TEACHERS' COMPETENCE OF ANALYSING THE USE OF MULTIPLE REPRESENTATIONS IN MATHEMATICS CLASSROOM SITUATIONS AND ITS ASSESSMENT IN A VIGNETTE-BASED TEST

Dissertation
zur Erlangung des Grades
eines Doktors der Philosophie (Dr. phil.)

der Pädagogischen Hochschule Ludwigsburg

vorgelegt von Marita Eva Friesen, geb. Hettinger,
aus Bad Mergentheim

Ludwigsburg
2017

Erstgutachter: Prof. Dr. Sebastian Kuntze

Zweitgutachter: Prof. Dr. Markus Vogel

Abschluss der mündlichen Prüfung: 10. Mai 2017

Danksagung

Mein Dank gilt an erster Stelle meinen beiden Betreuern und Gutachtern Prof. Dr. Sebastian Kuntze und Prof. Dr. Markus Vogel für die Möglichkeit, als abgeordnete Lehrerin im Rahmen des Forschungs- und Nachwuchskollegs EKoL (Effektive Kompetenzdiagnose in der Lehrerbildung) promovieren zu können. Ich danke Dir, Sebastian, vor allem für die Zeit, die Du Dir stets genommen hast, für alle Diskussionen und konstruktive Kritik, für Deine Förderung auf dem Weg in die mathematikdidaktische Community und die Unterstützung in wirklich allen Lagen. Danke für alles, was Du oftmals als selbstverständlich betrachtet hast und was für mich immer besonders war. Ich danke Dir, Markus, dass Du in unserem Projekt immer für mich als Ansprechpartner da warst, danke für alle Gespräche und Anregungen, danke für Deine Ermutigung und Begleitung von Heidelberger Seite aus. Ich danke allen Kollegiatinnen und Kollegiaten sowie den Betreuerinnen und Betreuern aus der ersten und zweiten Phase von EKoL für die Zusammenarbeit und Unterstützung innerhalb des Projektteams, insbesondere Prof. Dr. Markus Rehm als Leiter unseres Kollegs und Prof. Dr. Christoph Randler, der mich überhaupt erst auf das Projekt EKoL aufmerksam gemacht hat, sowie den beiden Ludwigsburger Kollegiaten Friederike Straub und Dr. Sebastian Goreth. Danke an Julia Ollesch für die schöne Zeit in unserem Matheteam und für die vielen gemeinsam verbrachten Tagungen. Ich danke allen Kolleginnen und Kollegen von der Pädagogischen Hochschule Ludwigsburg, die mich auf meinem Weg begleitet und unterstützt haben. Ein besonderer Dank gilt meinen wunderbaren Bürokollegen Dr. Christoph Till, Achim Schiller und Pelagia Papadopoulou-Tzaki, die immer für mich da waren. Ganz besonders bedanken möchte ich mich auch bei Prof. Dr. Anika Dreher und Dr. Ute Sproesser dafür, dass sie mich als akademische Schwestern unter ihre Fittiche genommen haben. Danke Dir, liebe Anni, für Deine Zeit, Hilfe und Unterstützung besonders am Anfang der Promotionszeit. Danken möchte ich auch den Kolleginnen und Kollegen aus dem Fach Mathematik des Staatlichen Seminars für Didaktik und Lehrerbildung in Ludwigsburg sowie Judith Geörg als Seminarleiterin. Vielen Dank für die Kooperation mit dem Projekt EKoL, für die Unterstützung bei der Entwicklung von Unterrichtsvignetten, für die Ermöglichung und Organisation von Befragungen am Seminar und an Realschulen. Besonders möchte ich mich an dieser Stelle bei Reiner Mecherlein bedanken, für seine Zeit, sein besonderes Interesse an unserer

Projektarbeit, für die gemeinsam durchgeführten Veranstaltungen am Seminar Ludwigsburg und für den stets bereichernden Austausch. Ein großes Dankeschön gilt meiner Kollegin und Freundin Christina Vierling, den Kolleginnen und Kollegen sowie Schülerinnen und Schüler der Realschule Boxberg und Andrea Fürle als Schulleiterin für die wunderbare Unterstützung. Ein ganz besonderes Dankeschön geht an meine Freundin Juliana Egete, die mit ihrem Talent und ihrer Kreativität die Figuren in den Comicvignetten zum Leben erweckt hat. Danke an Prof. Dr. Steffen Schaal, Prof. Dr. Laura Martignon und Prof. Dr. Orly Buchbinder für die Abnahme der mündlichen Prüfung. Bei Prof. Dr. Orly Buchbinder möchte ich mich ganz besonders dafür bedanken, dass sie die weite Anreise über den Atlantik auf sich genommen hat, um bei meiner Disputation dabei zu sein. Mein besonderer Dank gilt auch Elisabeth Grözinger für ihre wunderbare Arbeit als Tutorin. Ich danke allen Studierenden, Anwärterinnen und Anwärtern sowie den Lehrkräften, die sich bereit erklärt haben, an der Studie teilzunehmen. Bei meiner Familie und meinen Freunden möchte ich mich für die Unterstützung während der Promotionsphase ganz besonders bedanken. Ich bedanke mich von ganzem Herzen bei meinem Mann Alexander Friesen, der all die Zeit an meiner Seite war und mich stets dazu ermutigt hat, meine Vorhaben und Ideen umzusetzen.

Contents

1 Introduction

Whereas investigating teachers' professional knowledge and beliefs has a long tradition in mathematics education research (e.g., Shulman, 1986; Ball, Thames & Phelps, 2008; Kersting, 2008; Leder, Pehkonen & Törner, 2002), approaches related to the notion of *competence* form a more recent topic (e.g., Kunter et al., 2013; Kaiser et al., 2015). Although quite different perspectives on the concept of competence exist (e.g., Blömeke, Gustafsson & Shavelson, 2015), there is agreement that the competence approach aims at identifying the characteristics and qualities teachers need in order to meet the demands of their profession (Baumert & Kunter, 2013) and that those characteristics can be taught and learned in the framework of professional development (Weinert, 2001a, b). However, there is no "competence" per se and the definition of any competence construct requires identifying a relevant context, i.e. a content domain or a type or range of situations (Hartig, 2008). The starting point to define profession-related competences for teachers are thus the professional demands and tasks that teachers have to master in order to create effective teaching and learning situations in the classroom and enable students to achieve their learning objectives (Lindmeier, 2011; Koeppen et al., 2008; Baumert & Kunter, 2013). For mathematics teachers, demands related to the *use of multiple representations*, such as selecting representations for particular purposes were identified as core tasks of teaching (Hill, Schilling & Ball, 2004; Ball, Thames & Phelps, 2008). As numerous studies in this context have found changes between different representations of a mathematical object to be a major threshold for students' learning (Goldin & Shteingold, 2001; Ainsworth, 2006; Duval, 2006), special attention and help from the teacher is needed to support these cognitively complex processes and to avoid serious learning difficulties (e.g., Dreher & Kuntze, 2015a). In order to meet such professional requirements related to the use of multiple representations, mathematics teachers need a certain awareness as well as corresponding knowledge on the complex demands students have to master when dealing with multiple representations (Duval, 2006; Mitchell, Charalambous & Hill, 2014; Dreher & Kuntze, 2015a, b). Furthermore, they have to be able to *analyse classroom situations* in order to link such knowledge to relevant observations regarding the use of multiple representations, so that changes between representations can be identified and interpreted as potentially problematic for students' understanding (Friesen, Dreher & Kuntze, 2015; Friesen & Kuntze, 2016).

There is a broad consensus that teachers' analysing of classroom situations is an important prerequisite for providing students with adequate learning situations and learning support (e.g., Sherin, Jacobs & Philipps, 2011; Schoenfeld, 2011; Santagata & Yeh, 2016). Reviewing current frameworks of professional competence (e.g., Baumert & Kunter, 2013; Kaiser et al., 2015) reveals, however, that teachers' analysing of classroom situations regarding changes between multiple representations and related demands for students have so far not been sufficiently accounted for.

Consequently, the study presented here focuses on changes between multiple representations and aims at conceptualising teachers' analysing regarding the use of multiple representations in classroom situations as an important profession-related competence of mathematics teachers. As empirical evidence regarding the assessment of such a competence is still scarce, this study also aims at contributing to methodological questions related to the assessment of teachers' competence of analysing the use of multiple representations.

Therefore, a vignette-based instrument was developed, which comprises of six classroom situations in which dealing with changes between multiple representations plays a key role. Building on research by Dreher & Kuntze (e.g., 2015a), the classroom vignettes were situated in the context of learning fractions, as the integration of multiple representations is regarded as particularly significant for students' understanding in this content domain (e.g., Ball, 1993a; Charalambous & Pitta-Pantazi, 2007). Whereas it is widely acknowledged that vignette-based test instruments are particularly suitable to account for the context-specifity of competence constructs (Kaiser et al., 2015; Blömeke, Gustafsson & Shavelson, 2015), there is hardly any empirical evidence regarding the impact of different *vignette formats* and different *question formats* on teachers' analysing of classroom situations. Such item characteristics might, however, affect item difficulty and could thus cause limitations for the interpretation of assessment results (e.g., Hartig, 2008). It is therefore of particular importance to explore the role of item characteristics for the participants' analysing regarding the use of multiple representations.

With the aim to fill this gap in research, each of the six classroom situations was implemented in the format text, comic and video using a multiple matrix design with test booklets. Open-ended questions as well as rating-scale items were applied to each vignette in order to explore the role of such item characteristics. The test instrument was administered to N = 298 student teachers, pre-service teachers and in-service teachers. With the aim to account for the individual participants' competence of analysing as well as for relevant item characteristics, Rasch models were used to analyse the data (Bond & Fox, 2015).

Chapter two will outline the essential role of multiple representations for learning mathematics (section 2.1) and derive corresponding professional requirements for mathematics teachers (section 2.2). The state of research on corresponding professional knowledge (section 2.3) and teachers' analysing (section 2.4) of classroom situations regarding the use of multiple representations will be presented. In the next section, concepts of competence related to this study will be introduced (section 2.5). The competence of analysing regarding the use of multiple representations as used in this study will be defined (section 2.6), before different approaches to competence assessment will be described (section 2.7). *Chapter three* will present the research interest and the research questions of this study. Methods, design and the sample of this study will be described in *chapter four* before the results are presented in *chapter five.* In the final *chapter six*, the findings of the study will be summarised according to the research questions. Limitations of the study will be outlined. Conclusions and implications for further research will be described in the last section (section 6.6).

2 Theoretical background

The aim of this study is to conceptualise teachers' analysing of classroom situations regarding the use of multiple representations as a key aspect of mathematics teachers' professional competence and to develop a vignette-based test in order to assess this competence. The following chapters will accordingly provide insight into relevant theoretical concepts and the state of research related to this study.

2.1 The role of multiple representations for learning mathematics

The use of multiple external representations to support learning is widespread in traditional classroom settings as well as in computer-based environments (e.g., Lesh, Post & Behr, 1987; Ainsworth, Bibby & Wood, 2002; Ainsworth, 2006; Mayer, 2014). Most learning situations involve words and graphics, be it in the form of spoken language or printed texts, illustrations, graphs, diagrams, pictures or animations. Education research and cognitive science provide various explanations for the benefits of learning with multiple representations. Bruner (1966), for example, states that in relation to different subject matters, different ages and different learning styles among learners, any idea or body of knowledge can be presented in a form simple enough so that any particular learner can understand it: by action (enactive representation), by images or graphics (iconic representation) or by symbolic or logical propositions (symbolic representation). Ainsworth (1999, 2006) claims that by combining different representations, learners might benefit from the advantages of each of the individual representations and can choose the representation that best suits their individual needs for learning, provides the most efficient structure required for problem solving or enhances the best strategy. Cognitive flexibility theory (e.g., Spiro & Jehng, 1990) explains the benefits of multiple representations with the fact that constructing multiple perspectives on a subject and switching between them is fundamental for successful learning. Dual coding theory (e.g., Clark & Paivio, 1991) and the cognitive theory of multimedia learning (e.g., Mayer & Sims, 1994; Mayer, 2014) assume that by combining different representations endowed with different computational properties, learners can benefit from strengths of particular representations and balance out weaknesses of others.

For the learning of mathematics, however, the role of multiple representations goes beyond the fact that different representations can complement each other with regard to their computational efficiency or with respect to learner preferences. The reason for this lies in a characteristic which is specific to the domain of mathematics and puts dealing with multiple representations at the heart of any mathematical activity: In contrast to objects in many other domains of scientific knowledge, mathematical objects are abstract and can therefore not be directly perceived or observed with instruments (Duval, 2006). Consequently, the only way to gain access to those abstract mathematical objects and deal with them is using representations that can stand for them in many ways (Goldin, 2008). This makes mathematics the scientific domain in which the largest amount of representations can be found (Duval, 2006) and hence dealing with multiple representations is also an integral part of daily instruction in mathematics classrooms: Multiple representations are indispensable for the introduction of new mathematical objects, in problem-solving processes and in order to facilitate explanations (Stylianou, 2010; Leinhardt, 2001). Furthermore, it would not be possible to communicate about ideas and problems and to share strategies in the mathematics classroom without using multiple representations (Roth & McGinn, 1998) in the form of graphs, formulae, drawings, tables or diagrams.

However, any representation of a mathematical object does not stand alone but belongs to a system that is structured by common conventions that underlie it (Goldin & Shteingold, 2001). Duval (2006) describes those different systems of representation that can be mobilised in mathematical processes as *representation registers*. He distinguishes, for example, between natural language (used either orally, like in explanations or written, like in theorems or proofs), written symbolic systems, iconic and non-iconic shape configurations (drawings, sketches and configurations constructed with geometric tools), diagrams and graphs (Duval, 2006). The example in Figure 1 shows multiple representations from different registers that can stand for the mathematical object *one out of four equal parts*: It can be represented symbolically in the register of fraction notation or decimal notation, as a place on the number line in the number line register, in different verbal representations belonging to the register of language or as part of a rectangle in the fraction bar register. The example shows also, that different registers can emphasise different aspects of the corresponding mathematical object: *One out of four equal parts* can be viewed as part of a whole in the circle chart or as part of several wholes, as can be seen in the representation with the four sweets. Consequently, being able to use more than one representation of a mathematical object is essential, as any representation will express some but not all information of the related mathematical object, stress some aspects and rather hide others (Dreyfus, 2002).

