomposition of practice. The video-based simulation is validated based on the realization of these two principles. Thus, results regarding (i) a close relation to a typical real-world situation from the mathematics classroom and (ii) an adequate cognitive demand adapted to the preservice teachers' level of diagnostic skills are considered (Codreanu et al., 2020).

Investigating the realization of an approximation of practice, the first study showed findings regarding the preservice teachers' perception of authenticity, cognitive involvement, and motivation. The video clips showing one-on-one teacher-student interactions were perceived as authentic. This result is in line with other findings of videos showing classroom situations (Seidel et al., 2010). The perceived authenticity supports the idea that video clips can be used to authentically depict classroom situations (Blomberg et al., 2013; Brophy, 2004). However, a direct comparison of the same classroom situation presented via different media (e.g., real-life simulation, videos, text vignettes) is missing (see e.g., Fink et al., 2021). Thus, it remains unclear whether, for an on-the-fly diagnostic situation, a representation using, for example, text vignettes, which are less difficult to generate, can achieve comparable values (Friesen et al., 2018). Since the content of the video clips was scripted, the findings regarding the authenticity indicate that the developed simulated students show authentic behavior while handling a geometrical proof task (Böttcher & Thiel, 2018; Piwowar et al., 2017). This indicates that the abandonment of superficial characteristics, such as a real classroom environment, is still sufficiently immersive for preservice teachers. It thus seems much more important to carefully plan the behavior of the persons in the situation than to record real classroom environments (Dieker et al., 2009). Additionally, our results show that the simulation of the on-the-fly diagnostic situation, including the diagnostic task, was perceived as authentic. Particularly, the perception of the mathematic education experts shows that the diagnostic task, with its purpose, focus, and consequences, was appropriately chosen (Karst et al., 2017). This supports the idea that a well-defined diagnostic task and coherent and clear communication of these characteristics contribute to the perception of an authentic diagnostic situation. Furthermore, the participating preservice teachers reported a high level of cognitive involvement and motivation (Dankbaar et al., 2016; Schubert et al., 2001). Part of this may have been due to the fact that the participants were able to influence the course of actions within the simulation by deciding on the number of video clips to watch before finally diagnosing students' geometrical proving skills. This indicates that even small autonomous interactions with the situation may be helpful to cognitively engage and motivate participants (Ryan & Deci, 2000). The findings indicate that the created simulation is indeed an authentic approximation of practice that may be the fundament to support preservice teachers in acquiring skills that can be transferred from the simulated environment to real-world situations (Grossman et al., 2009).

The positive results concerning an approximation of practice were obtained, although the complexity of the diagnostic situation presented in the video-based simulation was simplified in multiple ways. Investigating the decomposition of practice, the preservice teachers' cognitive demand was considered (Grossman et al., 2009). For this purpose, the difficulty of the diagnostic task was determined and analyzed based on the achieved accuracy of the task (Leuders et al., 2018; Urhahne & Wijnia, 2021). Regarding the relative accuracy (McElvany et al., 2009; Schrader & Helmke, 1987), preservice teachers succeeded in ranking the students according to their level of geometrical proving skills. Distinguishing only four students with quite different levels of proving skills and ranking them did not seem particularly challenging. One reason for this may be that each student showed a different, quite homogeneous level of geometrical thinking (Kunimune et al., 2009; Usiskin, 1982). The simplification of student behavior to homogeneous profiles might have been too strong for such a ranking task. Another reason for the reduced difficulty on the ranking task could be the small number of students. In this case, settings with more students, such as the Simulated Classroom (Südkamp et al., 2008), might be more suitable for a more nuanced measurement of diagnostic skills. For the measurement of relative accuracy, the situation in the video-based simulation seems too much reduced. To determine absolute accuracy, the preservice teachers were asked to diagnose the students with regard to the three prerequisites: mathematical content knowledge, methodological knowledge, and problem-solving strategies. This task showed a medium level of difficulty for the preservice teachers. The medium level of difficulty indicates that the differentiated observing and diagnosing of individual prerequisites is feasible, but poses challenges for the preservice teachers. In addition, it encourages the use of video clips in representations of practice, as it was possible to present different prerequisites from mathematical content knowledge to problem-solving strategies in short video clips in such a way that they could be diagnosed. Overall, the findings suggest that the decomposition of practice was adequate for the target sample such that an overwhelming cognitive demand was avoided. Altogether, these results show that the video-based simulation balances an authentic approximation of a real-world classroom situation with an adequate decomposition of practice. Although both principles seem to be initially contrary to each other, they can be implemented in combination in one environment.

6.1.2 Empirical findings on preservice teachers' diagnostic activities

After validating the video-based simulation, the second goal of this dissertation was to gain insights into the process of diagnosing students' geometrical proving skills in an on-the-fly diagnostic situation. Thus, the preservice teachers' diagnostic activities were investigated in the second study using the video-based simulation. The diagnostic activities of noticing and knowledge-based reasoning were examined (Seidel & Stürmer, 2014). Similar to other studies, texts written by the preservice teachers served as a basis for the analyses of the diagnostic activities (Kramer et al., 2021; Wildgans-Lang et al., 2020). Noticing was examined in terms of focusing on relevant information, as indicated by analyses of prerequisites for geometrical proving (Sommerhoff, 2017; Ufer et al., 2008), for the diagnoses of the simulated students. Knowledge-based reasoning referred to diagnostic activities that are emphasized in the research on teachers' professional vision (Goodwin, 1994; van Es, 2011) and scientific reasoning and argumentation (Fischer et al., 2014).

As the relevant information about the students' geometrical proving skills were their level of the prerequisites (mathematical content knowledge, methodological knowledge, and problem-solving strategies) showed in the video clips (Reiss & Ufer, 2009; Sommerhoff et al., 2015; Ufer et al., 2008), it examined whether preservice teachers focus on these three aspects in their noticing. Findings from the second study suggest that preservice teachers were able to notice information about all three prerequisites. However, the degree of decomposition in the simulated situation was high compared to a real-world diagnostic situation. Many additional factors that influence the students' behavior in real-world situations, such as the students' beliefs about the function of proofs (Sommerhoff, 2017; Sommerhoff et al., 2020) or their motivational-affective characteristics (Kosel et al., 2021; Seidel, 2007), were kept mostly constant with all of the students with as little influence as possible on their geometrical proving skills. It was found that the preservice teachers rarely referred to factors aside from the three prerequisites. This supports the idea that research on this diagnostic situation is limited to findings regarding the assessment of the level of these students' prerequisites. Further, the results show that single cues, such as an indication of a student's prior knowledge on a mathematical concept, seem to be easier to notice than cues that are distributed over different actions, such as information on students' problem-solving strategies, which often have to be distilled from the interplay of several statements or actions. More insights into the noticing were gained by looking at the preservice teachers' patterns concerning different students. The preservice teachers varied the intensity of their foci across the students, adapting to the level of skills demonstrated by each student. This suggests that preservice teachers not only regarded the student holistically (J. Kaiser et al., 2013; Südkamp et al., 2012), but also noticed the particular prerequisites individually. However, these students had homogeneous and consistent profiles. For further research, the diagnosis of inconsistent profiles must be considered (Südkamp et al., 2018). Looking at the cognitive demand of the simulation in terms of noticing, it appears that the preservice teachers mastered this task well. This is in line with findings from the first study, that teachers were cognitively involved and motivated and thus were able to notice the relevant information and distinguish which information was particularly important to individual students.

