Diagnosing mathematical argumentation skills: A video-based simulation for pre-service teachers

Elias Codreanu[a], Sina Huber[a], Sarah Reinhold[b], Daniel Sommerhoff[c], Birgit Neuhaus[d], Ralf Schmidmaier[e], Stefan Ufer[c], Tina Seidel[a]

[a] Friedl Schöller Endowed Chair for Educational Psychology, School of Education, Technical University of Munich (TUM)
[b] Assistant Professorship of Technical Education, School of Education, Technical University of Munich (TUM)
[c] Mathematics Education, Faculty of Mathematics, Computer Science and Statistics, LMU Munich
[d] Biology Education, Faculty for Biology, LMU Munich
[e] Institute for Medical Education, University Hospital, LMU Munich

Abstract. Mathematical argumentations and proofs cause difficulties for secondary students (Healy and Hoyles 2000). To facilitate students' understanding of proofs, teachers' diagnostic skills are essential for adapting their teaching to the students' specific needs (Südkamp and Praetorius 2017). We developed a video-based simulation to investigate and promote pre-service teachers' diagnostic skills. Participants face a diagnostic task with short scripted video clips showing students working on a geometry proof with a teacher. Observing student–teacher interactions provides the basis for pre-service teacher participants' diagnosis of students' individual argumentation skills. This simulation is used in a first step to investigate pre-service teachers' diagnostic performance and the quality of their diagnoses, as well as their diagnostic processes. In a second step, the simulation will be expanded to use it as a learning environment to investigate the support of pre-service teachers' diagnostic skills through different kind of scaffolds.

Diagnosing Based on Student Observation

Every day, teachers face a variety of diagnostic situations where they gather information about learning prerequisites, processes, and outcomes of their students (Herppich et al. 2018; Praetorius et al. 2013; Ruiz-Primo and Furtak 2007; Thiede et al. 2015). This information serves as a basis for different pedagogical decisions like lesson planning, adaptive teaching, or grading students (Schrader 2013; Dünnebier et al. 2009; Südkamp et al. 2012; Vogt and Rogalla 2009). In particular, diagnostic decisions are indispensable for the continuous, on-the-fly adaption of teaching to students' prerequisites and ongoing learning processes. Across educational systems, these diagnostic situations arise within the everyday student–teacher interactions that dominate classrooms (Klug et al. 2013; Furtak et al. 2016; Kingston and Nash 2011; Birenbaum et al. 2006). In this context, teachers require professional vision by noticing and interpreting significant features of these classroom situations (Seidel and Stürmer 2014). Within the high density of interactions, they describe, explain, and predict features to come to meaningful conclusions for pedagogical actions.

For pre-service teachers, these high density interactions are oftentimes experienced as overwhelming, since they require the deliberate practice of diagnostic decision making (Levin et al. 2009). Therefore, many pre-service teachers struggle with getting into the profession (Stokking et al. 2003). Nevertheless, diagnostic skills in diagnostic situations in the classroom are rarely taught in teacher education. Initially, higher education is focused on conveying basic principles and conceptual knowledge, often separated in different study fields of content, pedagogical-content, and pedagogy and psychology. Given these structures, it is often unclear how these aspects of professional knowledge are related to specific diagnostic situations in classrooms (Alles et al. 2019). Therefore, new ways of supporting the acquisition of crucial skills like diagnostic skills are needed to prepare pre-service teachers to make reasonable diagnostic decisions before entering the classroom. Additionally, little is known about the processes involved in diagnostic decision making and differences in these processes along the learning trajectory (Herppich et al. 2018). Insights into these processes may be promising to identify characteristics for targeted interventions along this learning trajectory.