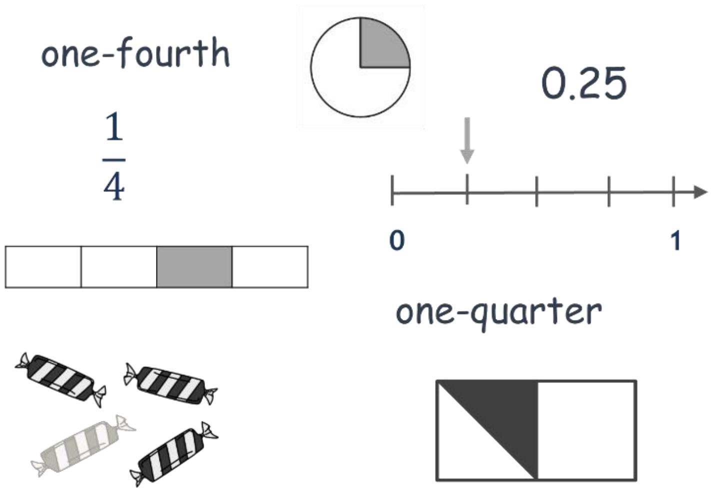


Figure 1: Multiple representations of the mathematical object *one out of four equal parts*

As any mathematical activity involves the use of representations, learners of mathematics have to manage cognitively challenging tasks related to the use of representations in order to understand a mathematical object and to be successful in problem solving: (1) They have to recognise the common mathematical object behind different registers of representation without confusing those representations with the object they stand for (Duval, 2006; Lesh, Post & Behr, 1987). (2) They have to learn about the conventions that underlie a register of representation in order to carry out transformations within a given register (Lesh, Post & Behr, 1987; Ainsworth, 2006). (3) They have to be able to translate a mathematical object from one representation register to another as many tasks involve at least two representation registers (Duval, 2006; Lesh, Post & Behr, 1987). Changing registers becomes also necessary whenever another representation appears to be more efficient in the process of problem solving (Dreyfus, 2002; Ainsworth, 2006).

These tasks related to the special role of multiple representations involve specific and complex cognitive requirements for any learner of mathematics. Numerous studies have shown that dealing with multiple representations can lead to serious difficulties in understanding if learners do not manage the complex cognitive tasks associated with their use (Lesh, Post & Behr, 1987; Janvier, 1987; Sierpinska, 1992;

Yerushalmy, 1991; Dreyfus, 2002; Ainsworth, Bibby & Wood, 1998; Ainsworth, 2006; Duval, 2006; Mitchell, Charalambous & Hill, 2014). When dealing with multiple representations, *transformations of representations* merit special attention as they are not only essential for problem solving but also for the construction of deep understanding in mathematics (Duval, 2006; Ainworth, 2006; Goldin & Shteingold, 2001; Deliyianni et al., 2016). According to Duval (2006, p. 111) transformations that are carried out within a representation register and which are specific to it can be described as *treatments* and transformations requiring the change of registers as *conversions*. Reducing a fraction to its lowest term within the representation register of fraction notation would accordingly be called a *treatment* (e.g., $\frac{4}{8} = \frac{1}{2}$), whereas converting a number from the representation register of fraction notation to the register of decimal notation would be referred to as a *conversion* (e.g., $\frac{4}{8} = 0.5$). Reading a fraction from the number line or representing a given fraction using a real-world situation (see Figure 1) involves also conversions, as each of these tasks requires changes of representation registers.

Conversions have, however, proved to be more complex than treatments as they contain problems of *discrimination* and *recognition*: In order to change registers, the learner must discriminate mathematically relevant features from those that are not relevant in a mathematical sense, transfer the information from one register to the other and also be able to recognise the same mathematical object behind the different representations (Duval, 2006). The consequence is that learners find carrying out conversions extremely difficult and often fail to see the links between different registers of representations (Ainsworth, 1999, 2006). Consequently, conversions are a source of problems in understanding at every level of teaching and in every domain of mathematics (Duval, 2006; Lesh, Post & Behr, 1987). Changing registers of representations can become even more difficult if the direction of the conversion is inverted and this inversion is more complex or new for the students (Deliyianni et al., 2016; Duval, 2006). The degree of familiarity with a certain representation and/or a specific topic can lead to more or less available cognitive resources and can thus enhance or impede the translation between different representation registers (Ainsworth, 2006). Accordingly, plotting a given fraction on a number line might, for example, prove to be harder than reading a fraction from a number line. It also adds to the complexity of conversions if a one-to-one mapping between the meaningful constituents of the source register and the target register is difficult to carry out (Duval, 2006). The more the source register and target register differ, the more difficult it will become for learners to translate between different representations and to recognise the common mathematical object behind them (Ainsworth, 2006): Changing successfully from the representation register of fraction notation to the register of decimal notation (e.g., $\frac{4}{8} = 0.5$) requires to recognise the common mathematical object and to transfer corresponding information although the different conventions in each of the used registers result in totally different numerals (4 and 8 compared to 0 and 5) and notation forms (vertical notation including a fraction bar compared to horizontal notation including a decimal point).

The importance of students' ability to carry out conversions when dealing with multiple representations is also reflected in National Standards and international student assessment. There is agreement that dealing with multiple representations does not only involve the selection and creation of different representations but also the ability to change representation registers. In the German National Standards, *using representations* is one of five overarching mathematical competencies in the curriculum for both primary and secondary schools. Related learning objectives define that students should be able to choose, develop and use appropriate representations of mathematical objects, change between different registers of representations, compare them and evaluate different representations in order to solve mathematical problems (KMK, 2004). The National Council of Teachers of Mathematics (e.g., NCTM, 2000) elevated the status of representations to one of five process standards of what mathematics instruction should enable students to do. It is recommended that students at all levels create and use representations to organise, record and communicate mathematical ideas and to select, apply and translate among mathematical representations to solve problems. In the OECD Programme for International Student Assessment (e.g., PISA 2012), using representations of mathematical objects to solve problems is seen as one of seven fundamental mathematical capabilities and thus an essential part of mathematical literacy (OECD, 2013). Using a variety of representations such as graphs, tables, diagrams, pictures, equations, formulae and concrete materials does not only entail the creation, selection and interpretation of multiple representations but also the flexible translation between them (ibid.).

Summing up, the development of students’ ability to change representation registers is not only a key task for learning with multiple representations (Ainsworth, 2006; Deliyianni et al., 2016) but can also be described as the “true challenge of mathematics education” (Duval, 2006, p. 128). Professional requirements for teachers regarding the use of multiple representations in the mathematics classroom merit hence special attention and will be derived in the following chapter.

2.2 Professional requirements related to the use of multiple representations in the mathematics classroom

Learning and understanding mathematics requires to obtain rich representations of a mathematical object and to be able to use them in problem solving in a flexible and controlled manner (Duval, 2006; Dreyfus, 2002; Acevedo Nistal et al., 2009). At the same time, the complex tasks and high cognitive demands students have to manage when dealing with multiple representations and carrying out conversions can hinder the learning process, diminish problem-solving performance and cause difficulties in students’ understanding (Duval, 2006; Acevedo Nistal et al., 2009). Against the background of this double role of multiple representations and their coordination for learning mathematics, the following *professional requirements for mathematics teachers* can be derived:

Teachers have to be aware of the special role multiple representations play in the mathematics classroom and that they can be aid and obstacle for learning mathematics (Dreher & Kuntze, 2015b; Acevedo Nistal et al., 2009). On the one hand, teachers have to provide the students with learning opportunities that contain multiple representations of a mathematical object in order to enhance their understanding and flexible problem-solving. On the other hand, they have to keep in mind that the cognitive demand of coordinating these multiple representations might also hinder understanding. When teachers demonstrate problem-solving processes using different representation registers or generate multiple representations in order to facilitate explanations (Dreher & Kuntze, 2015b; Stylianou, 2010; Leinhardt, 2001; Duval, 2006), they are required to balance this dilemma in a reflected and situation-specific way, taking into account the particular students, the content and the specific representations (Dreher & Kuntze, 2015b; Acevedo Nistal et al., 2009).

The use of representations is, however, not only in the charge of the teacher. Due to the fact, that representations are the only access to mathematical objects, multiple representations are also generated by students in the process of problem-solving, when they share strategies or negotiate ideas in the classroom (Roth & McGinn, 1998; Cuoco & Curcio, 2001; Duval, 2006). Student-generated representations can hence be used to gain insight into student reasoning and thinking in the mathematics classroom (Lesh, Post & Behr, 1987; Cuocu & Curcio, 2001; Stylianou et al., 2000; Mitchell, Charalambous & Hill, 2014). Consequently, teachers are required to carefully attend to the representations students generate and the way they use them in order to understand how they think and whether problems in understanding are related to unconnected conversions or the use of a specific representation register. Analysing students' representations and the way they use them can thus be considered as essential for providing students with adequate learning support that connects to their ideas.

As multiple teacher-generated and student-generated representations of mathematical objects can occur in the course of instruction, students are continuously faced with the complex cognitive task to coordinate different representation registers and teachers are continuously required to support their students in doing so. Students have to be provided with focused help to understand the relationships among multiple representations of the same mathematical object as well as the differences and similarities between different representation registers (Goldin & Shteingold, 2001; Dreyfus, 2002; Ainsworth, 2006; Deliyianni et al., 2016). Therefore, teachers are required to attend to situations that are relevant in this context and encourage their students to actively create connections between multiple representations of a mathematical object by relating meaningful constituents to each other (Duval, 2006; Bodemer & Faust, 2006; Renkl et al., 2013).

In order to meet the professional demands related to multiple representations outlined above, teachers do not only need corresponding *professional knowledge* (Duval, 2006; Goldin & Shteingold, 2001; Ball, Thames & Phelps, 2008; Mitchell, Charalambous & Hill, 2014) but they also have to link this knowledge to relevant observations in the classroom, what is referred to as *analysing classroom situations regarding the use of multiple representations* in this study (Friesen, Dreher & Kuntze, 2015; Friesen & Kuntze, 2016; Kuntze & Friesen, 2016).

In order to provide insight into relevant frameworks for such analysing, the next chapters will outline related theoretical frameworks and the state of research regarding teachers' *professional knowledge* on multiple representations (section 2.3) and corresponding *analysing of classroom situations* (section 2.4).

2.3 Mathematics teachers' professional knowledge on multiple representations

There is a broad consensus that a teacher's professional knowledge is a key component of professional competence (Baumert & Kunter, 2013) and an important prerequisite for teaching effectively (e.g., Petrou & Goulding, 2011; Ball, Thames & Phelps, 2008). With regard to representations, many approaches emphasise that teachers have to know about the essential role multiple representations play for the learning of mathematics and about the cognitive demands related to the change of representation registers (Duval, 2006; Goldin & Shteingold, 2001; Ball, Thames & Phelps, 2008; Mitchell, Charalambous & Hill, 2014). Only when teachers have knowledge, for example, about the difference between treatments and conversions, they can analyse related problems of mathematics comprehension and support their students in mapping meaningful constituents between different representation registers (Duval, 2006; Deliyianni et al., 2016; Dreher & Kuntze, 2015a). Although quite different frameworks on teachers' professional knowledge exist, most of them are based on Shulman (1986, 1987) and contain knowledge on representations as a core component of mathematics teachers' professional knowledge (Depaepe, Verschaffel & Kelchtermans, 2013). A closer review of prominent frameworks reveals, however, that specific knowledge related to the connection of multiple representation registers for the learning of mathematics and related cognitive demands for students is often not sufficiently taken into account.

2.3.1 Frameworks on teachers' knowledge regarding multiple representations

As the use of representations is considered to be subject-specific, corresponding knowledge is mostly conceptualised as a part of content-related professional knowledge. According to Shulman (1986, 1987), the content dimension of teacher knowledge is mostly organised in three categories: (1) Content knowledge (CK) including the knowledge of the subject, (2) curriculum knowledge and (3) pedagogical content knowledge (PCK). Pedagogical content knowledge (PCK) is regarded to go “beyond knowledge of subject matter per se to the dimension of subject matter knowledge *for teaching*” (Shulman, 1986, p. 9). Shulman (1986) identified two components that are central to PCK, namely *knowledge of instructional strategies and representations* and *knowledge of students’ (mis)conceptions*. Accordingly, multiple representations, which can be used in order to make subject matter comprehensible to students, form an essential part of teachers’ pedagogical content knowledge (Shulman 1986, 1987). Shulman’s conceptualisation of teachers’ PCK is, however, not specific to the subject of mathematics. Hence, the special role of coordinating multiple representation registers for learning mathematics and the complex cognitive demands related to unconnected conversions are not explicitly taken into account.

Fennema and Franke (1992) proposed a framework for mathematics teachers’ professional knowledge including four components: knowledge of content, knowledge of pedagogy, knowledge of students’ cognition and teachers’ beliefs. They regard teachers’ knowledge of mathematical representations as a central aspect of *content knowledge* as it enables the “translation of mathematics into understandable representations” (Fennema & Franke, 1992, p. 153). They suggest, for example, that most contents in school mathematics can be represented as real-world situations, concrete or pictorial representations in order to help students to learn the “abstract ideas of mathematics with understanding” (ibid., p. 154). However, taking this rather narrow perspective on representations as a kind of “helpful media” does neither capture the full range of representation registers nor their full meaning for the conceptual learning of mathematics. The consequence is that knowledge on the coordination of multiple representation registers and related cognitive demands for students are not included in this framework.