Results regarding the diagnostic activity of knowledge-based reasoning showed a different picture. Here, the preservice teachers' diagnostic activities were examined concerning the specific activities of describing, evaluating, explaining, and decision making (see also Kramer et al., 2021). These four activities were shown by the preservice teachers with different frequencies. Describing and evaluating were the most frequent reasoning activities, whereas explaining and decision making were less frequent. Hardly any other activities that could not be assigned to any of the four reasoning activities were found in the written texts. This is an indication that the newly developed coding scheme worked well for the simulation within its diagnostic situation. In particular, this demonstrated that the combination of frameworks with different perspectives—in this case, the frameworks from Seidel and Stürmer (2014) and Fischer et al. (2014)—helped describe the diagnostic process in specific diagnostic situations. Specifically, the inclusion of the activity of evaluating showed that many preservice teachers not only described objectively (Stürmer et al., 2013), but had already interpreted the noticed information (E. Bauer et al., 2020). However, most of the preservice teachers' diagnostic processes lacked explanations including references to mathematical concepts. This reveals a potential for further development in the diagnostic process among preservice teachers. Furthermore, the activity of decision making was rarely found. A diagnosis that does not inform a specific educational decision or action directly (high consequences of diagnosis; see Figure 2), but only depicts the level of students' skills, seems to be difficult. Therefore, such diagnoses should be promoted separately from the subsequent step of deriving concrete educational decisions, as this is additionally related to teachers' adaptive teaching skills (Beck et al., 2008; Brühwiler, 2017). Overall, the findings concerning knowledge-based reasoning are in line with prior research on diagnostic activities (E. Bauer et al., 2020; Stürmer et al., 2013). The pattern of knowledge-based reasoning activities was stable regardless of the student diagnosis. Since the preservice teachers had to diagnose the same student skills, they needed similar professional knowledge for the knowledge-based reasoning about the students' behavior and skills. This could be one explanation for the lack of differences in the distribution of knowledge-based reasoning activities. Altogether, the simulation revealed insights in the diagnostic process and thus showed potential for further development regarding the learning environment.

6.2 Implications for Teacher Education

From a practical perspective, diagnosing students' learning prerequisites, processes, and outcomes provides a basis to adapt teaching to the students' needs (Beck et al., 2008; Vogt & Rogalla, 2009). The findings from this dissertation are of practical relevance and have implications for research on teacher education and professional development. One main finding from this dissertation is that the video-based simulation is a best-practice example for designing and developing a representation of practice. The validation of this simulation in the first study examining the principles approximation of practice and decomposition of practice (Grossman et al., 2009) suggests that this representation of practice has the potential to be used (i) in educational research, as it allows for a contextualized and standardized method to measure preservice teachers' diagnostic skills, and (ii) in teacher training programs to facilitate the development of diagnostic skills, as the close relation to real-world professional practice allows for the acquisition of practice-oriented knowledge.

The video-based simulation shows that by combining staged videos with an interactive digital environment, an approximation of practice may be developed that is perceived as authentic by expert teachers and preservice teachers (c.f. Kramer et al., 2020; Seidel et al., 2010). The review of prior research (e.g., Südkamp & Praetorius, 2017) was of central importance in the development of this approximation of practice, as it provided multiple necessary categories to consider when creating the simulation. For example, a typical situation with a relevant diagnostic task was selected based on the characteristics of situations (Karst et al., 2017). In addition, the students' skills to be diagnosed were specified based on findings on students' prerequisites for handling mathematical proofs from mathematics education research (Reiss & Ufer, 2009; Sommerhoff et al., 2015; Weigand et al., 2014). Further, care was taken to imitate professional practice in the simulation and, thus, to enable preservice teachers to perform activities that are central to their later professional practice. These decisions led to a video-based simulation with which preservice teachers' diagnostic activities in an on-the-fly situation were examined. However, it remains to be seen how expert teachers actually act in such an on-the-fly situation, and thus what should be considered as an ideal diagnostic process that can be used as a prototype in higher education teaching. The research on professional vision (Seidel & Stürmer, 2014) and scientific reasoning and argumentation (Fischer et al., 2014) suggests certain activities, but their connections and an ideal process remain vague. Findings of expert teachers' diagnostic processes could contribute to sharpening and broadening such frameworks. This encourages expert-novice comparisons in further research (AERA, 2002) that supports the findings from this dissertation regarding the aspects wherein preservice teachers show potential to develop their diagnostic skills. Furthermore, it would be interesting for the acquisition of such skills to see how they develop throughout teacher training programs (Kunter et al., 2011). Longitudinal studies could provide additional insights for the development and acquisition of professional skills and how to support them.

The possibilities of staged videos (Dieker et al., 2009; van Es et al., 2020) and simulations (Frasson & Blanchard, 2012; Shannon, 1975) to decompose, reduce, and simplify situations from professional practice are manifold, such that an adequate demand can be achieved for different target groups. The video-based simulation showed an adequate cognitive demand for preservice teachers, but it is unclear which specific decisions in the development of the simulation (besides general considerations) enabled this cognitive demand. Further research should investigate systematically how it is possible to adapt the cognitive demand of a representation of practice to a group of learners in advance. A guideline for the development of such representations of practice would be helpful to make it easier to develop further simulation-based environments for educational research and higher education.

In addition to combining the two principles of the framework by Grossman et al. (2009) together in one representation of practice, simulations have several advantages compared to common methods in teacher training programs such as internships. Simulations offer a protected environment in which preservice teachers may be prepared for areas of teaching that are otherwise often neglected in teacher training. As a digital environment, such a simulation may be easily implemented in different teacher training programs. Smaller adaptations can be made. For example, the familiarization of the situation could be adapted to situations known to the preservice teachers (e.g., in other educational systems or school tracks). However, the video clips do not allow for later editing. Virtual reality alternatives would allow changes to be made to the simulated situation at any time and should therefore be considered in the development of further digital representations of practice (Wiepke et al., 2019). Digital environments such as video-based simulation offer the possibility of working in them repeatedly (Böheim et al., 2020). There are options to adjust the level of cognitive demand to the participants. In the video-based simulation, this could be done by varying the number of students, or the development of additional simulated students (showing inconsistent profiles). With the insights in the diagnostic process, instructional support like scaffolds or prompts can be developed and implemented in the simulation additionally to support the development of diagnostic skills.