Simulation as a Model of Reality

An environment to investigate and promote pre-service teachers' diagnostic skills should implement two aspects: First, following a practice-oriented approach, it should represent practice in an authentic way to engage pre-service teachers to get involved in the actual task (Schubert et al. 2001). This allows pre-service teachers to transfer their behavior from that environment to real-world teaching situations (Frank 2015). Second, reality must be decomposed and simplified in a way that it enables pre-service teachers to focus on particular aspects of classroom situations (Grossman et al. 2009). Therefore, decompositions of practices provide features that make diagnostic decision-making more accessible to pre-service teachers than in real-world classroom situations.

Due to its strengths regarding both aspects, video is becoming a frequently used medium in professional teacher education (Kang and van Es 2018; Gaudin and Chaliès 2015). Although videos show only one captured perspective of a classroom situation and, thus, have limitations in conveying the contextual background of the situation, it can give authentic insights in different teaching and learning situations (Blomberg et al. 2013). However, by taking a certain perspective, videos can also direct the observer's attention to significant features of the situation by so-called cues. Following the idea of a decomposition of practice in the case of diagnosing students' skills based on student observation in classrooms, videos should incorporate everyday student–teacher interactions with most relevant cues for diagnosis but only a few less relevant cues that call for teachers' attention in real-world classrooms. For diagnosing mathematical argumentation skills from a mathematics educational perspective, students' statements allowing conclusions, for example, of their understanding and knowledge of correct mathematical proof procedures are most relevant. General aspects like the situational motivation, for example, can be considered less relevant for this diagnosis. The reduction of the amount of the less relevant cues increases the capacity of teachers for deliberate action. The special form of scripted video allows for further targeted manipulation of these segments (Piwowar et al. 2017).

Not just the design of scripted videos but also their embedding in a simulated environment influences learning grounded in practice. Decomposing the situation by dividing a scripted video in particular scenes gives the opportunity to slow down the actual situation and to thereby reduce the density of interactions. Through decomposition of a situation, simulations provide researchers with insights into the processes and allow to gather data for further analyses of diagnostic skills. Such results may then help to develop evidence-based interventions for further training.

Mathematical Argumentation Skills

Mathematics is a relevant subject for diagnostic situations including student–teacher interactions because the traditional initiation-response-feedback discourse of teaching is the prevalent form of teacher-student dialogue (Lipowsky et al. 2009). In mathematics, as a proving science, handling mathematical argumentation and proofs as a special form of argumentation fulfilling strict standards (Stylianides 2007), is a crucial learning activity. Thus, mastering this activity is a central learning goal in many secondary school systems (Kultusministerkonferenz 2012). However, empirical studies have repeatedly shown that students have substantial problems when attempting to construct a mathematical proof (Healy and Hoyles 2000; Harel and Sowder 1998). In particular, the success of mathematical proving activities depends on different individual prerequisites, identified by studies as predictive for the performance in handling proofs (Sommerhoff et al. 2015; Schoenfeld 1992). These can be used in the diagnostic situation as an indication to diagnose the students' skills in handling argumentations and proofs. Based on their research, Ufer et al. (2008) and Sommerhoff et al. (2015) emphasize students' *mathematical knowledge base*, their *methodological knowledge*, and thier *problem-solving strategies* as important prerequisites. However, these three prerequisites can be divided in more specifical ways so that they can be used for the diagnostic process, as described in the following.

The mathematical knowledge base comprises three different sub-concepts (Weigand et al. 2014). First, knowledge of *concept properties* contains knowledge of features and terms, like features of the diagonals of parallelograms. The second sub-concept called *concept scope* considers the entirety of representatives of a mathematical term. This includes, for example, the knowledge that a square is also a representative of the term parallelogram. Third, the *concept network* refers to knowledge about the relationship of a concept to other concepts. Likewise, methodological knowledge that is the knowledge about the nature of proofs, and their use within mathematics and socio-mathematical norms regarding proofs can be divided into at least three components (Heinze and Reiss 2003): Knowledge of *proof scheme* contains knowledge about acceptable types of inferences in a proof. *Proof structure*, in contrast, refers to the overall logical structure of a proof, for example starting with the assumptions and ending with the assertion. Finally, *chain of conclusion* refers to the logic of the arrangement of individual arguments within the proof. With respect to problem-solving strategies, this research project focuses on two different aspects. First, *heuristic strategies* that jelp to solve a given problem task by reorganizing the task and changing the way of looking at it. Second, *metacognitive strategies* that allow the individual control of a problem-solving process by strategies such as monitoring and assessing the progress of the problem-solving process as well as drawing conclusions for one's own action.