In the TEDS-M study (*Teacher Education and Development Study: Learning to Teach Mathematics*; Blömeke, Kaiser & Lehmann, 2010), primary and secondary mathematics teachers in their final year of teacher education were assessed on their professional knowledge and beliefs. Based on the categories proposed by Shulman (1986, 1987), pedagogical content knowledge as understood in TEDS-M refers to knowledge about teaching and learning mathematics and covers the sub-areas

Knowledge of curricula and planning, *Instructional Designs* and *Interactive Knowledge* (e.g., Blömeke, 2013). Regarding the use of multiple representations, so-called core situations mathematics teachers are expected to cope with focus on teachers' knowledge related to *Planning appropriate methods for representing mathematical ideas* and *Explaining or representing mathematical concepts or procedures* (Blömeke, 2013). However, released TEDS-M items in this context (Tatto et al., 2008; 2012) do not indicate any reference to the requirement of supporting students in relating multiple representations of a mathematical object to each other when conversions have to be carried out. Consequently, even though representations are addressed in the approach of TEDS-M, it appears that teachers' knowledge on coordinating and connecting multiple representations of a mathematical object is not sufficiently taken into account.

In the COACTIV study (*Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*), teachers' professional knowledge is seen as a key dimension of professional competence (Baumert & Kunter, 2013). Five domains of professional knowledge (e.g., pedagogical content knowledge) are further differentiated into facets of knowledge, which are operationalised by specific indicators. With respect to the use of representations in the mathematics classroom, the PCK facet *Making content accessible: explanations and representations* describes teachers' ability to explain mathematical content by using appropriate representations (Kunter et al., 2013). Accordingly, twelve instructional scenarios were designed that require teachers to provide support for students' understanding. However, the emphasis was put on teachers' knowledge of multiple representations in the sense of providing "useful representations" (Krauss et al., 2013, p. 151) or "as many different ways as possible of explaining" (ibid., p. 152) in order to make mathematical content accessible. Thus, the released items in this context (Krauss et al., 2013) do not target teachers' knowledge on cognitive demands related to the use of multiple representations and the need for student support when conversions take place. Comparable to TEDS-M, teachers' knowledge on the coordination of multiple representations and the role of conversions for students' learning of mathematics is not sufficiently taken into account in the COACTIV study.

The MKT framework (e.g., Hill, Schilling & Ball, 2004; Hill, Rowan, Ball, 2005; Ball, Thames & Phelps, 2008) is a conceptualisation of mathematics teachers' knowledge that particularly highlights the specific role of connecting multiple representations to each other and to the common mathematical object behind them. The framework is built on an analysis of the mathematical demands of teaching and identifies *mathematical knowledge for teaching* (MKT) that is needed to perform the tasks of teaching mathematics to primary students. The MKT subdomain *specialised content knowledge* (SCK), which includes mathematical knowledge and skills unique to teaching mathematics, identifies *linking representations to underlying ideas and to other representations* as a central task of teaching mathematics (Ball, Thames & Phelps, 2008, p. 400). Teachers' knowledge on how to select, create and use mathematical representations effectively, for example by recognising advantages and disadvantages of using certain representations for particular purposes, is included in the MKT subdomain *knowledge of content and teaching* (KCT).

Building on the MKT framework, Mitchell, Charalambous and Hill (2014) generated a more detailed list of tasks entailed in dealing with representations in the mathematics classroom and inferred knowledge demands needed to successfully engage in this work. Derived from their analysis of classroom episodes, they particularly emphasise teachers' knowledge regarding the coordination of multiple representations, related student demands and the need for specific student support. Accordingly, *creating a context for connecting multiple representations* is described as a core task for mathematics teachers: It includes the identification of similarities and differences between representations as well as student support in using one representation to make sense of another. With respect to conversions, treatments and related demands for students, teachers should have knowledge on "the connections that can be made between and among representations to forge students' learning" (ibid., p. 55) and on "common errors that students can make when working on representations" (ibid., p. 56). They include an "awareness of the fact that representations are not transparent in and of their own" (ibid., p. 56) and emphasise that students need specific scaffolding when they work with representations. Referring to the dilemma of multiple representations being aid and obstacle for the learning of mathematics (Dreher & Kuntze, 2015b), it is argued that without those knowledge components described above "representation use can do more pedagogical harm than good" (Mitchell, Charalambous & Hill, 2014, p. 57).

Given the high relevance of teachers' knowledge regarding the coordination of multiple representations and related demands for learning mathematics, it is remarkable that corresponding empirical research is still scarce (Dreher, 2015). As Mitchell, Charalambous and Hill (2014) point out: "…most of our understanding about the tasks involved with using representations in instruction and the knowledge requirements imposed on teachers when using these aids is theoretical" (ibid., p. 37). The lack of empirical evidence in this specific context might also reflect what has been outlined above: In most of the prominent frameworks on mathematics teachers' professional knowledge, aspects related to the coordination of multiple representations and corresponding demands for students have so far not been sufficiently accounted for. There are, however, several studies by Dreher and Kuntze (2015a, b) and Dreher, Kuntze and Lerman (2016), in which these particular aspects of teacher knowledge for dealing with multiple representations were investigated in the case of pre-service and in-service teachers. As these studies form an important basis for the study presented here, underlying frameworks and main results will be presented in the following section.

2.3.2 Teachers' knowledge on the coordination of multiple representations: state of research

The knowledge framework used in the studies by Dreher and Kuntze (2015a, b) and Dreher, Kuntze and Lerman (2016) was introduced by Kuntze (2012; see Figure 2). It integrates three dimensions according to which different components of mathematics teachers' professional knowledge can be structured. The first dimension shows components of professional knowledge which are based on the domains by Shulman (1986, 1987). The spectrum between knowledge and teachers' beliefs resp. convictions were included in the model as a second dimension: They are considered to be aspects of professional knowledge, as distinguishing mathematics teachers' knowledge from their beliefs has often been found to be difficult (e.g., Pajares, 1992). The third dimension of the knowledge framework by Kuntze (2012) is constituted by different levels of globality (cf. Törner, 2002). Accordingly, knowledge and views about teaching mathematics can be very global, but they can also be related to a particular content domain in mathematics, a certain content or topic within this domain or even to a particular instructional situation.

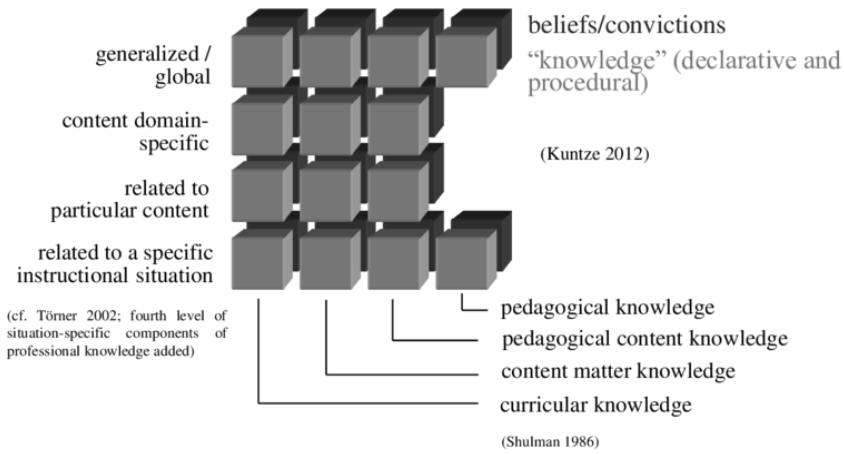


Figure 2: Components of professional knowledge (Kuntze, 2012, p. 275)

Kuntze (2012) emphasises that the different components might be interdependent, either mutually supportive or conflicting, and that therefore the cells of the model cannot be considered as strictly separable categories. Regarding the use of multiple representations in the mathematics classroom, Dreher and Kuntze (2015a, b) as well as Dreher, Kuntze and Lerman (2016) build their research mainly on the framework by Duval (e.g., 2006; cf. section 2.1). Therefore, the specific role of multiple representations for learning mathematics, the importance of the coordination of different representation registers and related cognitive demands for students were particularly taken into account.

In order to explore which reasons for using multiple representations in the mathematics classroom are important for teachers (Dreher & Kuntze, 2015b; Dreher, Kuntze & Lerman, 2016), pre-service and in-service teachers were asked to evaluate the importance of four possible reasons: (1) necessity for mathematical understanding (sample item: "*Enhancing the ability to change from one representation to another is essential for the development of mathematical understanding.*"), (2) motivation and interest, (3) support of remembering and (4) different learning types/input channels (e.g., Dreher, 2015, p. 32; Dreher, Kuntze & Lerman, 2016). The participants' evaluations showed that both, pre-service and in-service teachers rated the mathematics-specific reasons for using multiple representations as significantly less important than the other more general reasons (Dreher, Kuntze & Lerman, 2016; Dreher & Kuntze, 2015a). It can hence be concluded that pre-service and in-service teachers were not fully aware of the essential role that multiple representations and their coordination play for the development of mathematical understanding. However, going down a level of globality (cf. Kuntze, 2012) and focusing on teachers' views related to the role of multiple representations in the case of fractions yielded a clear difference between pre-service and in-service teachers (Dreher & Kuntze, 2015a): The in-service teachers attached significantly higher importance to the role of conversions for the students' understanding than the pre-service teachers did. Sample items of this study (e.g., *It is essential for their understanding that students master using different representations for fractions and that they can change between them*) reflect clearly the focus on teachers' views regarding the coordination of different registers when multiple representations are used for teaching fractions (e.g., Dreher & Kuntze, 2015b, p. 33).

The aim to investigate teachers' specific *content knowledge* on the coordination of multiple representations in the domain of fractions required the development of specific test items (e.g., Dreher, 2015; Dreher, Kuntze & Lerman, 2016). Based on the framework by Duval (e.g., 2006), eight corresponding items were developed. They have in common that given (incorrect) conversions between representations had to be checked and corrected or a conversion had to be carried out (see Figure 3 and 4 for sample items).

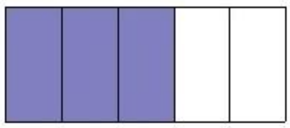
Please change the diagram, if necessary, so that $\frac{3}{5}$ of $\frac{1}{4}$ is shaded. Otherwise just tick the box on the right-hand side.		

Figure 3: Sample item # 1 of the knowledge test on conversions (Dreher, Kuntze & Lerman, 2016, p. 374)

Please change the score, if necessary, so that the home team has scored exactly $\frac{1}{5}$ of the goals. Otherwise just tick the box on the righ-hand side.	Standing in a soccer game: home away 1 : 5	

Figure 4: Sample item # 2 of the knowledge test on conversions (Dreher, Kuntze & Lerman, 2016, p. 374)

Thus, the teachers' content knowledge related to matching symbolic-numerical representations of fractions and related to operations with appropriate pictorial and content-related representations was assessed. The results showed that both pre-service and in-service teachers had gaps in their corresponding content knowledge: on average, less than half of the items were solved correctly by pre-service teachers (Dreher, Kuntze & Lerman, 2016). A comparison between pre-service and in-service teachers' results (Dreher & Kuntze, 2015a) revealed that although in-service teachers achieved significantly higher scores, they solved on average less than 60% of the items correctly.

Another study by Dreher and Kuntze (2015b) focused on the dilemma of multiple representations being aid and obstacle for learning mathematics (cf. section 2.2). As balancing this dilemma is considered to be essential for effective instruction and student support, in-service teachers were asked about their views on the role of multiple representations for teaching fractions ("multiple representations for understanding" vs. "fear of confusion by multiple representations"). A cluster analysis revealed that only about one sixth of the participating teachers took into account both sides of the dilemma. Another step of inquiry towards this dilemma was taken by the evaluation of the learning potential of tasks in which multiple representations occur. Six fraction tasks were accordingly designed as contrasting types (see Figure 5): The first type of tasks required carrying out conversions of representations and the coordination of representation registers. The second type was about solving a task in a symbolic register and provided also some potentially motivating pictures. Those pictures did, however, not represent the operation that had to be carried out.

Figure 5: Sample item # 3: evaluation of the learning potential of fraction tasks (Dreher & Kuntze, 2015b, p. 33)

In order to evaluate the learning potential of these tasks, corresponding statements had to be rated by the participants, for example: *The way representations are used in this task aids students' understanding* (Dreher & Kuntze, 2015b, p. 32). The results show clear differences between in-service teachers from different subsamples:

Teachers from academic-track schools evaluated the learning potential of tasks involving conversions of representations significantly higher than teachers from secondary schools for lower-achieving students. Again, only a minority of the participants accounted in their evaluations for the dilemma of multiple representations being aid and obstacle for the learning of mathematics (Dreher & Kuntze, 2015b). It can therefore be assumed that most teachers were not aware of the cognitive demands and potential learning difficulties related to the coordination of multiple representations. Dreher & Kuntze (2015b, p. 40) conclude that "the need for corresponding teacher professional development is evident from the results" of their study and call for more specific professional development programmes.

However, knowledge on the double role multiple representations play for learning mathematics as well as a corresponding awareness are not the only prerequisites for mathematics teachers in order to meet the professional requirements related to the use of multiple representations. It is claimed that based on their knowledge on the decisive role of the coordination of representation registers, teachers must "set up a mechanism of observation" (Duval, 2006, p. 121) in order to *see* students' difficulties related to conversions of representations. *Analysing classroom situations regarding the use of multiple representations* can hence be seen as key for teachers to master related requirements in the mathematics classroom: Only when teachers are sufficiently aware of the special role multiple representations play for the learning of mathematics, they will be able to identify relevant situations and to interpret them against the background of their corresponding knowledge. As teachers' analysing of classroom situations has proved to be highly consequential for their decision-making (Schoenfeld, 2011; Stahnke, Schueler & Roesken-Winter, 2016), it can be seen as an important prerequisite for teachers to provide their students with adequate learning support and learning opportunities adapted to their needs – also in the context of dealing with multiple representations in the mathematics classroom.

Following this argumentation, relevant frameworks and the state of research related to mathematics teachers' analysing of classroom situations will be outlined in the next chapter (2.4.1). As empirical evidence on teachers' analysing with respect to *the use of multiple representations* is especially relevant in the context of the study presented here, related research will be presented in section 2.4.2.