Other implications for practice are the findings of the dissertation on preservice teachers' diagnostic skills and their diagnostic activities. Research has long investigated the accuracy of teachers diagnosing students' skills (Karing, 2009; Südkamp et al., 2012; Urhahne & Wijnia, 2021). A greater emphasis in research on the underlying cognitive processes and (visible) diagnostic activities as causes of accurate and inaccurate diagnoses evolved recently (Herppich et al., 2018; Loibl et al., 2020). Before examining the teachers' diagnostic processes, methods must be developed to measure the diagnostic activities. In this regard, the present dissertation presents a coding scheme to get insights into the performed diagnostic activities in an on-the-fly diagnostic situation. Thus, written texts of preservice teachers were analyzed. The question arises of how close the measured activities are to the actual cognitive processes of the preservice teachers. A comparison of the results from this research with other measurement methods, such as the method of recording think-aloud protocols, is still pending (Charters, 2003; Cotton & Gresty, 2006).

The frameworks for professional competencies (Blömeke et al., 2015; Heitzmann et al., 2019; Herppich et al., 2018) assume a relationship between the activities in the process and the performance. In the context of the video-based simulation, this relationship needs further investigation. Likewise, for teacher education, it is worthy to investigate the influence of teachers' prerequisites on diagnostic activities. Overall, the results of this research show that preservice teachers had only a mediocre performance in a highly relevant, partially decomposed diagnostic situation. A comparable real-world situation has a higher demand, because there are many confounding factors, such as motivational-affective characteristics, that can interfere with a diagnosis (Huber & Seidel, 2018; Kosel et al., 2021). Similarly, class size was severely depleted with four simulated students, and these simulated students did not show any inconsistent profiles (Südkamp et al., 2018). Therefore, the results of this thesis can be used to raise the awareness of facilitating preservice teachers' diagnostic skills at an early stage of their education. Teacher training programs should aim to ensure that preservice teachers are already familiar with these diagnostic situations before they enter the profession. When it comes to promoting preservice teachers' diagnostic skills, further research should examine how transferable diagnostic skills are with regard to different situations. For example, the process could be practiced in parallel in other subject areas (Kron et al., 2021; Wildgans-Lang et al., 2020) or even in other subjects (Kramer et al., 2020; Pickal et al., 2020). A transfer of the diagnostic skills from simulated situations to real-world situations is decisive for simulated diagnostic situations. Furthermore, it is still largely unclear to what extent such simulations already function as learning environments without further instructional support.

6.3 Limitations and Further Research

Some limitations should be considered when interpreting the dissertation's results and their implications for further research and educational practice. To begin with, the two studies with their findings rely on data from only one specific on-the-fly formative diagnostic situation regarding the assessment of specific students' geometrical proving skills represented in the video-based simulation. Based on Schrader's (2011) definition of diagnostic skills, these skills refer to the ability to cope with different diagnostic situations. It remains unclear to what extent comparable situations allow a more reliable measurement of preservice teachers' diagnostic skills without other professional skills influencing the measurements in varying diagnostic situations. Each student in the video-based simulation and the diagnosis of the student's skills can represent a new situation, whereby these situations show high similarity in their characteristics (c.f. Karst et al., 2017). Further research on measurements of preservice teachers' diagnostic skills regarding additional students working on the same geometrical proof task is required to shed light on the quality of measurement of diagnostic skills in a video-based simulation. Another possibility to increase the quality of the measurement of preservice teachers' diagnostic skills would be the use of multiple situations with similar characteristics but different contents. Diagnosing in a situation where students are working on a different geometrical proof task from the same subject area, such as proving the inversion of Thales' theorem (ISB, 2021), could be such a new situation. It remains open for further research to what extent the measurement of diagnostic skills in variable situations is comparable, and to what extent preservice teachers' prerequisites such as professional knowledge influence these measurements.

Both studies included a relatively small sample of preservice teachers with a total number of $N = 79$ preservice teachers. The data collection suffered from technical problems such that data from some participants were lost or a comparison with the other data was impossible. Further, this sample is quite homogeneous, for example, concerning their professional knowledge and experiences. For further studies, preservice teachers from different stages in their teacher training programs could be recruited to represent a broader range of students and to increase the sample sizes.

To measure the performance in this dissertation, a score was computed for the absolute accuracy, which cannot reflect all differences in the performance of the preservice teachers. Absolute accuracy was calculated considering the agreement with an expert solution on the Likert response scales, whereby the deviation from this solution was not taken into account. Thus, the score did not consider whether someone gave an assessment that was only slightly or completely different from the expert solution. One possibility to obtain a higher validity would be to weigh the participants' answers based on the distribution of experts' answers from a preceding expert study (Hammer, 2015).

For further validation, the video-based simulation could have been compared with other representations of practice. When the same diagnostic situation is presented as text or with actors in a live situation, a direct comparison of the perceived approximation of practice could be realized (Fink et al., 2021). This could extend the results of the dissertation on the realization of an approximation of practice by highlighting which steps (e.g., the use of staged video clips) mainly contribute to this realization.

Data for the analyses of the diagnostic activities limits conclusions regarding relations of the diagnostic activities. The preservice teachers' final written diagnoses were used for the analysis. These texts were possible to code low inferently. The measurement of activities through the existing texts allowed us to gain knowledge about the process without interfering significantly. To get better measurements of the actual cognitive activities of the preservice teachers, the coding scheme has to be further developed in order to be used for texts that were written during the process.

6.4 Conclusion

The acquisition of diagnostic skills in higher education is one goal of teacher training programs, but the use of simulation-based learning environments at university is scarce. This dissertation introduces a newly developed video-based simulation as a representation of practice for investigating preservice teachers' diagnostic skills to build a basis for simulation-based learning environments. The dissertation contributes to the field of research on teachers' diagnostic skills in an on-the-fly diagnostic situation regarding the diagnosis of students' geometrical proving skills. First, the dissertation has a *conceptual value* since the process of developing and designing a representation of practice was based on the framework by Grossman et al. (2009), showing that the framework with its principles can guide the designing of representations of practice. This should inspire researchers in the educational field to employ guidelines for the development of simulation-based environments for higher education. Further, it was shown that the three student prerequisites are suitable to represent geometrical proving skills in diagnostic situations. The findings from the studies with the video-based simulation contribute to the understanding of preservice teachers' diagnostic skills by showing different diagnostic activities and, thus, indicating the preservice teachers' difficulties in such diagnostic situations. Second, the validation of the video-based simulation adds *methodological value* to the field, since it represents an authentic environment in which professional skills can be investigated in a standardized and valid way. In addition, the development of a coding scheme to investigate diagnostic activities based on texts written by preservice teachers could provide further insights into the diagnostic process. With a domain-specific view of noticing in connection with multiple perspectives on knowledge-based reasoning, collective insights into the diagnostic process can be achieved. Third, the video-based simulation has a *practical value* since it can be used in higher teacher education to support the knowledge about, and the acquisition of, diagnostic skills. Furthermore, there are possibilities to enrich the video-based simulation with additional support in the form of scaffolds, further learning tasks, or possibilities for reflection at many points.