Prior research indicates that students typically differ widely regarding each of these eight aspects, resulting in a range of difficulties when attempting mathematical proofs (Reiss and Ufer 2009). It is a difficult task for a teacher to diagnose based on brief student–teacher interactions and possibly a brief look at students' notes, what the reasons for students' difficulties are and, hence, which kind of teacher support will help each individual student.

Guiding Questions in Designing the Simulation

Both, measuring and supporting teachers' diagnostic skills based on simulations require high standards regarding the simulations' authenticity and the content of the embedded videos. The development of the video-based simulation presented in this chapter was thus guided by the following questions:

1) To what extent can we authentically represent a diagnostic situation within student–teacher interactions around mathematical argumentation in a scripted video-based simulation?
2) To what extent can the decomposition of the diagnostic situation in the video-based simulation reveal insights into the diagnostic processes of the participants?

Conception of the Scripted Videos

To simulate a setting to diagnose individual students' mathematical argumentation skills in a simulated classroom situation, we developed scripted videos with small group student–teacher interactions following Dieker et al. (2009)'s recommendations. In the first step, we identified relevant content of the situation with essential features (*selection of practice*). Second, we developed a contextual frame for all of the recordings as well as detailed scripts for the particular scenes (*vignette script development*). Third, we collected video footage and edited it to create a representation of teaching practice (*video production*).

Selection of the Practice. Based on their importance, we decided to focus on three students' individual prerequisites for the video production that are predictive for their performance in handling proofs (Ufer et al. 2008): (a) students' *mathematical knowledge base*, (b) their *methodological knowledge*, and (c) their *problem-solving strategies*. All three prerequisites have been shown to affect students' skills to handle geometrical proofs and could be portrayed in brief video clips. We considered the three sub-concepts of the mathematical knowledge base, the three sub-concepts of the methodological knowledge, and the two aspects of problem-solving strategies as a theoretical fundament to design the student profiles.

In the following step, we outlined four student profiles that vary in the availability of the predictive prerequisites of students' skills to handle argumentations and proofs (in total eight aspects). To determine different levels of availability of the predictors in the student profiles, van Hiele's model for describing the development of geometric thinking of children provided valuable additional guidance (Usiskin 1982). According to this model, students on a first level recognize and judge figures by their appearance. A student on the second level can identify properties of figures, while a student on the third level can already follow simple deductions. Only at level four, a student has acquired enough understanding to meaningfully construct proofs. We concentrated on these four levels and specified the student profiles on the basis of their knowledge and abilities according to the eight predictive concepts. For example, profile one and two know little about acceptable of inferences (*proof scheme*), whereas profile three is mostly and profile four totally aware of this (see Fig. 1).

Mathematical Knowledge Base				
Concept Properties	★	★ ★	★ ★ ★ ★	★ ★ ★ ★
Concept Scope	★	★ ★	★ ★ ★ ★	★ ★ ★ ★
Concept Network	★	★ ★	⃠	★ ★ ★ ★
Methodological Knowledge				
Proof Scheme	★	★	★ ★ ★	★ ★ ★ ★
Proof Structure	★	★ ★	★ ★ ★	★ ★ ★
Chain of Conclusion	★	★	★ ★	★ ★ ★
Problem-Solving Strategies				
Heuristic Strategies	★ ★ ★	★ ★ ★ ★	★ ★ ★ ★	★ ★ ★ ★
Metacognitive Strategies	★	★ ★	★ ★ ★	★ ★ ★ ★

⃠ not observable ★ mostly not available ★ ★ insufficiently available ★ ★ ★ mostly available ★ ★ ★ ★ available

Figure 1. Student profiles. Four student profiles with their specific predictive prerequisites for the performance in handling proofs.