2.4 Teachers' analysing of classroom situations

Teachers' analysing of classroom situations has become an important focus of mathematics education research during the last 20 years (Stahnke, Schueler & Roesken-Winter, 2016). It is assumed that *attending to* aspects of classroom interactions that influence student learning and *reasoning about* those events inform teaching decisions, which are based on "the analysis of these observations" (van Es, 2011, p. 135; Schoenfeld, 2011). Various studies provide evidence that teachers' analysing of classroom situations is highly relevant for instructional quality and student learning (e.g., Kersting 2008; Kersting et al., 2012) and that it can be successfully fostered in professional development programmes (e.g., Sherin & van Es, 2009). There is also empirical evidence that teachers' expertise and experience positively influence their analysing (Stahnke, Schueler & Roesken-Winter, 2016). Research into teacher expertise (e.g., Berliner, 2001) show that expert teachers perceive classroom situations more quickly, more accurately and more holistically than novice teachers. Comparably, teachers' analysing of classroom situations is described as component of expert practice, as expert teachers are supposed to have "heightened sensivities to particular aspects of their work, as well as techniques for analyzing, using, and inquiring into these features of their practices" (van Es, 2011, p. 135).

A review of studies on teachers' analysing reveals that different frameworks and concepts exist to describe where teachers attend to when observing classroom situations, what they see and what sense they make of what they see (Sherin, Jacobs & Philipp, 2011; Stahnke, Schueler & Roesken-Winter, 2016). For this reason, main concepts and related research will be presented in the following section.

2.4.1 Teachers' analysing of classroom situations: concepts and state of research

When teachers' analysing of classroom situations is investigated, different terms are used for the same concepts as well as similar concepts are described in various terms (Mason, 2016). There is also hardly a consensus on the processes that are involved in teachers' analysing and concepts range from including perception only (e.g., Star & Strickland, 2008) to connecting perception with interpretation (e.g., van Es & Sherin, 2002) to including perception, interpretation and decision-making (e.g., Kaiser et al., 2015). Moreover, reviewing related articles reveals two main approaches to teachers' analysing, namely (1) *analysing in-action* where the teachers are themselves actors in the situation they observe and (2) what can be described as *armchair analysing*, as teachers are mere spectators of a classroom situation that is presented to them, for example, as a video clip.

(1) Approaches where the teachers are themselves actors in the instructional scene they are observing are, for example, referred to as *professional vision in-action* (Sherin, Russ & Colestock, 2008) or *teachers' in-the-moment noticing* (Sherin, Russ & Colestock, 2011). Mostly set in the context of mathematics education reform (van Es, 2011), these approaches try to capture the processes through which teachers manage the ongoing information in the midst of instruction and highlight the importance of teachers' ability to respond to students' mathematical thinking during instruction. As adaptive and responsive teaching requires teachers to attend to the ideas that students raise in class and how these ideas relate to the mathematical objectives of the ongoing lesson, it is informed by teachers' *attending to particular events* (e.g. students' strategies) and *making sense of those events* (Sherin, Jacobs & Philipp, 2011). Related studies use, for example, in-the-moment videotaping (e.g., Sherin, Russ & Colestock, 2008) and are often designed as case studies including only a few teachers (cf. Stahnke, Schueler & Roesken-Winter, 2016). Mostly, the following three processes are investigated: *attending to* students' strategies, *interpreting* student responses and solution methods in terms of the mathematical understanding they reveal as well as *deciding in the moment how to respond* on the basis of students' understanding (e.g., Jacobs et al., 2011; Jacobs & Empson, 2016).

(2) Approaches where the teachers are only spectators of classroom situations (e.g., presented as videos) investigate, for example, *teachers' abilities to analyse teaching* (e.g., Santagata, Zannoni & Stigler, 2007; Santagata & Yeh, 2016), *teachers' professional vision* (Sherin & van Es, 2009) or *teachers' ability to notice classroom interactions* (e.g., Sherin & van Es, 2005). Related studies often stem from the context of professional development programmes which aim, for example, at preparing teachers to conduct systematic analysis of their own teaching (e.g., Santagata, 2011; Santagata & Yeh, 2016) or at developing teachers' noticing in the context of a video club (e.g., van Es & Sherin, 2008; Sherin & van Es, 2009). In this context, *noticing*

is mostly considered to involve two interrelated processes (van Es & Sherin, 2002; Sherin, Jacobs & Philipp, 2011): attending to particular events in a classroom situation (*selective attention*) and making sense of those events which is considered as knowledge-based (*knowledge-based reasoning*). An important component of this research is the focus on teachers' analyses of student learning and the relationship between teaching moves and learning results (van Es, 2011). Teachers' *ability to analyse teaching* (e.g., Santagata, Zannoni & Stigler, 2007) includes (1) attending to the details of the teaching-learning process, (2) elaborating on these details to examine the impact of teacher decisions on student progress towards learning goals, and (3) proposing improvements in the form of alternative strategies.

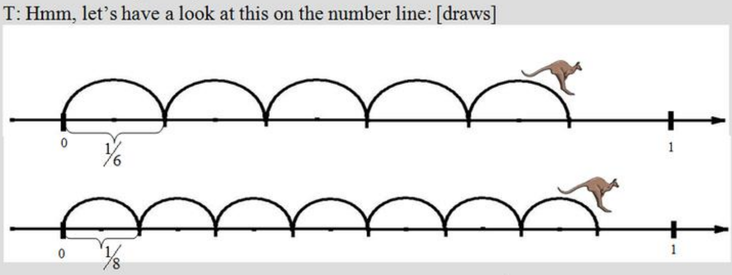
Another important approach is the implementation of analysing tasks in standardised test instruments. Results of teachers' analysis of classroom situations are then, for example, taken as indicators of so-called *professional vision* (e.g., Stürmer & Seidel, 2015), *situation-specific skills* (e.g., Blömeke, Gustafsson & Shavelson, 2015), *usable knowledge* (Kersting, 2008; Kersting et al., 2016) or aspects of *professional competence* (e.g., Bruckmaier et al., 2016; Knievel, Lindmeier & Heinze, 2015; Kaiser et al., 2015; Hoth et al., 2016). In this context, research focusing on teachers' analysing of classroom situations regarding the use of multiple representations and in particular the coordination of representation registers is, however, still very scarce. Dreher & Kuntze (2015a, b) took a step into this direction of inquiry by relating the framework of teacher noticing (e.g., van Es & Sherin, 2002) to teachers' professional requirements when dealing with multiple representations in the mathematics classroom. As their research can be considered as fundamental for the study presented here, main results will be described in the following.

2.4.2 Research on teachers' analysing of classroom situations regarding the use of multiple representations

Building up on the noticing framework and focusing on the use of multiple representations, Dreher and Kuntze (2015a, b) described the analysing processes they investigated as *theme-specific noticing*. Accordingly, teachers' theme-specific noticing encompasses paying attention to conversions in student-teacher interactions and evaluating these conversions by drawing on corresponding professional knowledge. Dreher & Kuntze (2015a, b) assessed theme-specific noticing in a vignette-based test comprising of four fictitious classroom situations. Each of the short transcripts (see Figure 6 for a sample vignette) shows a student-teacher interaction in which a student asks a question or reveals a misconception when solving a fraction task. The teachers' reaction involves a so-called unnecessary change of representation registers without making explicit connections between the used representations and can thus be considered as potentially hindering for the students' understanding. The participating teachers were asked to evaluate the teachers' reaction according to the following question: "*How much does this response help the student? Please evaluate the use of representations in this situation and give reasons for your answer.*" (Dreher & Kuntze, 2015b, p. 98).

S: The fractions $\frac{5}{6}$ and $\frac{7}{8}$ are of equal size! There is always one piece of the whole missing.

T: Hmm, let's have a look at this on the number line: [draws]



Now you can see that the kangaroo that jumps five times $\frac{1}{6}$ does not get as far as the kangaroo that jumps seven times $\frac{1}{8}$, can't you?

Figure 6: Sample item # 4: text-based vignette for the assessment of theme-specific noticing (Dreher & Kuntze, 2015b, p. 99)

An answer was considered to provide an indicator for theme-specific noticing if both, the change of representations and its potentially negative effect for students' understanding was mentioned by a participant. The findings revealed that pre-service as well as in-service teachers showed on average a relatively low frequency of theme-specific noticing (Dreher & Kuntze, 2015a). In this context, low specific content knowledge related to conversions was found to be hindering for pre-service teachers' theme-specific noticing. The answers of in-service teachers indicated theme-specific noticing about twice as often as those of pre-service teachers. The in-service teachers' theme-specific noticing showed also a significant relation to the global view that the ability to carry out conversions is essential for the development of mathematical understanding (Dreher & Kuntze, 2015a).

A qualitative in-depth analysis of the participants' answers revealed that the teachers drew on different components of their professional knowledge from different levels of globality (cf. Figure 2; Kuntze, 2012) when they evaluated the teachers' reaction in the four vignettes: Successful theme-specific noticing was either informed by knowledge from a single level of globality, such as situation-specific knowledge, or by a combination of different levels (Dreher & Kuntze, 2015a). This can be taken as evidence that teachers' professional knowledge and awareness regarding the special role of multiple representations for the learning of mathematics formed important prerequisites for their analysis of corresponding classroom situations.

However, as interrelations were mostly found to be low, it appears that there is "no simple relationship" between teachers' theme-specific noticing and their related professional knowledge (Dreher, 2016, p. 107). Similar results were obtained in a study by Charalambous (2008), who could not find any significant relations between pre-service teachers' noticing regarding the use of multiple representations and corresponding aspects of their professional knowledge. Both findings indicate that possessing a certain knowledge related to the use of multiple representations does not simply translate into corresponding analysing processes. Such considerations are supported by findings which show that teachers' knowledge, as measured in related tests, is not always activated during classroom analysis (Santagata & Yeh, 2016). Similarly, so-called *usable knowledge* of teaching as indicated by video analysis was found to only partially overlap with mathematics knowledge for teaching as measured by a corresponding multiple-choice survey (Kersting et al., 2012; 2016). Kersting et al. (2016) conclude that so-called *usable knowledge* indicated by successful classroom analysis requires "both individual knowledge components and an overarching ability to access and apply those components that are most relevant to a particular teaching episode" (Kersting et al., 2016, p. 106).

Teachers' knowledge regarding the use of multiple representations can hence be regarded as a necessary but insufficient prerequisite to meet related professional demands. What is needed, is the connection of such knowledge with relevant classroom observations in order to identify unconnected conversions and interpret them situation-specifically as potentially obstructing for students' understanding (e.g., Friesen, Dreher & Kuntze, 2015; Kuntze & Friesen, 2016), which is referred to as *analysing the use of multiple representations in classroom situations* in this study. The element of being capable to apply available dispositions such as knowledge to particular and variable situations is also an essential part of Weinert's (2001b) *competence* definition. He describes competences as "cognitive abilities and skills possessed by or able to be learned by individuals that enable them to solve particular problems, as well as the motivational, volutional and social readiness and capacity to utilize the solutions successfully and responsibly in variable situations" (ibid., p. 309). As the competence approach is also used to describe the characteristics teachers need in order to meet the demands of their profession (e.g., Baumert & Kunter, 2013) *analysing the use of multiple representations in classroom situations* can be described as an important profession-related *competence* of mathematics teachers.

As large-scale studies investigating competences of mathematics teachers, such as TEDS-M and COACTIV, are mainly based on Weinert's approach, the concept of competence as introduced by Weinert (1999, 2001a, b) will be outlined in the following chapter. The competence frameworks of TEDS-M and COACTIV will subsequently be explored with respect to aspects related to teachers' analysing of classroom situations regarding the use of multiple representations.

2.5 Concepts of competence in mathematics education research

Since the 1990s, there has been great progress in competence measurement through large-scale student assessments such as the *Third International Mathematics and Science Study* (TIMSS) and the *Programme for International Student Assessment* (PISA). In order to provide a conceptual basis for the comparison of achievement, the notion of competence was developed in particular for large-scale assessment in the context of PISA (e.g., Weinert, 1999). The guiding question was, what skills and abilities young adults would need at the end of education to be able to cope with the demands and challenges of their future life and to play a constructive role as citizens in society (cf. Klieme, Hartig & Rauch, 2008). Accordingly, Weinert (1999) suggested that a competence should be defined by the range of situations which have to be mastered. Assessment of competences should, consequently, be done by confronting the student with a sample of such (simulated) situations (Weinert, 1999). Weinert initially deferred from giving a unified competence definition and limited himself to pragmatic conclusions regarding the use of the concept of competence. One of those conclusions is that the concept of competence should be used when referring “to the necessary prerequisites available to an individual or a group of individuals for successfully meeting complex demands” (Weinert, 2001a, p.62). Thereby, Weinert aimed at differentiating competences from *skills*, which can be automatised to meet less complex demands. Accordingly, the concept of competence should be used “when the necessary prerequisites for successfully meeting a demand are comprised of cognitive and (in many cases) motivational, ethical, volitional and/or social components” (ibid., p. 62). The structure of a competence construct should consequently be derived from the logical and psychological structure of the corresponding demands: “What is required is a prototypical, typical and/or specific characterisation of classes of performance demands, performance criteria, and indicators of competencies.” (Weinert, 2001a, p. 57). Another important aspect in Weinert’s concept is the idea that learning processes are a necessary condition for the acquisition of competences. These “minimal criteria for a pragmatic definition of the concept of competence” (ibid.) were later reflected in Weinert’s (2001b) fundamental competence definition in which competences are seen as “cognitive abilities and skills possessed by or able to be learned by individuals that enable them to solve particular problems, as well as the motivational, volutional and social readiness and capacity to utilize the solutions successfully and responsibly in variable situations” (ibid., p. 309). Although this competence definition encompasses motivational, volitional and social abilities, Weinert proposed that cognitive and motivational aspects of competence might better be assessed as separated constructs allowing the empirical analysis of their interaction (cf. Koeppen et al., 2008).