7 References

AERA. (2002). *Standards for educational and psychological testing*. American Educational Research Association.

Alles, M., Apel, J., Seidel, T., & Stürmer, K. (2019). How candidate teachers experience coherence in university education and teacher induction: The influence of perceived professional preparation at university and support during teacher induction. *Vocations and Learning*, *12*(1), 87–112. https://doi.org/10.1007/s12186-018-9211-5

Alonzo, A. C. (2011). Learning progressions that support formative assessment practices. *Measurement: Interdisciplinary Research & Perspective*, *9*(2-3), 124–129. https://doi.org/10.1080/15366367.2011.599629

Altmann, A. F., & Nückles, M. (2017). Empirische Studie zu Qualitätsindikatoren für den diagnostischen Prozess. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 142–150). Waxmann.

Artelt, C., & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift Für Pädagogische Psychologie*, *23*(34), 157–160. https://doi.org/10.1024/1010-0652.23.34.157

Barnhart, T., & van Es, E. (2015). Studying teacher noticing: Examining the relationship among pre-service science teachers' ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education*, *45*, 83–93. https://doi.org/10.1016/j.tate.2014.09.005

Bauer, E., Fischer, F., Kiesewetter, J., Shaffer, D. W., Fischer, M. R., Zottmann, J. M., Sailer, & Michael (2020). Diagnostic activities and diagnostic practices in medical education and teacher education: An interdisciplinary comparison. *Frontiers in Psychology*, *11*, 562665. https://doi.org/10.3389/fpsyg.2020.562665

Bauer, J., & Prenzel, M. (2012). Science education: European teacher training reforms. *Science*, *336*(6089), 1642–1643. https://doi.org/10.1126/science.1218387

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift Für Erziehungswissenschaft*, *9*(4), 469–520. https://doi.org/10.1007/s11618-006-0165-2

Beck, E., Baer, M., Guldimann, T., Bischoff, S., Brühwiler, C., Müller, P., Niedermann, R., Rogalla, M., & Vogt, F. (2008). *Adaptive Lehrkompetenz: Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Pädagogische Psychologie *und Entwicklungspsychologie: Vol. 63*. Waxmann.

Behrmann, L., & Glogger-Frey, I. (2017). Produkt- und Prozessindikatoren diagnostischer Kompetenz. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 134–142). Waxmann.

Behrmann, L., & van Ophuysen, S. (2017). Das Vier-Komponenten-Modell der Diagnosequalität. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 38–42). Waxmann.

Bennack, J., & Jürgens, E. (2002). Schulpraktika in Lehramtsstudiengängen. In H.-U. Otto, T. Rauschenbach, & P. Vogel (Eds.), *Erziehungswissenschaft: Lehre und Studium* (pp. 143–160). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-93238-9_12

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, *21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Blomberg, G., Renkl, A., Sherin, M. G., Borko, H., & Seidel, T. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal of Educational Research Online*, *5*(1), 90–114.

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. (2013). Assessment of competencies in higher education. *Zeitschrift Für Psychologie*, *221*(3), 202. https://doi.org/10.1027/2151-2604/a000148

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies. *Zeitschrift Für Psychologie*, *223*(1), 3–13. https://doi.org/10.1027/2151-2604/a000194

Boero, P. (Ed.). (2007). *New directions in mathematics and science education: Vol. 2. Theorems in school: From history, epistemology and cognition to classroom practice*. Sense Publishers.

Böttcher, F., & Thiel, F. (2018). Evaluating research-oriented teaching: A new instrument to assess university students' research competences. *Higher Education*, *75*(1), 91–110. https://doi.org/10.1007/s10734-017-0128-y

Brophy, J. E. (2004). *Using video in teacher education. Advances in research on teaching*. Elsevier JAI.

Brühwiler, C. (2017). Diagnostische und didaktische Kompetenz als Kern adaptiver Lehrkompetenz. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 123–133). Waxmann.

Carlson, J., Daehler, K. R., Alonzo, A. C., Barendsen, E., Berry, A., Borowski, A., Carpendale, J., Kam Ho Chan, K., Cooper, R., Friedrichsen, P., Gess-Newsome, J., Henze-Rietveld, I., Hume, A., Kirschner, S., Liepertz, S., Loughran, J., Mavhunga, E., Neumann, K., Nilsson, P., . . . Wilson, C. D. (2019). The Refined Consensus Model of Pedagogical Content Knowledge in Science Education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 77–94). Springer Singapore. https://doi.org/10.1007/978-981-13-5898-2_2

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal*, *12*(2), 68–82.

Chernikova, O., Heitzmann, N., Fink, M. C., Timothy, V., Seidel, T., & Fischer, F. (2020). Facilitating diagnostic competences in higher education: A meta-analysis in medical and teacher education. *Educational Psychology Review*, *68*, 157–196. https://doi.org/10.1007/s10648-019-09492-2

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-Based Learning in Higher Education: A Meta-Analysis. *Review of Educational Research*, *90*(4), 499–541. https://doi.org/10.3102/0034654320933544

Chinnappan, M., Ekanayake, M. B., & Brown, C. (2011). Specific and general knowledge in geometric proof development. *SAARC Journal of Educational Research*, *8*, 1–28.

Cochran-Smith, M., & Zeichner, K. M. (2005). Studying teacher education. *The Report of the AERA Panel on Research and Teacher Education.*

Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, *95*, 103146. https://doi.org/10.1016/j.tate.2020.103146

Codreanu, E., Sommerhoff, D., Huber, S. A., Ufer, S., & Seidel, T. (2021). Exploring the process of preservice teachers' diagnostic activities in a video-based simulation. *Frontiers in Education*, *6*, 626666. https://doi.org/10.3389/feduc.2021.626666

Cook, D. A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J. H., Wang, A. T., Erwin, P. J., & Hamstra, S. J. (2011). Technology-enhanced simulation for health professions education: A systematic review and meta-analysis. *JAMA*, *306*(9), 978–988. https://doi.org/10.1001/jama.2011.1234

Correa, J. M., Martínez-Arbelaiz, A., & Aberasturi-Apraiz, E. (2015). Post-modern reality shock: Beginning teachers as sojourners in communities of practice. *Teaching and Teacher Education*, *48*, 66–74. https://doi.org/10.1016/j.tate.2015.02.007