To ensure that the students' remain comparable, all students were working on the same geometry proof task in the video clips: They had to prove that opposite sides of a parallelogram are of equal length, based on the information that pairs of sides of a parallelogram are parallel. Beginning to handle proofs implies that the students don't pay a lot of attention to norms and standards of proofs on an abstract level. Thus, it is not expected that all aspects, especially for methodological knowledge, become important in the proving process. Likewise, the conceptual network may not be required during the proving process for all four students. However, as the task allows to portray a wide range of different ways to process the task, it was chosen as a basis for the simulation.

Vignette Script Development for Staged Videos. Students' working time on the geometrical proof task was split into eight smaller video scenes, each with a length of approximately one minute. Thus, the number of scenes was the same for all students and was enough to provide participants with the opportunity to observe each students multiple times. The scripts for these scenes contain detailed dialogues between the teacher and one of the students, as well as blueprints of the students' sketches and other records in the exercise books. The teacher's input in the scenes was reduced to a minimum, so the teacher's function focuses on eliciting students talk about their thoughts. Thus, typical questions and requests of the teacher were "What do you mean by that?" or "Can you explain what you have done here?"

Considering the eight predictive prerequisites, answers and statements given by the students were generated according to their profile. Besides the verbal student–teacher interaction, cues for the aspects could also be found in the sketches and records of the students. With this, we controlled the distribution of the cues that were relevant for the diagnosis. During the creation of the scenes, the attribution was continuously reviewed in an internal review process, to provide salient cues for the provided prerequisites in the video scenes. These cues were distributed as evenly as possible over the eight scenes in order to portray an authentic conversation. This resulted in a distribution, which allowed at least one (often more than one) salient cue for each aspect occurring no later than in the fourth video scene.

Production of Staged Videos. The videotaping was completed with one trained teacher and four student volunteers from the eighth grade. Teachers and students were provided with the scripts prior to filming and were given time and guidance to familiarize themselves with their role, the script, and each other. While the scene of the teacher and one student was filmed, the other students practiced their next scene with an instructor of the video production team. During the shot, the actors followed the scripts with as much fidelity as natural behavior allowed in that moment. The research team ensured that the main cues of the scripts were successfully captured on video. To capture both, the verbal student–teacher interaction and the written records, two different camera perspectives were used at the same time: One from the front showing the conversation, one from above showing the exercise book of the student. In the editing process, the scenes were cut from the two camera perspectives according to the appropriate camera angle. After the production, the final video scenes were reviewed by two independent researchers concerning the perceptibility of the cues initially implemented in the scripts. In a subsequent consensus process based on the final video scenes, the characteristics of the profiles according to the predictive prerequisites were determined in four ordinal categories (see Fig. 1).

Design of the Simulation

The presented video-based simulation has an underlying structure consisting of four main parts (see Fig. 2). It starts by familiarizing the participants with the situation the simulation depicts, an assistant teacher observing student–teacher interactions during a student exercise on a geometric proof (*diagnostic situation*). Before participants start working with the tool, they are introduced into the task to be accomplished in this diagnostic situation: the simulated teacher asks them to assess the students' mathematical argumentation skills so that he can choose tasks for individual learning support in the subsequent class based on the participants' observations (*diagnostic task*). After that, participants can work independently in the simulated classroom situation to gather information about the students by watching video scenes and taking notes (*diagnostic process*). These notes form the basis for the final diagnosis of each student, which participants formulate in the last section of the tool to provide the simulated teacher a basis for his further lesson planning (*diagnostic outcome*).