Many large-scale studies investigating profession-related competences of mathematics teachers, such as TEDS-M and COACTIV, built up on Weinert’s (2001b) competence definition and specified it for teachers of mathematics. The *Teacher Education and Development Study: Learning to Teach Mathematics* (TEDS-M), for example, looked at how primary and secondary mathematics teachers were trained and what kinds of competences they had acquired at the end of their training (e.g., Blömeke & Kaiser, 2014). Figure 7 shows the theoretical framework of *teacher competence* in TEDS-M (Blömeke, 2013), which is mainly based on quality aspects of teaching as outlined by Bromme (1997) and the competence definition by

Weinert (2001b). In TEDS-M, *teacher competence* is accordingly considered to comprise of both, cognitive abilities and affective-motivational characteristics, even if this combination makes it difficult to consider "teacher competence" as a homogeneous construct – especially when keeping in mind Weinert's (2001b) requirement of competence being related to a coherent set of professional demands.

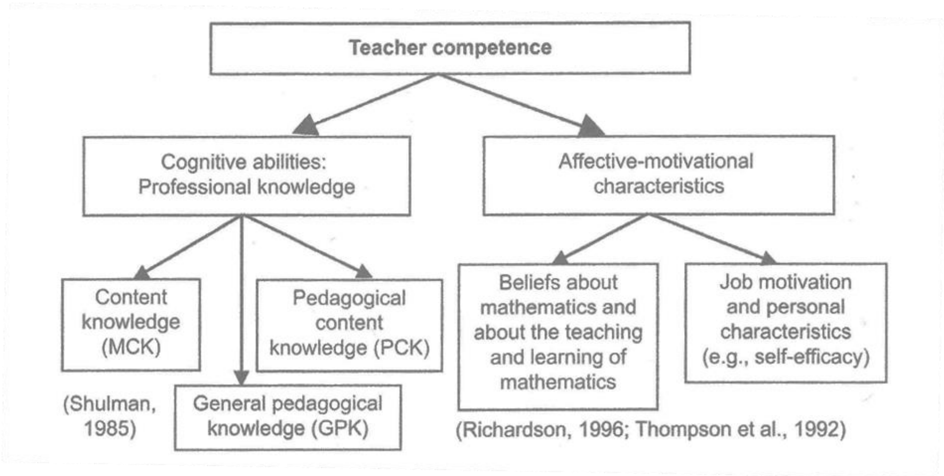


Figure 7: The TEDS-M model of teacher competence (Blömeke, 2013, p. 94)

In order to assess teacher competence in TEDS-M, so-called *core tasks* which mathematics teachers are expected to solve according to national standards were designed (e.g., KMK, 2004; NCTM, 2000; cf. Blömeke, 2013). With respect to the use of representations in the mathematics classroom, the core task *Explaining or representing mathematical concepts or procedures* referred to demands occurring during instruction (Blömeke, 2013). However, published TEDS-M items in this context (Tatto et al., 2008; 2012) do not indicate reference to demands such as supporting students in relating multiple representations of a mathematical object to each other or balancing the double role of multiple representations being help and obstacle to students' learning (Dreher & Kuntze, 2015b). Although *Analyzing or evaluating students' mathematical solutions or arguments* and *Analyzing the content of students' questions* (e.g., Blömeke, 2013, p. 95) were also included as core tasks in TEDS-M, there appears to be no relation to analysing the use of multiple representations.

Also based on Weinert (2001b) and on research into teacher expertise (e.g., Berliner, 2001), the competence concept of the COACTIV study (*Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*) aims at identifying so-called characteristics and qualities that teachers need in order to meet the demands of their profession. Accordingly, teachers' *professional competence* is defined as "the individual's ability to cope with specific occupational situations" (Baumert & Kunter, 2013, p. 42). Figure 8 shows the COACTIV model of *professional competence* which describes the qualities needed to succeed in the teaching profession from a multidimensional perspective: The COACTIV competence model comprises of four aspects of competence (professional knowledge, beliefs, motivation and self-regulation), with the aspect of professional knowledge (Shulman, 1986, 1987) emphasised as key dimension. Individual attributes in these competence aspects are considered to provide the necessary basis for effective teaching practice. Five domains of professional knowledge (e.g., pedagogical content knowledge) are further differentiated into facets of knowledge, which are operationalised by specific indicators. With respect to the use of representations in the mathematics classroom, the PCK facet *Making content accessible: explanations and representations (explanatory knowledge)* relates to teachers' knowledge on how to explain mathematical content by using multiple representations.

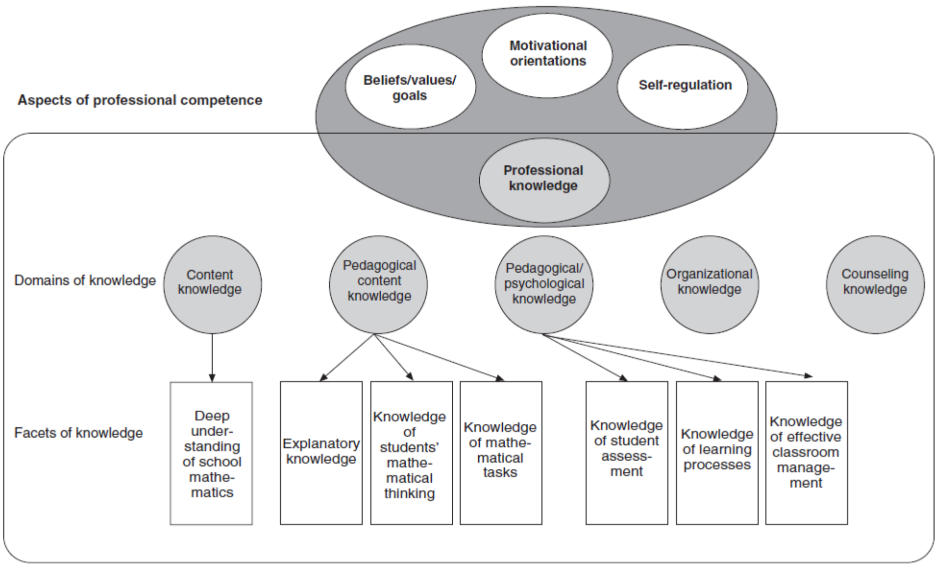


Figure 8: The COACTIV model of professional competence
(Baumert & Kunter, 2013, p. 29)

Accordingly, twelve instructional scenarios were designed that require teachers to provide support for students' understanding. Released items show that emphasis was put on teachers' knowledge on representations in the sense of providing "useful representations" (Krauss et al., 2013, p. 151) or "as many different ways as possible of explaining" (ibid., p. 152) in order to make mathematical content accessible. This suggests that professional demands related to balancing the dilemma of multiple representations being help and hindrance for the learning of mathematics (Dreher & Kuntze, 2015b) were not taken into consideration. Furthermore, teachers' analysing of classroom situations regarding the use of representations was completely neglected in the COACTIV competence model (Baumert & Kunter, 2013).

It is also remarkable that the competence approaches in TEDS-M and COACTIV have a strong emphasis on teachers' cognitive dispositions and regard related competence aspects strongly related to subfacets of pedagogical content knowledge (e.g., *Explaining or representing mathematical concepts or procedures;* Blömeke, 2013 or *Making content accessible: explanations and representations;* Baumert & Kunter, 2013). However, defining different kinds of knowledge facets and developing measures to investigate their relationship to each other, to teaching and to students' learning does not appear to be enough as "still missing are studies that help us better understand how knowledge gets activated and used in classroom settings" (Kersting et. al., 2016, p. 99). Also Kaiser et al. (2015, p. 371) point to the current need of research in teacher education and call for "a more comprehensive understanding of competence by linking cognitive with situated aspects".

TEDS-FU, a follow-up study of TEDS-M (e.g., Kaiser et al., 2015), has taken a first step into this direction of inquiry by including approaches related to teachers' *noticing* as described in section 2.4.1. Besides assessing aspects of professional knowledge based on the test instruments used in TEDS-M, a video analysis test component was developed in the framework of TEDS-FU in order to assess competence in a more situated way. Following the competence model introduced by Blömeke, Gustafsson & Shavelson (2015), three situated teacher competence facets were distinguished: (1) perceiving particular events in an instructional setting, (2) interpreting the perceived activities and (3) decision-making, for example as proposing alternative instructional strategies (Kaiser et al., 2015). These so-called *situation-specific skills* are supposed to comprise "all aspects" important for quality in mathematics teaching (Kaiser et al., 2015, p. 374) and comprise, for example, the potential for cognitive activation of the learners and individual learning support. However, a closer examination of released sample items reveals that competence aspects related to teachers' analysing of classroom situations with respect to unconnected conversions of representations or corresponding support for students were not sufficiently taken into account.

Summing up, current concepts of teacher competence aim at describing the characteristics and qualities mathematics teachers need in order to successfully meet the demands of their profession (e.g., Baumert & Kunter, 2013). Regarding the use of multiple representations, however, the competence frameworks of TEDS-M, TEDS-FU and the COACTIV study appear not to have captured the full spectrum of such characteristics. In particular, competence aspects related to teachers' analysing regarding the coordination of different representation registers and related cognitive demands for students have so far not been conceptualised and investigated in the competence frameworks described in this section. Consequently, the study presented here aims at following this need for research. Derived from the professional requirements related to dealing with multiple representations in the mathematics classroom (see section 2.2), teachers' *analysing regarding the use of multiple representations* will be defined as an important *profession-related competence* which mathematics teachers need in order to meet such requirements.

2.6 The competence of analysing the use of multiple representations in mathematics classrooms

The starting point to conceptualise profession-related competences for teachers are professional demands that teachers have to master in order to create effective teaching and learning situations in the classroom and enable students to achieve their learning objectives (Weinert, 2001b; Koeppen et al., 2008; Baumert & Kunter, 2013). In section 2.2, essential professional requirements related to the use of representations have been described in detail. Due to the special role multiple representations play for the learning of mathematics, teachers have, for example, to support students in relating different registers of representations to each other and they have to balance the dilemma of multiple representations being aid and obstacle for the learning of mathematics (Dreher & Kuntze, 2015b). The analysing of classroom situations regarding the use of multiple representations can thus be regarded as an essential prerequisite for teachers in order to meet such demands.

As described above in section 2.4.2, analysing classroom situations is regarded as an awareness-driven and knowledge-based process in which observations made in a classroom situation are connected with relevant criterion knowledge (Kuntze, Dreher & Friesen, 2015). Building up on the concept of theme-specific noticing (Dreher & Kuntze, 2015a, b; Sherin, Jacobs & Philipps, 2011) and considering analysing as a process, teachers' analysing of classroom situations regarding the use of multiple representations comprises of the following three components: *identification* of a relevant situation aspect with regard to the use of multiple representations, *interpretation* of the identified aspect based on corresponding professional knowledge and the *articulation* of the analysing result (see Figure 9). The process of analysing is, however, not seen as a linear process in which its components *identification*, *interpretation* and *articulation* follow each other in strict sequence. It might as well contain jumps between its components or even simultaneous processes (Friesen, Dreher & Kuntze, 2015; Friesen & Kuntze, 2016).

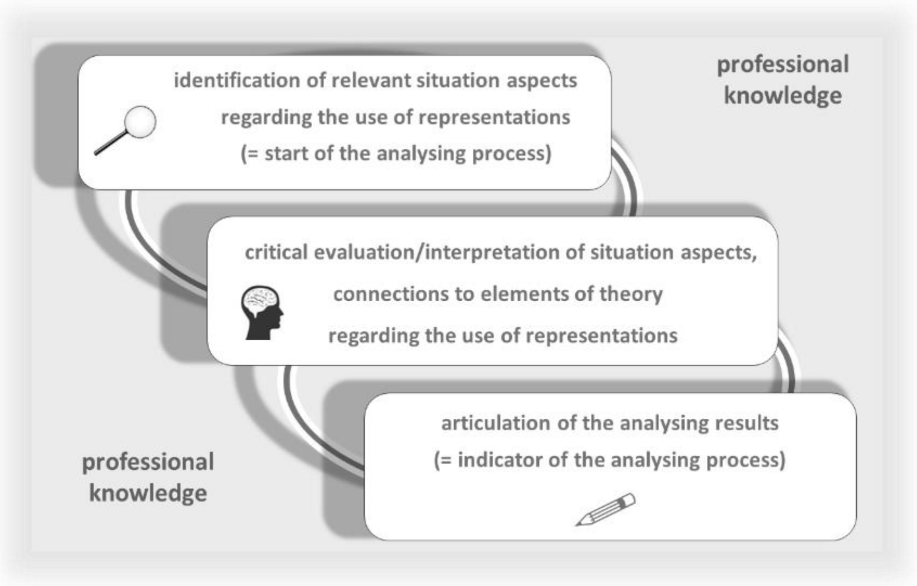


Figure 9: Analysing the use of representations as a knowledge-based process (Friesen & Kuntze, 2016, p. 260; cf. Friesen, Dreher & Kuntze, 2015)

The *identification* of a relevant situation aspect within a mathematics classroom situation can be considered to be the starting point of an analysing process in most cases. With respect to the use of multiple representations and related cognitive demands for learners, conversions of representations and observations around them can be considered as particularly relevant situation aspects. The identification of relevant situation aspects, such as unconnected conversions is supposed to be facilitated by an observer's knowledge on the key role multiple representations and their coordination play for the learning of mathematics. The *interpretation* of an identified situation aspect comprises the connection between that aspect and corresponding knowledge regarding the use of multiple representations. Such knowledge can provide the observer of a classroom situation with criteria to evaluate the identified situation aspect. Professional knowledge on the cognitive demands related to conversions can, for example, be used to evaluate situation-specifically whether an identified conversion of representations might be a potential obstacle for a student's understanding. The *articulation* of the analysing result is marked by the expression of criteria-based explanation and argumentation. With respect to dealing with multiple representations, it can reveal the situation aspects an observer has identified as relevant and can give insight into the criterion knowledge an observer has drawn on in his or her argumentation (Friesen & Kuntze, 2016; Kuntze & Friesen, 2016).