Cotton, D., & Gresty, K. (2006). Reflecting on the think-aloud method for evaluating e-learning. *British Journal of Educational Technology*, *37*(1), 45–54. https://doi.org/10.1111/j.1467-8535.2005.00521.x

Dankbaar, M. E. W., Alsma, J., Jansen, E. E. H., van Merrienboer, J. J. G., van Saase, J. L. C. M., & Schuit, S. C. E. (2016). An experimental study on the effects of a simulation game on students' clinical cognitive skills and motivation. *Advances in Health Sciences Education: Theory and Practice*, *21*(3), 505–521. https://doi.org/10.1007/s10459-015-9641-x

Darling-Hammond, L., & Bransford, J. (2007). *Preparing teachers for a changing world: what teachers should learn and be able to do*. San Francisco: Jossey-Bass.

de Jong, T. (2006). Computer simulations. Technological advances in inquiry learning. *Science*, *312*(5773), 532–533. https://doi.org/10.1126/science.1127750

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation, and Accountability*, *28*(3), 251–272. https://doi.org/10.1007/S11092-015-9233-6

Derry, S. J., Sherin, M. G., & Sherin, B. L. (2014). Multimedia learning with video. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 785–812). Cambridge University Press.

Dicke, T., Elling, J., Schmeck, A., & Leutner, D. (2015). Reducing reality shock: The effects of classroom management skills training on beginning teachers. *Teaching and Teacher Education*, *48*, 1–12. https://doi.org/10.1016/j.tate.2015.01.013

Dieker, L. A., Lane, H. B., Allsopp, D. H., O'Brien, C., Butler, T. W., Kyger, M., Lovin, L., & Fenty, N. S. (2009). Evaluating video models of evidence-based instructional practices to enhance teacher learning. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *32*(2), 180–196. https://doi.org/10.1177/0888406409334202

European Commission. (2013). *Supporting teacher competence development: for better learning outcomes*. http://ec.europa.eu/assets/eac/education/experts-groups/2011-2013/teacher/teachercomp_en.pdf.

Fink, M. C., Reitmeier, V., Stadler, M., Siebeck, M., Fischer, F., & Fischer, M. R. (2021). Assessment of diagnostic competences with standardized patients versus virtual patients: Experimental Study in the Context of History Taking. *Journal of Medical Internet Research*, *23*(3), e21196. https://doi.org/10.2196/21196

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Neuhaus, B. J., Dorner, B., Pankofer, S., Fischer, M. R., Strijbos, J.-W., Heene, M., & Eberle, J. (2014). Scientific Reasoning and Argumentation: Advancing an Iinterdisciplinary Research Agenda in Education. *Frontline Learning Research*, *2*(3), 28–45. https://doi.org/10.14786/flr.v2i3.96

Fischer, F., & Opitz, A. (Eds.). (in press). *Learning to diagnose with simulations: Examples from teacher education and medical education*. Springer.

Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., Reiss, K., Stürmer, K., Siebeck, M., Schmidmaier, R., Seidel, T., Ufer, S., Wecker, C., & Neuhaus, B. (2018). Systematizing professional knowledge of medical doctors and teachers: development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences*, *8*(4), 207. https://doi.org/10.3390/educsci8040207

Fraser, S., Beswick, K [Kim], & Crowley, S. (2019). Making tacit knowledge visible: Uncovering the knowledge of science and mathematics teachers. *Teaching and Teacher Education*, *86*, 102907. https://doi.org/10.1016/j.tate.2019.102907

Frasson, C., & Blanchard, E. (2012). Simulation-based learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning.* Springer US.

Friesen, M., Kuntze, S., & Vogel, M. (2018). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß, & T. Dörfler (Eds.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (pp. 153–177). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20121-0_8

Funder, D. C. (2012). Accurate Personality Judgment. *Current Directions in Psychological Science*, *21*(3), 177–182. https://doi.org/10.1177/0963721412445309

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., León, V. de, Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*, *44*(3), 267–291. https://doi.org/10.1007/s11251-016-9371-3

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41–67. https://doi.org/10.1016/j.edurev.2015.06.001

Gegenfurtner, A., Lewalter, D., Lehtinen, E., Schmidt, M., & Gruber, H. (2020). Teacher Expertise and Professional Vision: Examining Knowledge-Based Reasoning of Pre-Service Teachers, In-Service Teachers, and School Principals. *Frontiers in Education*, *5*, 59. https://doi.org/10.3389/feduc.2020.00059

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *96*(3), 606–633.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: a cross-professional perspective. *Teachers College Record*, *111*(9), 2055–2100.

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, *45*(1), 184–205. https://doi.org/10.3102/0002831207312906

Gruber, H., Mandl, H., & Renkl, A. (2000). Was lernen wir in der Schule und Hochschule: Träges Wissen? In H. Mandl (Ed.), *Die Kluft zwischen Wissen und Handeln: Empirische und theoretische Lösungsansätze* (1st ed., pp. 139–156). Hogrefe.

Hammer, S. (2015). *Professionelle Kompetenz von Mathematiklehrkräften im Umgang mit Aufgaben in der Unterrichtsplanung: Theoretische Grundlegung und empirische Untersuchung* [Dissertation]. Ludwig-Maximilians Universität, München.

Harel, G., & Sowder, L. (1998). Students' Proof Schemes: Results from Exploratory Studies. *CBMS Issues in Mathematics Education*, *7*, 234–283.

Hartmann, S. (1996). The world is a process: Simulations in the natural and social sciences. In R. Hegselmann, U. Mueller, & K. G. Troitzsch (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Springer Netherlands.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, *31*(4), 396–428.

Heinze, A., & Kwak, J. Y. (2002). Informal prerequisites for informal proofs. *Zentralblatt Für Didaktik Der Mathematik*, *34*(1), 9–16. https://doi.org/10.1007/BF02655688

Heinze, A., & Reiss, K. (2003). *Reasoning and proof: methodological knowledge as a component of proof competence*. www.lettredelapreuve.itlCERME3PapersiHeinze-paperl.pdf

Heitzmann, N., Seidel, T., Hetmanek, A., Wecker, C., Fischer, M. R., Ufer, S., Schmidmaier, R., Neuhaus, B., Siebeck, M., Stürmer, K., Obersteiner, A., Reiss, K., Girwidz, R., Fischer, F., & Opitz, A. (2019). Facilitating Diagnostic Competences in Simulations in Higher Education A Framework and a Research Agenda. *Frontline Learning Research*, 1–24. https://doi.org/10.14786/flr.v7i4.384

Helmke, A. (2007). *Unterrichtsqualität: Erfassen, Bewerten, Verbessern* (5. Aufl.). 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. https://doi.org/10.1016/0742-051X(87)90010-2

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. https://doi.org/10.1016/j.tate.2017.12.001

Hertel, J. P., & Millis, B. J. (2002). *Using simulations to promote learning in higher education: An introduction* (1st ed.). *Enhancing learning series*. Stylus Pub.