Diagnostic Situation. The situation chosen for this simulation is an everyday classroom situation in mathematical lessons (Lipowsky et al. 2009). Students are independently working on a task, in this case, a geometry proof, while the teacher walks from student to student to monitor and support their progress in short student–teacher interactions. At the beginning of the simulation, participants are familiarized with their role in this simulation: they are observing teacher and students in their interactions in the role of a pre-service intern. Besides this, they receive information about the overall topic, prior lessons, and learning context to get acquainted with the classroom situation as well as with the content discussed in the lesson. Taking over the role of an intern is familiar to the participants (pre-service teachers), so it presumably does not require a big effort to put themselves in this role. Thus, the scenario is likely to support the immersion into the simulation (Slater and Wilbur 1997). Furthermore, when acting in real classrooms, interns face similar challenges and possibilities as those embedded in the diagnostic process later in the simulated situation. This parallelism between the role in real-world situations and in the simulated diagnostic situation is expected to lead to a higher authenticity of the learning environment (Schubert et al. 2001).

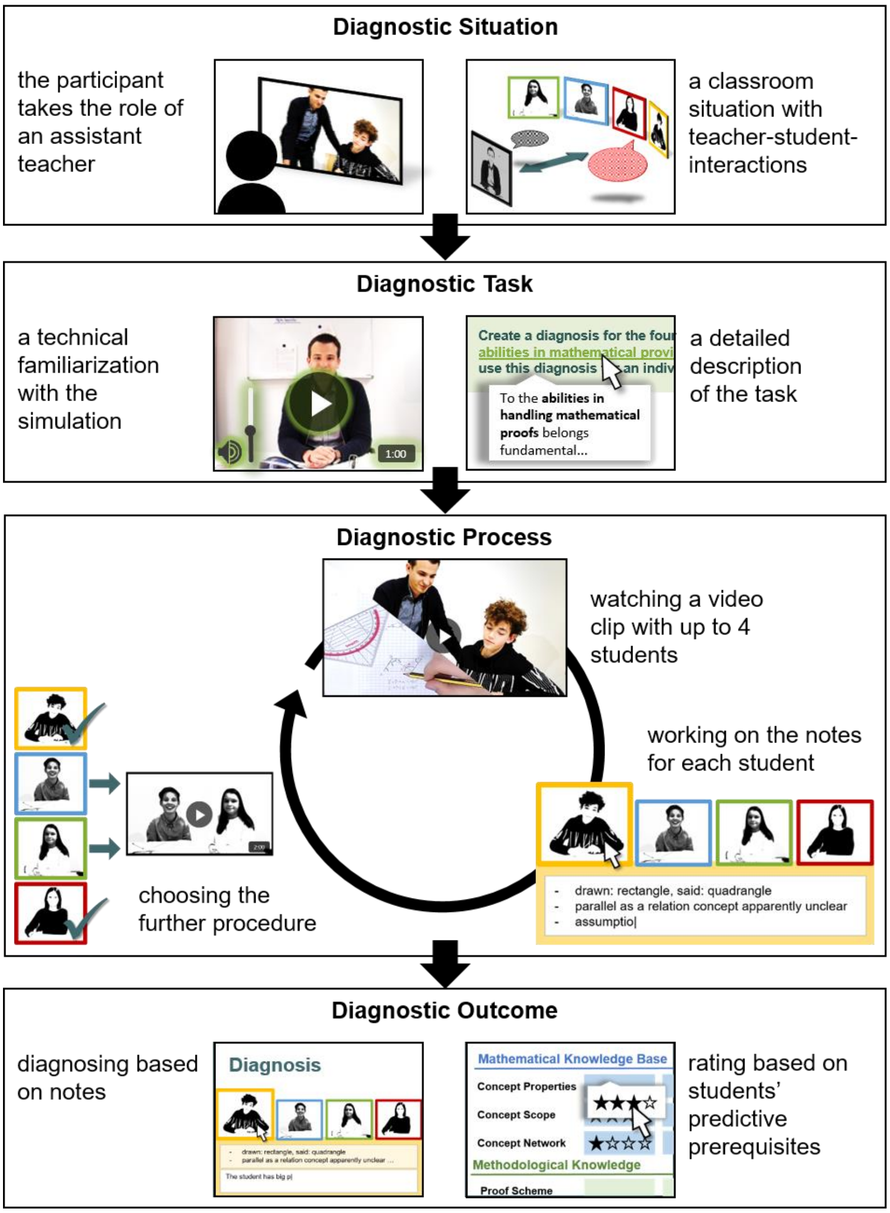

Figure 2. Design of the video-based simulation.

After the introduction to their role, participants get information about the different steps a teacher considers in the preparation of a lesson. Information about the prior knowledge of the whole class and the subject matter of the previous lessons of the class is provided. In addition, participants get the possibility to familiarize themselves with the proof task of the upcoming lesson. To cushion a potential lack of curricular knowledge in pre-service teachers at an early stage of education (Shulman 1986), participants have the choice to look into the main topics of the curriculum relevant for the upcoming classroom situation.

Diagnostic Task. After familiarizing themselves with the diagnostic situation, the simulated teacher presents the diagnostic task to the participants: They are asked to diagnose the level of handling geometrical proofs of four specific students in order to give the simulated teacher ideas for individual student support in the subsequent remedial lesson.

In the presentation of the diagnostic task, we strive for two goals: participants should know both, the specific task in detail and its embedding in the simulation. In a first step, the presentation of the task in a short video clip familiarizes participants with the technical side of the simulation. For example, participants have the possibility to play and pause but not to rewind the video to stay close to reality. With the familiarization, we minimize technical complications later in the diagnostic process. Additionally, participants get to know