Thus, analysing the use of multiple representations can be seen as an essential prerequisite for teachers in order to meet related professional demands in mathematics classroom situations: Teachers have to link their observations with relevant criterion knowledge to find out situation-specifically whether, for example, introducing a further register of representation facilitates an explanation or could rather lead to problems in students' understanding. In order to make sure that students are supported in connecting different registers of representations to each other, teachers have to identify conversions and interpret them as potentially obstructive for students' learning, in particular if the registers remain unconnected. This relates, for example, to situations where further representations are introduced by teachers, as well as to situations where multiple representations are used by students when they share ideas or communicate during problem-solving processes.

Accordingly, analysing regarding the use of multiple representations in classroom situations can be described as an important profession-related *competence* for mathematics teachers in two ways: First, analysing as described above requires the connection of situation aspects related to the use of multiple representations with relevant criterion knowledge. Therefore, it can be seen as a competence in the sense of Weinert (2001b), who emphasises the ability to apply available dispositions such as knowledge to particular and variable situations as an essential part of his competence definition (section 2.5). Second, the successful analysing of classroom situations regarding the use of multiple representations can be regarded as an important prerequisite for teachers to master corresponding demands in the mathematics classroom (section 2.2). Only when a teacher is, for instance, able to identify and interpret potentially obstructing conversions based on criterion knowledge, he or she will be able to provide students with adaptive support in connecting different registers of representations to each other. Accordingly, and in line with Weinert's (2001b) competence definition, the *competence of analysing the use of multiple representations* is defined as a teacher's ability to successfully link observations in a mathematics classroom situation to relevant criterion knowledge regarding the use of multiple representations so that unconnected conversions are identified and situation-specifically interpreted with respect to their role as potential learning obstacle.

As this study aims at providing evidence on the assessment of such a competence, it is explicitly taken into consideration that the participating teachers are not themselves actors in the situation they observe and that they analyse other teachers teaching. Analysing the use of multiple representations is consequently conceptualised as what can be described as *armchair analysing* (see also section 2.4.1). The fundamental difference between armchair analysing and analysing in-action is also emphasised by Mason (2016, p. 221) who argues that "(…) perception and interpretation of classroom incidents involving other teachers is a far cry from acting in-the-moment in a classroom". Teachers who are only spectators of a classroom situation are, for example, unable to interact with the students and probe students' thinking (Kersting et al., 2016). They also lack the social context in which the observed teacher works like, for example, certain classroom practices and community values (Santagata & Yeh, 2016). At the same time, teachers who are involved in arm-chair analysing are not faced with the whole complexity of the ongoing information in a classroom and are not exposed to the pressure of decision-making or responding to students in the midst of instruction (Sherin, Jacobs & Philipp, 2011). As armchair-analysing entails the opportunity to sit back and try to make sense of what is going on in a classroom, it is supposed to enhance analysing processes as described above while such processes might be impeded by the pressure of decision-making and acting in real teaching situations. Another important aspect is that teachers have to manage multidimensional tasks and face simultaneity during instruction (e.g., Doyle, 1977) while armchair analysing can facilitate the process of analysing by focusing on a certain aspect, like in this case the use of multiple representations.

Regarding the assessment of the competence of analysing the use of multiple representations, the articulated analysing results are of special interest: Although they might only reflect parts of the analysing process, they can be taken as a key *indicator* for an observer's competence of analysing regarding the use of representations (cf. Oser, Salzmann & Heinzer, 2009). From an analysing result that contains

(1) reference to an unconnected, potentially obstructing conversion and

(2) a criteria-based argumentation based on relevant professional knowledge,

an observer's competence of analysing a classroom situation regarding the use of multiple representations can be inferred.

In order to compare teachers regarding their competence of analysing the use of representations, it is, however, necessary to obtain analysing results from several classroom situations (Shavelson, 2013; Hartig et al., 2012). Moreover, it is important to take into account specific characteristics of the classroom situations the teachers are presented with, as the teachers' analysing regarding the use of multiple representations might not only be influenced by their corresponding competence. Specific situational demands related to the *content* of a classroom situation, to its *presentation format* (e.g., video) or the *response format* in which teachers articulate their analysing results (e.g., open-ended format) could enhance or hinder teachers' analysing regarding the use of representations and thus impede the interpretation of analysing results. Related research is, however, still scarce – even beyond the specific competence construct described above. Consequently, the next chapter provides insight into methodological opportunities and challenges related to the assessment of teachers' competence of analysing the use of multiple representations in classroom situations.

2.7 Approaches to competence assessment

Competences are characterised by the range of situations and tasks which have to be mastered and the assessment of competence should consequently be done by confronting the test person with a sample of such (simulated) situations (Weinert, 1999). The aim of such a task sample is to elicit a test result from which an individual's level of competence can be inferred (Shavelson, 2013). How to sample tasks for the assessment of a competence construct is consequently a crucial question, as it has critical implications for the interpretation of the obtained test results. Therefore, the construct definition plays a decisive role as it describes the competence constituents that are required to master specific demands. Shavelson (2013) also emphasises that task samples should be approved by experts in the field to be representative for the domain of interest. Tasks included in competence measurement should consequently be related to demands in real-life situations which are typical for the domain of interest (Weinert, 2001b) and at the same time be standardised, providing all participants with the same tasks, scoring rubrics and testing conditions (Shavelson, 2013).

In teacher education research, test instruments entailing *classroom vignettes* have been found to meet such criteria (e.g., Lindmeier, 2013; Brovelli et al., 2014; Oser, Salzmann & Heinzer, 2009). Accordingly, the next sections provide insight into the use of classroom vignettes in the assessment of profession-related competences of teachers. Different vignette formats (video, text, animation, comic) as well as different question formats (open-ended, closed-ended) will be reviewed with respect to the assessment of teachers' competence of analysing regarding the use of multiple representations.

2.7.1 Vignette-based competence assessment

Vignettes present typical or problematic teaching situations involving specific professional demands or problems (Brovelli et al., 2014; Rehm et al., 2014). In order to meet such demands, corresponding professional competences are required. Teachers are, for example, asked to evaluate presented classroom situations according to certain criteria and make suggestions for improvement (Tepner & Dollny, 2014; Brovelli et al., 2014; Oser, Salzmann & Heinzer, 2009) or they are prompted to directly continue a lesson (e.g., Lindmeier, 2011; Knievel, Lindmeier & Heinze, 2015; Bruckmaier et al., 2016). The aim is to elicit a test result from which a teacher's level of the competence under investigation can be inferred (Shavelson, 2013). In contrast to direct classroom observations, vignette-based test instruments allow to assess competences under standardised conditions as the test persons' responses to the same classroom situations become comparable (Kaiser et al., 2015; Oser, Salzmann & Heinzer, 2009). In addition, vignette-based instruments enable the systematic assessment of competence also with larger samples of teachers (Borko, 2016).

2.7.2 The role of video vignettes

Competence assessments entailing *video vignettes* have become particularly prominent as they are supposed to provide representative job situations close to real teaching (Blömeke, Gustafsson & Shavelson, 2015; Kaiser et al., 2015; Seidel & Prenzel, 2007). Video vignettes have the potential to present complex classroom settings in a relatively authentic way and can thus also positively affect teachers' motivation to engage with presented classroom situations (Sherin, 2004; Seidel et al., 2011). In this context, Goldman (2007) uses the term *immersion* to describe the effect that videos provide the observer with enough information to be "inside" of an event. Observing video vignettes can also facilitate the *resonance* with a teacher's own teaching experiences, teaching practices of colleagues or teaching methods that are perceived as typical in a certain cultural setting (Seidel et al., 2011; Kleinknecht & Schneider, 2013). The experience of authenticity, motivation, immersion and resonance during video analysis is therefore supposed to be a facilitator of teachers' knowledge activation (Seidel et al., 2011). This is of particular interest for the assessment of teachers' competence of analysing in this study, since analysing classroom situations regarding the use of multiple representations is regarded as a knowledge-based process (section 2.6). Consequently, the test-takers' engagement with the analysed classroom situations can be regarded as an important prerequisite for their analysing regarding the use of multiple representations.

In the following, current studies using video-based approaches for competence measurement (e.g., TEDS-FU, Kaiser et al., 2015; COACTIV video, Bruckmaier et al., 2016) will be reviewed with respect to how video vignettes were implemented to assess the competence aspects under investigation. The design of the video-based test instruments differ, for example, regarding the number and length of the presented video clips, the mathematical topics and grades they cover, the camera perspective and time constraints. In TEDS-FU (Kaiser et al., 2015), for example, the secondary school teachers were presented with three video clips covering different grades (class 9 and 10), school types (academic track and comprehensive schools) and topics (functions, volumes, surfaces). The three video clips in the COACTIV video study comprised the topics inequalities including fractions, proportionality and mean values (Bruckmaier et al., 2016). The duration of the COACTIV video vignettes was up to two minutes each, those of TEDS-FU lasted up to 3.5 minutes (Kaiser et al., 2015). Lindmeier (2011) used four video vignettes lasting between 24 seconds and 2 minutes whereas Oser, Salzmann and Heinzer (2009) applied only one video lasting about 5 minutes. In order to simulate a real-life situation, in some of the studies the videos could only be watched once and rewinding or pausing was not possible (e.g., Lindmeier, 2011, Kaiser et al., 2015). For the assessment of so-called action-related competencies (e.g., Lindmeier, 2011) and the competence to recognise students' errors (Kaiser et al., 2015) time limits for teachers' reactions to the video clips were additionally implemented. Other studies did not set any time constraints (e.g., Bruckmaier et al., 2016; Oser, Salzmann & Heinzer, 2009) and explicitly allowed the teachers to view a video clip more than once. They hereby tried to compensate for the fact that the teachers are unable to interact with the students as they would be in a real teaching situation (e.g., Kersting et al., 2016) and gave them the opportunity to investigate details of a classroom situation or even revise and correct their own judgements (Oser, Salzmann & Heinzer, 2009). While some studies presented the test persons with videos recorded in real teaching situations (e.g., Kersting, 2008), others used staged videos with scripted plots (Kaiser et al., 2015) or re-enacted classroom situations (Bruckmaier et al., 2016). Different perspectives of the camera were used in order to capture different competence facets: while a point-of-view mode simulating the teachers' perspective was applied to assess action-related "direct" teacher competencies (e.g., Lindmeier, 2011), a third-person perspective showing another teacher acting in the classroom was used to assess more "reflective competencies" (e.g., Kaiser et al., 2015, p. 377). Oser, Salzmann and Heinzer (2009) used three camera perspectives simultaneously (the teacher, the entire classroom and the student to whom the teacher relates in a particular time) to assess competence in a so-called *advocatory approach:* The test persons were presented with a video vignette showing another teacher teaching his class and were asked to evaluate the observed teaching behaviour regarding certain criteria. The evaluations were then taken as indicators of the observing teachers' own competence regarding these criteria. In studies comprising several video vignettes, different lessons with different teachers and different students were enclosed in the test instruments (e.g., Kaiser et al., 2015) in order to avoid interactive effects. In most of the studies, context information about the class and the lessons prior to the one shown in the video were given before the video was played (e.g., Kersting, 2008; Kaiser et al., 2015; Bruckmaier et al., 2016).

As many current studies argue for the use of video in competence assessment, video-based vignettes could also be implemented in a test instrument assessing teachers' competence of analysing the use of multiple representations. However, a closer look at the studies outlined above show that one of the main potentials of video vignettes, namely to present "real-life, that is unstructured situations" (cf. Blömeke, Gustafsson & Shavelson, 2015, p. 9) often remains unused: In order to provide the test-takers with several classroom situations within a limited test time, most studies apply rather short video clips based on scripted plots as the implemented professional demands have to be presented in a compact form within a few minutes. Moreover, the characteristics of "maximum situativity and authenticity" that are commonly assigned to video vignettes (e.g., Brovelli et al., 2014 p. 555; Voss, Kunter & Baumert, 2011) have to be considered with caution as "instead of experiencing the real situation, a certain video-sequence is shown to the test person *as if* it were real" (Kaiser et al., 2015, p. 384). The focus of the camera allows to take only a certain perspective of the classroom and many real-life features such as classroom atmosphere or interaction with students cannot be provided by a video vignette.

Although the characteristic of authenticity is often considered as one of the main reasons for using video vignettes, only very few studies have investigated how authentic or genuine the observers perceived the presented classroom videos in a test instrument (e.g., Seidel et al., 2011; Herbst, Aaron & Erickson, 2013). With respect to the assessment of the competence of analysing in this study, it has also to be taken into consideration whether essential features entailed in video vignettes, such as high individuality, temporality (Herbst et al., 2011a) and a large amount of context information (Friesen & Kuntze, 2016) are beneficial for teachers' analysing regarding the use of multiple representations.

2.7.3 Beyond video: vignettes in the format text and animation

As outlined above, using video vignettes in the assessment of teacher competence appears to have merits as well as pitfalls. For this reason, the study presented here aims at investigating if formats other than video can be used to assess teachers' competence of analysing the use of representations in mathematics classrooms. One possible format are text vignettes as used by Dreher and Kuntze (e.g., 2015a, b) in order to assess teachers' theme-specific noticing. They applied four transcript-like text vignettes with fictitious classroom situations to elicit the observers' ability to notice potentially obstructing demands of conversions of representations for students' understanding. Their study did, however, not include the investigation of the participants' engagement with the presented text vignettes, for example in terms of authenticity or immersion (cf. Seidel et al., 2011). Although Dreher and Kuntze (2015a) argue for the use of text-based vignettes in order to avoid unintended disturbing factors due to the higher complexity in video clips, they acknowledge that video vignettes might represent classroom situations "in a more realistic way" (Dreher, 2015, p. 113) and call for corresponding research.

In the context of teacher education, Syring et al. (2015) compared pre-service teachers' work with cases in the two formats text and video. Therefore, they investigated the effects of both formats on cognitive load and motivational-emotional processes. They found that working with video vignettes appeared to cause higher cognitive load than working with text vignettes. While the effect on the pre-service teachers' motivation was comparable, the immersion in and joy of learning appeared stronger in the video analysis than in that of text vignettes.