Hoge, R. D., & Coladarci, T. (1989). Teacher-Based Judgments of Academic Achievement: A Review of Literature. *Review of Educational Research*, *59*(3), 297–313. https://doi.org/10.3102/00346543059003297

Hosenfeld, I., Helmke, A., & Schrader, F.-W. (2002). Diagnostische Kompetenz. Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Unterrichtsstudie SALVE. *Zeitschrift Für Pädagogik*, *45*.

Huber, S. A., & Seidel, T. (2018). Comparing teacher and student perspectives on the interplay of cognitive and motivational-affective student characteristics. *PloS One*, *13*(8), e0200609. https://doi.org/10.1371/journal.pone.0200609

Ingenkamp, K.-H., & Lissmann, U. (2008). *Lehrbuch der Pädagogischen Diagnostik* (6. Auflage). Beltz.

ISB. (2021). *LehrplanPLUS Gymnasium. Fachlehrplan Mathematik*. https://www.lehrplanplus.bayern.de/fachlehrplan/gymnasium/7/mathematik

Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M., & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers—cognitive versus situated approaches. *Educational Studies in Mathematics*, *94*(2), 161–182. https://doi.org/10.1007/s10649-016-9713-8

Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, *28*, 73–84. https://doi.org/10.1016/j.learninstruc.2013.06.001

Kang, H., & van Es, E. A. (2018). Articulating design principles for productive use of video in preservice education. *Journal of Teacher Education*, *5*(1), 1-14. https://doi.org/10.1177/0022487118778549

Karing, C. (2009). Diagnostische Kompetenz von Grundschul- und Gymnasiallehrkräften im Leistungsbereich und im Bereich Interessen. *Zeitschrift Für Pädagogische Psychologie*(23), 197–209. https://doi.org/10.1024/1010-0652.23.34.197

Karst, K., & Dickhäuser, O. (2016). Das Ganze und die Summe der Teile: Aggregation schülerspezifischer Lehrerurteile zur Messung schülerglobaler Urteilsakkuratheit. *Vortrag Auf Der 50. Tagung Der Deutschen Gesellschaft Für Psychologie in Leipzig*.

Karst, K., & Förster, N. (2017). Ansätze der Modellierung diagnostischer Kompetenz - ein Überblick [Approaches to the modeling of diagnostic competences—an overview]. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 19–20). Waxmann.

Karst, K., Klug, J., & Ufer, S. (2017). Strukturierung diagnostischer Situationen im inner- und außerunterrichtlichen Handeln von Lehrkräften. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 102–114). Waxmann.

Koeppen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current issues in competence modeling and assessment. *Zeitschrift Für Psychologie / Journal of Psychology*, *216*(2), 61–73. https://doi.org/10.1027/0044-3409.216.2.61

Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review*, *6*(1), 3–34. https://doi.org/10.1007/BF00155578

Kosel, C., Wolter, I., & Seidel, T. (2021). Profiling secondary school students in mathematics and German language arts using learning-relevant cognitive and motivational-affective characteristics. *Learning and Instruction*, *73*, 101434. https://doi.org/10.1016/j.learninstruc.2020.101434

Kramer, M., Förtsch, C., Seidel, T., & Neuhaus, B. J. (2021). Comparing two constructs for describing and analyzing teachers' diagnostic processes. *Studies in Educational Evaluation*, *68*, 100973. https://doi.org/10.1016/j.stueduc.2020.100973

Kramer, M., Förtsch, C., Stürmer, J., Förtsch, S., Seidel, T., & Neuhaus, B. J. (2020). Measuring biology teachers' professional vision: Development and validation of a video-based assessment tool. *Cognet Journal.* Advance online publication. https://doi.org/10.1080/2331186X.2020.1823155

Kron, S., Sommerhoff, D., Achtner, M., & Ufer, S. (2021). Selecting Mathematical Tasks for Assessing Student's Understanding: Pre-Service Teachers' Sensitivity to and Adaptive Use of Diagnostic Task Potential in Simulated Diagnostic One-To-One Interviews. *Frontiers in Education*, *6*, Article 604568. https://doi.org/10.3389/feduc.2021.604568

Kultusministerkonferenz. (2004). *Standards für die Lehrerbildung: Bildungswissenschaften*. KMK.

Kultusministerkonferenz. (2012). *Bildungsstandards im Fach Mathematik für die allgemeine Hochschulreife*. KMK.

Kunimune, S., Fujita, T., & Jones, K. (2009). Strengthening studnets' understanding of 'proof' in geometry in lower secondary school. *Proceedings of CERMCE 6*, 756–765.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann.

Leuders, T., Dörfler, T., Leuders, J., & Philipp, K. (2018). Diagnostic Competence of Mathematics Teachers: Unpacking a Complex Construct. In T. Leuders, K. Philipp, & J. Leuders (Eds.), *Diagnostic Competence of Mathematics Teachers* (pp. 3–31). Springer International Publishing. https://doi.org/10.1007/978-3-319-66327-2_1

Leuders, T., & Loibl, K. (2020). Processing Probability Information in Nonnumerical Settings - Teachers' Bayesian and Non-bayesian Strategies During Diagnostic Judgment. *Frontiers in Psychology*, *11*, 678. https://doi.org/10.3389/fpsyg.2020.00678

Levin, D. M., Hammer, D., & Coffey, J. E. (2009). Novice teachers' attention to student thinking. *Journal of Teacher Education*, *60*(2), 142–154. https://doi.org/10.1177/0022487108330245

Lin, F.-L. (2005). Modeling students' learning on mathematical proof and refutation. *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*(1), 3–18.

Lindmeier, A. (2013). Video-vignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In U. Riegel & K. Macha (Eds.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (pp. 45–61). Waxmann Verlag.

Loibl, K., Leuders, T., & Dörfler, T. (2020). A Framework for Explaining Teachers' Diagnostic Judgements by Cognitive Modeling (DiaCoM). *Teaching and Teacher Education*, *91*, 103059. https://doi.org/10.1016/j.tate.2020.103059

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core Practices and Pedagogies of Teacher Education. *Journal of Teacher Education*, *64*(5), 378–386. https://doi.org/10.1177/0022487113493807

McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Horz, H., & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern [Teachers' diagnostic skills to judge student performance and task difficulty when learning materials include instructional pictures]. *Zeitschrift Für Pädagogische Psychologie*, *23*(34), 223–235. https://doi.org/10.1024/1010-0652.23.34.223

NCTM. (2000). *Principles and Standards for School Mathematics.* National Council of Teachers of Mathematics (Ed.), Reston, VA: NCTM.