the teacher they are accompanying in the simulation later on. In the second step, the diagnostic task is described in detail focusing on the following two aspects: (a) *who* is the diagnosis for and *how* should it, therefore, look like and (b) what is the diagnosis' *purpose* and which *components* should it, therefore, entail? Considering that the participants most likely have little experience with diagnosing student skills and abilities and the wording in this field, we deliver a detailed description of the task to be worked on in the following diagnostic process. To go into detail regarding aspect (a), it is pointed out that a diagnosis should include descriptions, explanations, predictions, and a decision (Blömeke et al. 2015; Seidel and Stürmer 2014). To meet aspect (b), a description for the ability to handle geometrical proofs is given containing the predictive prerequisites implemented in the video clips (see Fig.1).

Diagnostic Process. In the diagnostic process, the participants observe four students, which simulates a reduced classroom setting. The process is divided into several cycles. Each cycle starts with watching one video clip containing scenes with student–teacher interactions between one of the students and the teacher. In the first cycle, participants observe all four students in a row. While observing the students, participants have the possibility to take notes. By clicking on the picture of each student, they can enter the notes for each student in the respective text box. This provides the opportunity to work individually on the notes for each student. At the end of a cycle, participants must choose to continue the procedure. They can decide in case of each individual student whether they want to observe more interactions with this student and hence, run through another cycle of the diagnostic process for them, or conclude the diagnostic process for this particular student. Therefore, if a participant decides, for example, to continue observing two of the four students, the next cycle covers only the further processing of the proof task of these two students. The text box can only be opened for the two remaining students. Again, participants then decide whether to continue to observe each of the remaining students in a third cycle. This continues until participants choose to conclude the observation process for all four students or after a maximum of eight cycles.

In the first cycle, participants start with an empty text box for their note taking. In the following cycles, the notes from the previous cycles are already displayed in the text box, so that participants can continue working on the notes based on the previous notes. These notes serve as individual support throughout the entire diagnostic process. However, the maximum number of scenes participants can watch is limited to 20. Thus, they have to deal with portioning the scenes on the four students. This makes it possible to measure the efficiency of the diagnostic process of the participants.