Besides text and video formats, current research has also focused on animations with cartoon characters as possible vignette format in teacher education and research (e.g., Lande & Mesa, 2016). In the context of professional development, it has been shown that not only video clips but also animated vignettes could achieve better learning effects than vignettes in text format (Moreno & Ortegano-Layne, 2008). Comparably, Herbst, Aaron and Erickson (2013) found that an animated vignette can be just as useful as a video to engage pre-service teachers in classroom situations. They administered the same classroom situation in three different vignette formats to N=61 participants: (1) as a 6 minute unedited video clip, (2) as an animation with a nondescript character set and (3) as an animation with a more complex character set entailing greater character individuality (see Figure 10).

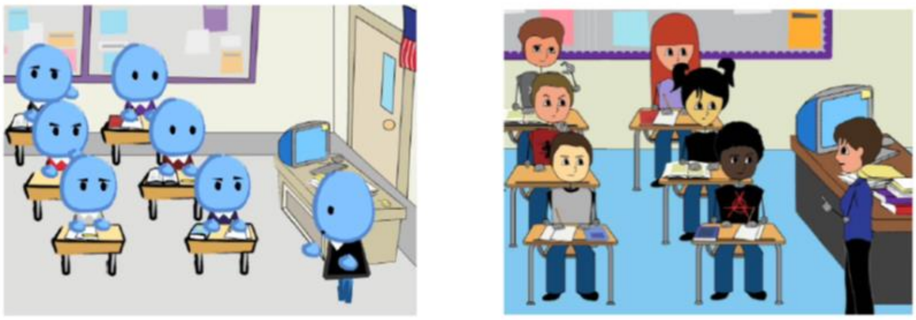


Figure 10: Screenshots of a classroom situation represented in two animation formats (Herbst, Aaron & Erickson, 2013, p. 6)

Both animations depicted the same episode of the geometry lesson captured in the video and combined an audio track with a video track. The presentation of the classroom vignette was followed by a two-hour interview arranged in the six sections genuineness, projectiveness, mathematics, noticing, reflectiveness and alternativity. While the pre-service teachers who watched the video saw it significantly more genuine than pre-service teachers who viewed either of the animated vignettes, no significant differences were found in any of the other measures. Herbst, Aaron and Erickson (2013) thus concluded that the vignette formats animation and video are comparably effective in enabling pre-service teachers, for example, to notice aspects of subject matter and pedagogy or to project a classroom situation to their own experiences. In the case of in-service teachers, Herbst and Kosko (2014) investigated N=68 experienced practitioners and compared the conversations that the teachers had in response to representations of instruction in the formats video and animation. They found that the video elicited more uses of inclination (e.g., "I'd rather do X") while the animated vignettes elicited more uses of appropriateness (e.g., "She should have done X"). Although the video and the animations dealt with situations about "doing proofs" in highschool geometry, they represented different instructional situations and the animations were no "translation" of the video vignette like in the study by Herbst, Aaron & Erickson (2013). The results of the study suggested, however, that animations are just as useful as videos to elicit modal statements about instructional practice.

2.7.4 Comic-based vignettes: a format “in between”

Herbst et al. (2016) used comic-based storyboards to assess teachers’ instructional decision-making and suggest comic-based vignettes as format “in between” (see Figure 11). Regarding temporality, a comic strip with speech bubbles can be compared to a text vignette whereas regarding individuality, it is closer to an animation or video (Herbst et al., 2011a). Similar to text vignettes, comics allow to sketch numerous and systematic variations of a classroom situation that can hardly be found and recorded in reality (Herbst & Kosko, 2014). Comparable to videos and animations, individual characteristics that might be important to fully comprehend a situation can be added in comics, without leading to lengthy descriptions that would be necessary in text vignettes.

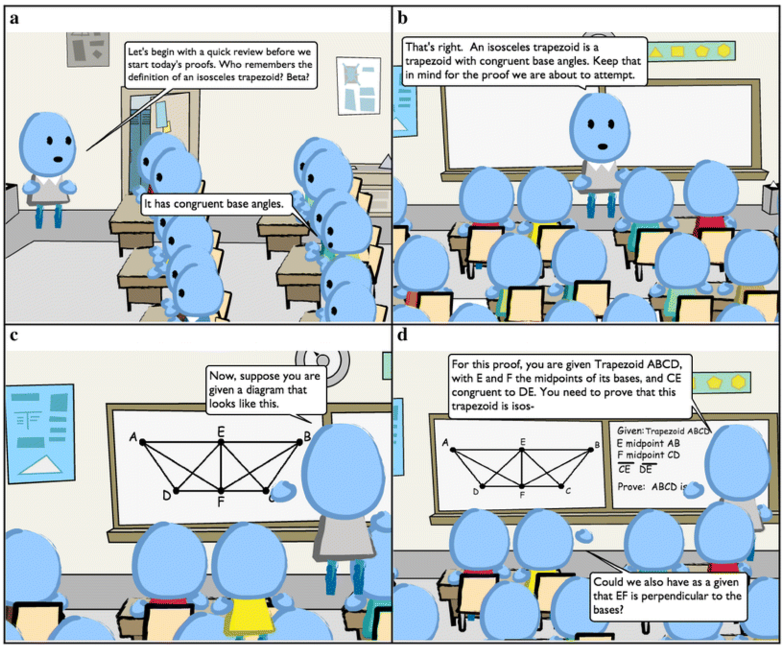


Figure 11: Classroom situation represented as comic vignette (Herbst et al., 2016, p. 174)

At the same time, unnecessary context information that might be hindering for the engagement with and analysis of a classroom situation can be left out (Friesen & Kuntze, 2016; Friesen, Kuntze & Vogel, 2017). Low individuality as provided by nondescript characters in comic vignettes might also help to project an observers’ individual teaching experience on a classroom situation and could thus facilitate the engagement with a vignette in terms of immersion, authenticity, motivation and resonance (Herbst & Kosko, 2014; Seidel et al., 2011).

2.7.5 Open vs. closed: The question of question format

Besides deciding on vignette formats in order to assess teachers' competence of analysing the use of multiple representations, the development of a corresponding test instrument requires to choose suitable question formats. A common classification of question formats is the differentiation between closed-ended (or selected-response) and open-ended (or constructed-response) questions (e.g., Haberkorn, Pohl & Carstensen, 2016). While the answers to open-ended questions have to be generated by the test-takers themselves, closed-ended formats require to select one or several answers out of pre-determined correct and incorrect options (multiple choice). Another form of closed-ended question formats consists of rating scales involving statements which the test persons are asked to rate, for example, according to a Likert-type scale (Vogt, 1999).

The characteristics of different types of question formats are not only related to aspects of test validity and reliability, but they are also supposed to have an impact on item difficulty and can even affect the dimensionality of a competence construct (e.g., Hartig, 2008). In the context of competence measurement, considerations with regard to validity support the implementation of open-ended items since the test persons have to generate the answers themselves and respond in their own words (e.g., Kaiser et al., 2015). However, solving open-ended questions often takes more time and reduces the amount of tasks that can be implemented within a given testing time. Moreover, the coding of high-inferent items can involve problems regarding reliability (Wood & Power, 1987). Due to high test efficiency, closed-ended formats are the most widely used item types in achievement tests of large-scale studies (Haberkorn, Pohl & Carstensen, 2016). Closed-ended items are, however, considered to be liable to guessing, either blind, through the elimination of alternatives or even using the format or length of the items provided (Wood & Power, 1987). It has also been found that multiple-choice items can sometimes be answered without reading the respective stem or text passage, which raises doubts regarding validity (Rauch & Hartig, 2010).

Hartig et al. (2012) point out that different response formats can affect item difficulty, as generating own answers to open-ended questions might be more complex than selecting between pre-determined answers. Several studies, including TIMSS (*Third International Mathematics and Science Study*), provide evidence that the empirical difficulty of open-ended items is often significantly higher than that of multiple-choice items (Rauch & Hartig, 2010; Hartig, 2008). Regarding the assessment of competences, different question formats can even cause items to measure different latent traits, thus leading to unintended multidimensionality (Hartig & Höhler, 2009). This could, for example, be the case if solving an open-ended item in a test instrument is not only more demanding but requires also somewhat different or additional latent abilities than solving a corresponding closed-ended item (Hohensinn & Kubinger, 2011; Hartig, 2008; Rauch & Hartig, 2010).

For assessing the competence of analysing the use of multiple representations, open-ended question formats appear to be highly applicable for several reasons: Since the test-takers have to generate the answers themselves, they have to make explicit which events of the classroom situation they refer to in their analysis and their argumentation can indicate what criteria they draw on when they evaluate those events. This reflects very well the quality of analysing and gives insight whether and how the subject of analysis is connected with relevant criterion knowledge (cf. chapter 2.6). Corresponding closed-ended formats would require the selection of pre-determined options and would thus make the subject of analysis (use of multiple representations) more explicit. The given answers might therefore provide the test-takers with corresponding criteria they can use for their analysis. In this context, an exploratory study by Friesen, Dreher and Kuntze (2014) provided evidence that N = 31 pre-service teachers' analysis of two video vignettes regarding the use of multiple representations were substantially influenced by the applied question formats: Answering open-ended questions, most pre-service teachers described the use of multiple representations as shown in the classroom videos as entirely supportive for the students' understanding, whereas answering multiple-choice questions strongly increased the number of negative and critical evaluations.

Summing up, different *question formats* as well as different *vignette formats* can be applied in order to assess teachers' competence of analysing classroom situations regarding the use of representations. From the content perspective, Dreher & Kuntze (e.g., 2015a) have shown that *text vignettes* can be used to assess pre-service and in-service teachers' noticing regarding the use of representations in mathematics classrooms. From the perspective of competence assessment, *video vignettes* appear to be highly applicable as they have the potential to present professional demands related to the use of representations in a relatively authentic way and can facilitate the observers' engagement with the presented classroom situations. Current research suggests *comic vignettes* as intermediate vignette format providing the potential of integrating advantages attributed to texts and videos. Regarding different question formats, *open-ended* and *closed-ended formats* can be implemented to assess teachers' analysing regarding the use of multiple representations in the mathematics classroom (Friesen, Dreher & Kuntze, 2014; Dreher & Kuntze, 2015a, b). However, there are good reasons to assume that the individual characteristics of different vignette and question formats correspond to diverging demands which can influence the test-takers' analysis of the corresponding classroom situations (Hartig et al., 2012; Friesen, Dreher & Kuntze, 2014).

However, there is so far very little empirical research that is format-aware and makes the effect of different vignette formats (text, video, comic) and different question formats (open-ended, closed-ended) comparable. The studies presented above differ, for example, regarding domains (educational sciences, social studies, mathematics) and vignette formats (different kinds of texts, videos, comics, animations). Moreover, the number of classroom situations involved in studies that compare different vignette formats is often rather low: Whereas Syring et al. (2015) used four vignettes from one classroom in their test instrument, Moreno and Ortegano-Layne (2008) as well as Herbst, Aaron and Erickson (2013) compared preservice teachers' responses to only one classroom situation. There is also need for research with respect to the comparison of different question formats used in competence tests: Many studies in this context are not situated in the context of teacher education, refer to domains other than mathematics or compare different question formats that do not have a common stimulus in form of a reference text or vignette (e.g., Rauch & Hartig, 2010). Although current studies assessing aspects of teacher competence use a combination of open-ended, rating-scale and multiple-choice formats (e.g., Kaiser et al., 2015), there is hardly any systematic empirical evidence on different question formats which aim at eliciting the same competence aspect while being related to the same classroom situation in a vignette-based test.

Therefore, it would be particularly insightful to compare teachers' responses to various classroom situations that are situated in similar contexts (regarding topic and grade) while each situation is represented in different vignette formats (text, comic, video) and responses are collected in open-ended and closed-ended questions. The study presented here aims at filling this gap in research by investigating (1) how different vignette formats (text, comic, video) impact on teachers' *engagement with* and *analysing of* the use of multiple representations in (2) six classroom situations, which are all situated in class 6 in the content domain of fractions, using (3) both open-ended and closed-ended questions for each classroom vignette. Corresponding research questions will be presented in the next chapter.

3 Research interest and research questions

The main research interest of this study is to investigate teachers' competence of analysing the use of multiple representations in mathematics classroom situations. Consequently, the following research questions address essential aspects related to the vignette-based assessment of such a competence and the quality of its measures.

As a vignette-based approach is chosen in this study in order to assess the competence of analysing the use of multiple representations, it is of particular importance to investigate if the participants of the study are sufficiently engaged with the classroom situations they analyse. Therefore, it should be examined how the participants perceive the implemented vignettes in terms of motivation, immersion and resonance and how they rate the vignettes regarding their authenticity (Seidel et al., 2011; Herbst, Aaron & Erickson, 2013). Accordingly, the first research question is:

1 How do the participants of the study perceive the presented classroom vignettes in terms of authenticity, immersion, motivation and resonance?

For teachers' engagement with the classroom vignettes, the format of these vignettes might play an important role. Video-based vignettes are attributed the potential to enhance teachers' engagement with presented classroom situations (Seidel et al., 2011). There is, however, little empirical evidence if comics and text vignettes are perceived comparably authentic and if they are comparably useful to engage pre-service and in-service teachers in terms of motivation, immersion and resonance. Accordingly, this study aims at answering the following research question:

2 How are different vignette formats (text, comic, video) perceived in terms of authenticity, immersion, motivation and resonance?

Since this study aims at assessing teachers' competence of analysing regarding the use of multiple representations in a newly-developed test, it should be investigated if all items measure the same construct before test scores are assigned based on the participants' analysing results. Controlling for unidimensionality can accordingly be regarded as a quality check in order to examine if the theoretical conceptualisation as a competence construct holds up empirically (Bond & Fox, 2015). Accordingly, the third research question of this study is:

3 Is it possible to model the competence of analysing regarding the use of multiple representations empirically as a unidimensional construct?