Nutley, S., Walter, I., & Davies, H. T. O. (2003). From Knowing to Doing. *Evaluation*, *9*(2), 125–148. https://doi.org/10.1177/1356389003009002002

Oser, F., & Renold, U. (2005). Kompetenzen von Lehrpersonen — über das Auffinden von Standards und ihre Messung. In I. Gogolin, H.-H. Krüger, D. Lenzen, & T. Rauschenbach (Eds.), *Standards und Standardisierungen in der Erziehungswissenschaft* (pp. 119–140). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80769-4_9

Paas, F., Tuovinen, J. E., Tabbers, H., & van Gerven, P. S. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, *66*(1), 23–41. https://doi.org/10.1007/s10649-006-9057-x

Pickal, A., Chinn, C., Girwidz, R., Neuhaus, B. J., & Wecker, C. (2020). Diagnosing scientific reasoning skills in inquiry settings: The role of pre-service teachers' professional knowledge. In M. Gresalfi & I. S. Horn (Eds.), *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS)* (Vol. 4, pp. 2265–2268). International Society of the Learning Sciences.

Piwowar, V., Barth, V. L., Ophardt, D., & Thiel, F. (2017). Evidence-based scripted videos on handling student misbehavior: The development and evaluation of video cases for teacher education. *Professional Development in Education*, *44*(3), 369–384. https://doi.org/10.1080/19415257.2017.1316299

Pólya, G. (1973). *How to solve it: A new aspect of mathematical method* (New Princeton science library ed.). *Princeton science library*. Princeton Univ. Press.

Praetorius, A.-K., Koch, T., Scheunpflug, A., Zeinz, H., & Dresel, M. (2017). Identifying determinants of teachers' judgment (in)accuracy regarding students' school-related motivations using a Bayesian cross-classified multi-level model. *Learning and Instruction*, *52*, 148–160. https://doi.org/10.1016/j.learninstruc.2017.06.003

Praetorius, A.-K., Lipowsky, F., Karst, K., Lazarides, R., & Ittel, A. (2012). Diagnostische Kompetenz von Lehrkräften: Aktueller Forschungsstand, unterrichtspraktische Umsetzbarkeit und Bedeutung für den Unterricht [Diagnostic competences of teachers: Current state of research, teaching practicability and importance for teaching]. In R. Lazarides & A. Ittel (Eds.), *Differenzierung im mathematisch-naturwissenschaftlichen Unterricht: Implikationen für Theorie und Praxis* (pp. 115–146). Klinkhardt.

R Core Team. (2020). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org

Radkowitsch, A., Fischer, M. R., Schmidmaier, R., & Fischer, F. (2020). Learning to diagnose collaboratively: Validating a simulation for medical students. *GMS Journal for Medical Education*, *37*(5), Doc51. https://doi.org/10.3205/zma001344

Reiss, K., Hellmich, F., & Thomas, J. (2002). Individuelle und schulische Bedingungsfaktoren für Argumentationen und Beweise im Mathematikunterricht. In M. Prenzel & J. Doll (Eds.), *Bildungsqualität von Schule: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen* (pp. 51–64). Beltz.

Reiss, K., & Törner, G. (2007). Problem solving in the mathematics classroom: the German perspective. *ZDM*, *39*(5-6), 431–441. https://doi.org/10.1007/s11858-007-0040-5

Reiss, K., & Ufer, S. (2009). Was macht mathematisches Arbeiten aus? Empirische Ergebnisse zum Argumentieren, Begründen und Beweisen. *Jahresbericht Der Deutschen Mathematiker-Vereinigung*, *111*(4), 1–23.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, *44*(1), 57–84.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037//0003-066X.55.1.68

Schack, E. O., Fisher, M. H., & Wilhelm, J. A. (Eds.). (2017). *Research in Mathematics Education. Teacher noticing: Bridging and broadening perspectives, context, and framework*. Springer. https://doi.org/10.1007/978-3-319-46753-5

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*(3), 205–221. https://doi.org/10.1007/BF01323044

Schoenfeld, A. H. (1985). *Mathematical problem solving* (2 [Dr.]. Acad. Press.

Schoenfeld, A. H. (1992). Learning to think mathematically: problem solving, metacogntion, and sense making in mathematics. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Simon & Schuster.

Schoenfeld, A. H. (2011). Noticing Matters. A lot. Now what? In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 223–238). Routledge.

Schrader, F.-W. (2011). Lehrer als Diagnostiker. In E. Terhart (Ed.), *Handbuch der Forschung zum Lehrerberuf* (pp. 683–698). Waxmann.

Schrader, F.-W. (2013). Diagnostische Kompetenz von Lehrpersonen [Diagnostic compenteces of teachers]. *Beiträge Zur Lehrerbildung*, *31*(2), 154–165.

Schrader, F.-W. (2017). Diagnostische Kompetenz von Lehrkräften: Anmerkungen zur Weiterentwicklung des Konstrukts [Diagnostic competences of teachers: Comments on the further development of the construct]. In A. Südkamp & A.-K. Praetorius (Eds.), *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen* (pp. 247–256). Waxmann.

Schrader, F.-W., & Helmke, A. (1987). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen [Diagnostic competences of teachers: components and effects]. *Empirische Pädagogik*, *1*, 27–52.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence*, *10*(3), 266–281.

Seidel, T. (2007). The role of student characteristics in studying micro teaching–learning environments. *Learning Environments Research*, *9*(3), 253–271. https://doi.org/10.1007/s10984-006-9012-x

Seidel, T., Blomberg, G., & Stürmer, K. (2010). Observer: Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht [Observer: Validation of a video-based instrument for investigating the professional perception of lessons]. *Zeitschrift Für Pädagogik*, *56*(56), 296–306.

Seidel, T., & Krapp, A. (Eds.). (2014). *Pädagogische Psychologie* (6., vollst. überarb. Aufl.). Beltz.

Seidel, T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*, *51*(4), 739–771. https://doi.org/10.3102/0002831214531321

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, *27*(2), 259–267. https://doi.org/10.1016/j.tate.2010.08.009

Shannon, R. E. (1975). *Systems Simulation, The Art and Science*. Inc., Englewood Cliffs.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008a). On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers. *Applied Measurement in Education*, *21*(4), 295–314. https://doi.org/10.1080/08957340802347647

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008b). On the Impact of Curriculum-Embedded Formative Assessment on Learning: A Collaboration between Curriculum and Assessment Developers. *Applied Measurement in Education*, *21*(4), 295–314. https://doi.org/10.1080/08957340802347647

Sherin, M. G., Russ, R. S., & Colestock, A. A. (2011). Accessing mathematics teachers' in-the-moment noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 79–94). Routledge.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616.