Diagnostic Outcome. Finally, after the participants complete the diagnostic process for all students, they have to submit their diagnosis in two different ways. First, they are asked to formulate a diagnosis for each student in an open text box. Therefore, their notes from the diagnostic process are shown above the text box. The participants can copy parts of the notes, summarize their points, or use the notes as an aid to remembering the situations in the video clips. Analogous to the note page, they can work on the four diagnoses in any order. Second, participants are asked to assess the students regarding their mathematical knowledge base, their methodological knowledge, and their problem-solving strategies. Participants have to rate the availability of each of the eight predictive prerequisites on a four-point Likert-scale, which allows assessing the diagnostic outcome in two different ways for a more differentiated consideration (see Fig. 1). Additionally, participants are asked to rank the student profiles according to their level of mathematical argumentation skills from weakest to strongest.

Discussion and Outlook

The video-based simulation developed in this project provides an innovative way to investigate and promote pre-service teachers' diagnostic skills regarding students' mathematical argumentation skills. The purpose-built conceptualisation of scripted videos as well as the specific design of the simulation give an indication that the environment represents practice authentically and allows participants to immerse in the situation. This supports the transfer of the behavior to real-world situations (Frank 2015). The geometric task and the student–teacher interactions around this task as a facet of practice chosen for the scripted videos resemble situations found in real-world mathematics classrooms. Important prerequisites of students in terms of mathematical knowledge base and methodological knowledge are contained in the four meticulously designed student profiles (Ufer et al. 2008). Finally, the videos were filmed with student volunteers, to enrich the actual script with their natural behavior. These prerequisites for their argumentation skills were implemented considering the authentic representation of the situation. In the simulation itself, we separated the content-related and technical familiarization from the part where the participants actually work on the given task. Hence, all information required to work undisturbed by interruptions and all additional instructions needed for the simulated situation is provided before the actual diagnostic process starts. This leads to the possibility to immerse deeper into the situation. In empirical analyses, expert teachers' as well as pre-service teachers' ratings of authenticity and immersion with regard to scripted videos and the simulation as a whole will evaluate if participants experience this as a convincing representation of real-world classrooms. We will involve expert teachers because of their wealth of experience in classroom situations and beside this novice teachers as they belong to the targeted group the simulation was developed for.

The specific conception and design of scripted videos and simulation embedding both contribute to the decomposition of practice in a way to allow for extraction of features regarding the diagnostic process of the participants. The scripted videos show only four simulated students whose profiles only differ in important prerequisites of successfully handling geometrical proofs. This makes it easier for participants to focus on and distinguish between students than in a twenty-plus student classroom with a continuous and diverse composition of those prerequisites. The deliberate absence of classroom management issues such as handling disturbances gives participants the opportunity to concentrate on the more relevant instead of less relevant cues in their diagnostic process. Adding time to the participants' observation by writing notes slow down the ongoing classroom actions. While a real-world classroom does not provide specific times to take notes on what teachers notice and interpret from this, the simulation addresses these processes. Furthermore, the instruction to describe observations in the notes as well as to interpret them helps to deal in detail with what has been observed. This way, complexity of the situation is reduced and important mental steps are recorded. Additionally, the notes give insight into the use and performance of the participants reasoning (Seidel and Stürmer 2014). These data serve as a basis to find key features in the diagnostic process to develop targeted support in the simulation. Thus, analyzing pre-service teachers' diagnostic process should reveal differences where instructional support like expert modeling and self-explanatory prompts can be set.

We expect further findings of the processes and the variables that have an influence on the performance in the simulation by investigating individual prerequisites of the participants like their knowledge base or interest and self-concept. Regarding these findings, the simulation will be expanded from a tool to assess diagnostic skills to a tool that also has the feature to foster those skills.

Acknowledgements

The research presented in this contribution was funded by a grant of the Deutsche Forschungsgemeinschaft (DFG) to Tina Seidel, Stefan Ufer, Birgit Neuhaus, and Ralf Schmidmaier (Grant number 2385, NE 1397/11-1).