Sommerhoff, D. (2017). *The inidividual cognitive resources underlying students' mathematical argumentation and proof skills: From theory to intervention* [Dissertation]. Ludwig-Maximilians-University, Munich.

Sommerhoff, D., Brunner, E., & Ufer, S. (2020). How beliefs shape the selection of proofs for classroom instruction. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Interim proceedings of the 44th conference of the international group for the psychology of mathematics education* (pp. 556–565). Khon Kaen.

Sommerhoff, D., & Ufer, S. (2019). Acceptance criteria for validating mathematical proofs used by school students, university students, and mathematicians in the context of teaching. *ZDM*, *51*(5), 717–730. https://doi.org/10.1007/S11858-019-01039-7

Sommerhoff, D., Ufer, S., & Kollar, I. (2015). Research on mathematical argumentation: a descriptive review of PME proceedings. In K. Beswick & Muir, T. Wells, J. (Eds.), *Proceedings of 39th Psychology of Mathematics Education conference* (pp. 193–200). PME.

Stokking, K., Leenders, F., Jong, J. de, & van Tartwijk, J. (2003). From student to teacher: Reducing practice shock and early dropout in the teaching profession. *European Journal of Teacher Education*, *26*(3), 329–350. https://doi.org/10.1080/0261976032000128175

Stürmer, K. (2011). *Voraussetzungen für die Entwicklung professioneller Unterrichts-wahrnehmung im Rahmen universitärer Lehrerausbildung* [Dissertation]. Technische Universität München, München.

Stürmer, K., Seidel, T., & Schäfer, S. (2013). Changes in professional vision in the context of practice. *Gruppendynamik Und Organisationsberatung*, *44*(3), 339–355. https://doi.org/10.1007/s11612-013-0216-0

Stylianides, A. J. (2007). Proof and Proving in School Mathematics. *Journal for Research in Mathematics Education*, *38*(3), 289–321.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*(3), 743–762. https://doi.org/10.1037/a0027627

Südkamp, A., Möller, J., & Pohlmann, B. (2008). Der Simulierte Klassenraum: Eine experimentelle Untersuchung zur diagnostischen Kompetenz [The simulated classroom: An experimental study on diagnostic competence]. *Zeitschrift Für Pädagogische Psychologie*, *22*(34), 261–276. https://doi.org/10.1024/1010-0652.22.34.261

Südkamp, A., & Praetorius, A.-K. (Eds.). (2017). *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 94. Diagnostische Kompetenz von Lehrkräften: Theoretische und methodische Weiterentwicklungen*. Waxmann.

Südkamp, A., Praetorius, A.-K., & Spinath, B. (2018). Teachers' judgment accuracy concerning consistent and inconsistent student profiles. *Teaching and Teacher Education*, *76*, 204–213. https://doi.org/10.1016/j.tate.2017.09.016

Tekkumru-Kisa, M., & Stein, M. K. (2017). Designing, facilitating, and scaling-up video-based professional development: Supporting complex forms of teaching in science and mathematics. *International Journal of STEM Education*, *4*(1), 27. https://doi.org/10.1186/s40594-017-0087-y

Thiede, K. W., Brendefur, J. L., Osguthorpe, R. D., Carney, M. B., Bremner, A., Strother, S., Oswalt, S., Snow, J. L., Sutton, J., & Jesse, D. (2015). Can teachers accurately predict student performance? *Teaching and Teacher Education*, *49*, 36–44.

Ufer, S., Heinze, A., & Reiss, K. (2008). Individual predictors of geometrical proof competence. In O. Figueras & A. Sepúlveda (Eds.), *Proceedings of the joint meeting of the 32nd conference of the international group for the psychology of mathematics education* (Vol. 1, pp. 361–368). PME.

Ufer, S., & Neumann, K. (2018). Measuring Competencies. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 433–443). Routledge Taylor & Francis Group.

Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review*, *32*, 100374. https://doi.org/10.1016/j.edurev.2020.100374

Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry: CDASSG Project.* The University of Chicago.

van Es, E. A. (2011). A framework for learning to notice student thinking. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134–151). Routledge.

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. https://doi.org/10.1080/07370008.2017.1317125

van Es, E. A., & Sherin, M. G. (2002). Learning to Notice: Scaffolding new Teachers‘ interpretations of classroom interactions. *Journal of Technology and Teacher Education*, *10*(4), 571–596.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, *24*(2), 244–276. https://doi.org/10.1016/j.tate.2006.11.005

van Es, E. A., Tekkumru-Kisa, M., & Seago, N. (2020). Leveraging the Power of Video for Teacher Learning. In O. Chapman (Ed.), *The international handbook of mathematics teacher education* (pp. 23–54). Brill Sense. https://doi.org/10.1163/9789004418967_002

Villaume, W. A., Berger, B. A., & Barker, B. N. (2006). Learning motivational interviewing: Scripting a virtual patient. *American Journal of Pharmaceutical Education*, *70*(2), 33. https://doi.org/10.5688/aj700233

Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education*, *25*(8), 1051–1060. https://doi.org/10.1016/j.tate.2009.04.002

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner, T., & Jäncke, P. (2004). MEC spatial presence questionnaire (MEC-SPQ): short documentation and instructions for application. *Report to the European Community, Project Presence: MEC (IST-2001-37661)*. http://www.ijk.hmt-hannover.de/presence

Vu, N. V., & Barrows, H. S. (1994). Use of Standardized Patients in Clinical Assessments: Recent Developments and Measurement Findings. *Educational Researcher*, *23*(3), 23–30. https://doi.org/10.3102/0013189x023003023

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Weigand, H.-G., Filler, A., Hölzl, R., Kuntze, S., Ludwig, M., Roth, J., Schmidt-Thieme, B., & Wittmann, G. (2014). *Didaktik der Geometrie für die Sekundarstufe I* (2. Auflage). *Mathematik Primarstufe und Sekundarstufe I + II*. Springer Spektrum. https://doi.org/10.1007/978-3-642-37968-0

Wiepke, A., Richter, E., Zender, R., & Richter, D. (2019). Einsatz von Virtual Reality zum Aufbau von Klassenmanagement-Kompetenzen im Lehramtsstudium. In N. Pinkwart & J. Konert (Eds.), *DELFI 2019. Bonn: Gesellschaft für Informatik e.V.* (pp. 133–144). https://doi.org/10.18420/delfi2019_319

Wildgans-Lang, A., Scheuerer, S., Obersteiner, A., Fischer, F., & Reiss, K. (2020). Analyzing prospective mathematics teachers' diagnostic processes in a simulated environment. *ZDM*, *52*(2), 241–254. https://doi.org/10.1007/s11858-020-01139-9

Wissenschaftsrat. (2014). *Bedeutung und Weiterentwicklung von Simulationenin der Wissenschaft*. Dresden.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225–240.