References

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*, 87–112. doi:10.1007/s12186-018-9211-5

Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgway, J., et al. (2006). A learning integrated assessment system. *Educational Research Review, 1*, 61–67. doi:10.1016/j.edurev.2006.01.001

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. J. (2015). Beyond dichotomies. *Zeitschrift für Psychologie, 223*, 3–13. doi:10.1027/2151-2604/a000194

Dieker, L. A., Lane, H. B., Allsopp, D. H., O'Brien, C., Butler, T. W., Kyger, M., et al. (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*, 180–196. doi:10.1177/0888406409334202

Dünnebier, K., Gräsel, C., & Krolak-Schwerdt, S. (2009). Urteilsverzerrungen in der schulischen Leistungsbeurteilung. *Zeitschrift für Pädagogische Psychologie, 23*, 187–195. doi:10.1024/1010-0652.23.34.187

Frank, B. (2015). *Presence messen in laborbasierter Forschung mit Mikrowelten: Entwicklung und erste Validierung eines Fragebogens zur Messung von Presence*. Masterarbeit, Duisburg. Accessed 16 August 2017.

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., León, V. de, Morrison, D., et al. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science, 44*, 267–291. doi: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. doi:10.1016/j.edurev.2015.06.001

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.

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

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

Heinze, A., & Reiss, K. (2003). Reasoning and proof: methodological knowledge as a component of proof competence. www.lettredelapreuve.itlCERME3PapersiHeinze-paperl.pdf. Accessed 8 August 2017.

Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., et al. (2018). Teachers' assessment competence: integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education, 76*, 181–193. doi:10.1016/j.tate.2017.12.001

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

Kingston, N., & Nash, B. (2011). Formative assessment: a meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*, 28–37. doi:10.1111/j.1745-3992.2011.00220.x

Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of teachers: A process model that accounts for diagnosing learning behavior tested by means of a case scenario. *Teaching and Teacher Education, 30*, 38–46. doi:10.1016/j.tate.2012.10.004

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

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

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction, 19*, 527–537. doi:10.1016/j.learninstruc.2008.11.001

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*, 369–384. doi:10.1080/19415257.2017.1316299

Praetorius, A.-K., Berner, V.-D., Zeinz, H., Scheunpflug, A., & Dresel, M. (2013). Judgment confidence and judgment accuracy of teachers in judging self-concepts of students. *The Journal of Educational Research, 106*, 64–76. doi:10.1080/00220671.2012.667010

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*, 57–84. doi:10.1002/tea.20163

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). New York: Simon & Schuster.

Schrader, F.-W. (2013). Diagnostische Kompetenz von Lehrpersonen. *Beiträge zur Lehrerbildung, 31*(2), 154–165.

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

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

Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher, 15*(2), 4–14.

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., 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). Hobart: 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*, 329–350. doi:10.1080/0261976032000128175

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*, 743–762. doi:10.1037/a0027627

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

Thiede, K. W., Brendefur, J. L., Osguthorpe, R. D., Carney, M. B., Bremner, A., Strother, S., et al. (2015). Can teachers accurately predict student performance? *Teaching and Teacher Education, 49*, 36–44. doi:10.1016/j.tate.2015.01.012

Ufer, S., Heinze, A., & Reiss, K. (2008). Individual predictors of geometrical proof competence. *PME 32 and PME-NA XXX, 1*(4), 361–368.

Usiskin, Z. (1982). *Van Hiele levels and achievement in secondary school geometry: CDASSG Project*. Accessed 3 July 2017.

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

Weigand, H.-G., Filler, A., Hölzl, R., Kuntze, S., Ludwig, M., Roth, J., et al. (2014). *Didaktik der Geometrie für die Sekundarstufe I* (2nd ed., Mathematik Primarstufe und Sekundarstufe I + II). Berlin: Springer Spektrum.