In this study, the competence of analysing multiple representations in classroom situations is described as an important profession-related competence of mathematics teachers. Teachers have to be able to identify unconnected conversions and interpret these conversions situation-specifically with respect to their role as potential learning obstacles in order to provide students with corresponding learning support. For this reason, the competence of analysing is regarded as an important prerequisite to meet professional demands related to the use of multiple representations in the mathematics classroom. It is consequently of particular interest, how teachers analyse classroom situations in a vignette-based test and whether the competence of analysing can be inferred from their analysing results. This leads to the next research question:

4 How do the participating pre-service and in-service teachers analyse the classroom situations regarding the use of multiple representations?

Competences are considered to be teachable and learnable in the framework of professional development (Weinert, 2001b; Stahnke, Schueler & Roesken-Winter, 2016) and analysing classroom situations regarding the use of multiple representations has been described as an important aspect of teacher expertise (Dreher & Kuntze, 2015a). There are consequently good reasons to assume that the competence of analysing differs amongst different levels of professionalisation. Accordingly, the presented study aims at investigating teachers at four important levels of professional development: the beginning of university studies, the beginning and end of pre-service teachers' induction phase at schools and as more experienced in-service teachers. This leads to research question number five:

5 Are there any differences in the participants' competence of analysing related to different levels of teacher professionalisation?

As outlined above, it is assumed that individual characteristics of the different classroom vignettes can have an impact on teachers' analysing regarding the use of representations. Hartig (2008) suggests that not only the content of a vignette, but also specific characteristics such as the vignette format and the question format can play a role for the difficulty of an item. Vignettes in the format text, comic and video differ, for example, with respect to their temporality, the level of individuality or the amount of context information, which might enhance or impede the analysing of the presented classroom situation. There is, however, hardly any empirical evidence on the relation between different vignette formats or question formats and teachers' analysing regarding the use of multiple representations. The aim of this study is, accordingly, to find answers to the following research questions:

6 How is the participants' analysis of the presented classroom vignettes related to different vignette formats (text, comic, video)?

7 How is the participants' analysis of the presented classroom vignettes related to different question formats (open-ended, closed-ended)?

The following chapter describes the design and methods that were used in order to find answers to these research questions.

4 Design, methods and sample

4.1 Design of the vignette-based test instrument

4.1.1 Common content and structure of the classroom vignettes

In order to find answers to the research questions of this study, a vignette-based test instrument was designed, for which in a first step nine classroom sequences were developed. The aim was to design the classroom sequences in such a way, that they would reflect professional requirements related to the use of multiple representations, in particular the requirement to identify unconnected conversions and to interpret them situation-specifically with respect to their role as potential learning obstacles. At the same time, the classroom situations should represent a variety of situations within a common content in order to enhance the comparability of corresponding analysing results. Following this argumentation, all classroom sequences were situated in the context of learning fractions in grade 6. Thus, the study can build on research by Dreher & Kuntze (e.g., 2015a, b) who chose this content focus due to the fact that the integration of multiple representations is regarded as particularly significant for students' conceptual understanding of fractions (e.g., Ball, 1993a; Charalambous & Pitta-Pantazi, 2007). The key role of multiple representations for learning fractions is also reflected in National Standards: Students are supposed to use fractions along with visual fraction models and to represent fractions and corresponding operations on the number line, in fraction bars and area models (e.g., KMK 2004; NCTM, 2000). The main reason for this is that fractions have "many faces" (Padberg, 2012, p. 32), meaning that various kinds of representations are needed to represent the many aspects encompassed in the notion of fractions, such as part-whole, operator, ratio, quotient or measure (e.g., Padberg, 2012; Dreher & Kuntze, 2015a). In order to master the tasks connected to the conceptual learning of fractions, students must consequently be able to integrate a variety of fraction representations reflecting the different aspects encompassed in the notion of fractions (Padberg, 2012). For this reason, teachers' are particularly required to identify unconnected conversions and interpret them situation-specifically with respect to their role as potential learning obstacles, which makes teachers' competence of analysing essential for meeting the professional demands in teaching fractions.

All classroom situations in the designed test instrument have a similar structure showing a teacher and his or her class working in small groups. In order to focus on the participants' analysing regarding the use of multiple representations in the classroom situations, the plots were designed on purpose in such a way that the teachers' support of the students is not in line with the theory regarding the use of multiple representations as outlined above (cf. Friesen & Kuntze, 2016; Friesen, Kuntze & Vogel, 2017): Each classroom situation starts with a teacher being asked for help by a group of students who has already started to solve a given problem using a certain register of representation (e.g., number line or decimal notation). With the aim to support the students' understanding, the teacher shifts away from the register of representation at hand, introduces an additional register and uses this register for the following explanations. The teacher fails, however, to offer any support in relating the additional register of representation to the register the students have already been using and thus exposes them to unexplained changes of representations. In doing so, several important requirements related to the use of multiple representations in the mathematics classroom are not met by the teacher and can be analysed in the designed classroom situations: First, the teacher misses the opportunity to help the students to solve their problem in the register they have already started to use. As this would have been possible in all of the situations, the teacher does not show sufficient awareness of the dilemma that introducing an additional representation is not always a help, but might also be an obstacle for students' understanding. Second, the teacher introduces the additional register of representation without relating it to the students' representation register. Thereby, the teacher does not consider the high cognitive demands related to unconnected conversions and takes the risk that the explanation impedes students' learning rather than enhancing it. Third, the teacher solves the given problem in the additional register without relating the solution back to the register the students have started with. Thereby, the students are left with a solution in a different register from the one they have started from and the teacher misses the opportunity to support them in making connections between the different representations generated by students and the teacher. As the teachers in the classroom vignettes do not meet essential professional requirements related to the use of multiple representations, the students are repeatedly exposed to unconnected and unexplained conversions and the teachers' reaction could lead to further problems in their understanding.

In this way, the vignettes focus directly on the teachers' competence of analysing the use of multiple representations as defined in section 2.6: The teachers have to identify unconnected conversions and interpret them situation-specifically with respect to their role as potential learning obstacles.

4.1.2 Validation of the classroom vignettes: expert rating

In order to explore validity after the vignette design, the classroom situations were presented to N=5 expert teachers who are not only experienced practitioners but also hold positions as teacher educators for pre-service teachers who are in their induction phase at secondary schools. Therefore, these expert teachers can be expected to be well experienced in observing and analysing classroom situations. Based on the method of cognitive interviewing (Willis, 2005), they were separately asked to analyse the designed vignettes and to evaluate the teacher's reaction to the students' question regarding the use of multiple representations in each classroom situation. In addition, the expert teachers judged the authenticity of the designed classroom situations, for example regarding the questions the students asked and the representations that were used by students and teachers. According to these expert ratings, six classroom situations were chosen for the test instrument in which the support given by the teachers was identified as potentially impeding for the students' understanding due to the unexplained change of representations as outlined above. These classroom situations were also rated as highly authentic and representative for mathematics classrooms in grade 6 by the experts.

4.1.3 The classroom vignettes of the study: an example

To provide more insight into the content and plot of the vignettes, vignette number five is shown as an example in Figure 12 and will be described in more detail in the following (cf. Friesen & Kuntze, in press).

In this classroom situation, the students struggle with converting an improper fraction into a mixed number. They have already started to solve the problem by changing registers, namely from the given register of representation (fraction number $\frac{13}{5}$) to a division (*13:5*). As they do not know how to continue, the teacher explains that they can write the remainder of the division as a fraction. As this idea involves a conversion from the division register (13:5=*2R3*) back to the fraction register ($2\frac{3}{5}$), which the students obviously are not able to carry out, the teacher introduces two further registers of representation: The problem is now represented in a real-world situation where 13 pizza slices are put together in a way that they form two whole pizzas and three slices.

PRACTICE LESSON GRADE 6: Rewriting improper fractions as mixed numbers

S1: Can you please help us here? We have a question…

(Teacher comes to the students' desk.)

S2: We want to rewrite this fraction as a mixed number. And because the fraction bar means the same as dividing, we have started to write it like this *(shows entry in their notes)*. But we don't know how to go on now…

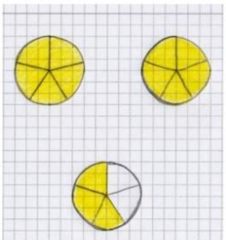
$\frac{13}{5} = 13:5$

T: Ok, we call this a division with a remainder and, as you know, you can write the remainder as a fraction.

S2: Well, honestly, I still don't get it…

T: Let's explain it with a pizza then, ok? I`ll make a quick drawing for you.

(Teacher draws and explains.)



T: Here you have 13 pizza slices. And now we put them together: five slices make one whole pizza. Then you get two pizzas and three slices are left: Two wholes and three fifths!

Figure 12: Vignette # 5 in text format

While telling the pizza story, the teacher uses another register of representation by drawing fraction circles that stand for the pizzas. The situation finishes with the teacher verbally providing a mixed number as the solution to the initial problem: "Two wholes and three fifths".

On the surface, it seems that the teachers' learning support has finally led to the correct solution of the given problem. Moreover, the teachers' idea to move away from symbolic representations to the potentially motivating pizza story and coloured fraction circles seems to be a student-oriented approach. However, analysing this classroom vignette against the theoretical background of dealing with multiple representations in the mathematics classroom as outlined above leads to a different result: Throughout the situation, the students are hardly supported in relating different registers of representations to each other when changes of representations occur. The first conversion from the fraction to the division register is initiated by the students themselves, but they cannot complete it. Instead of supporting the students in doing so, the teacher unnecessarily changes registers again by introducing the pizza story and the circular pies, however, without making any connections to the registers used before. It remains, for example, unexplained why there are thirteen pizza slices and why always five slices make one whole. The problem is finally solved by the teacher in the pizza register, again without making any connections back to the registers the students were struggling with at the beginning of the situation. The teacher does, for example, not explain, how the solution "two wholes and three fifth", which is only verbally expressed, is related to the solution in the pizza register (two whole pizzas and three slices are left) or in the division register (2R3), where fifths do not appear at all. For these reasons, the teachers' reaction can hardly be regarded as a support for students' understanding, but might rather be seen as a potential obstacle for the successful integration of multiple registers of representation in the process of students' fraction learning.

4.1.4 Different vignette formats: text, comic and video

In order to investigate the role of different vignette formats for the participants' engagement with the classroom situations and their analysis regarding the use of multiple representations, each of the six situations was implemented as text, comic and video, resulting in 18 vignettes altogether. The texts were used as blueprints to design the comics and the comics provided the storyboards for the video recordings. In order to avoid dependencies between the video vignettes, each video was recorded in another classroom showing six different teachers and their learning groups from grade 6. The teachers in the videos represent different age groups and there are male as well as female teachers. After editing the video recordings, the comics and texts were adapted, so that the conversations in the classroom situations would have the same wording and the representations used by students and teachers would look the same in each vignette format (cf. Friesen, Kuntze & Vogel, 2017; example in Figure 13).

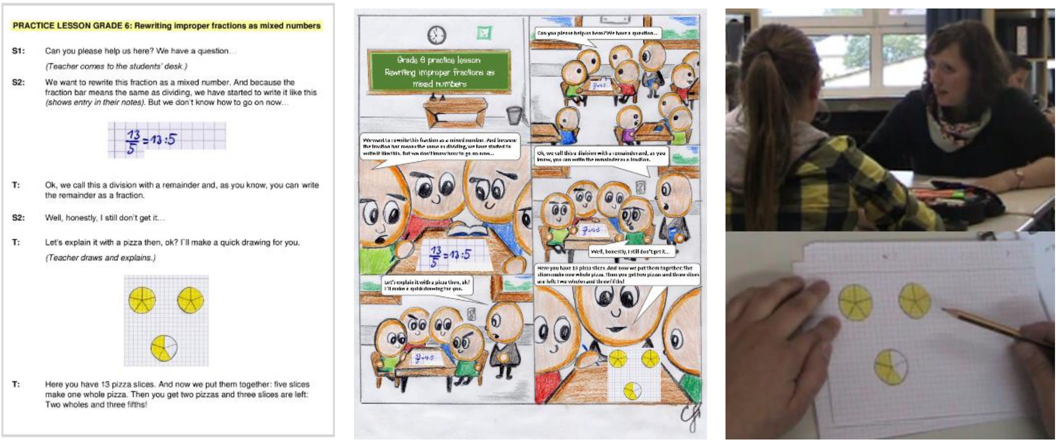


Figure 13: Vignette # 5 as text, comic and video; comic drawn by Juliana Egete

Although the corresponding text, comic and video represent the same classroom situation, they differ from each other regarding aspects of individuality, temporality, modality, the positioning of relevant information regarding representations and the amount of context information in general (Friesen & Kuntze, 2016; Friesen & Kuntze, in press). Specific characteristics regarding those aspects might, however, either facilitate or impede the test-takers' engagement with the vignettes and their analysing regarding the use of representations. Therefore, the particular characteristics and demands of the text, comic and video vignettes implemented in this study will be described in the following.

Analysing *text vignettes* requires from the test-takers to get engaged with the classroom situations although only basic information is given. Providing only little individuality regarding the characters and the classroom, the text format might on the one hand facilitate the test-takers' resonance to own teaching experiences and help to focus the analysis on the use of representations. The reduced amount of context information could, on the other hand, make it difficult for the test-takers to immerse into the situations and could make the situations look less authentic. As each participant can read the text vignettes according to his or her own speed, stop or go back to read more carefully, the temporality of a real classroom is not preserved and analysing could become easier. In addition, the representations used by the students and the teachers clearly stand out from the surrounding text lines and relevant information might thus be more accessible for analysis.

The *video vignettes* provide in contrast high individuality as they show specific students, teachers and classrooms which might contribute to the perceived authenticity of the situations. Video might also be particularly motivating for the test-takers. However, video vignettes provide the observers with a large amount of context information that could also impede the analysis regarding the use of representations. When the teacher explains and draws at the same time, visual and acoustic information has to be perceived simultaneously with a temporality close to a real classroom situation. Analysing video vignettes requires from the test-takers to extract the relevant information regarding the use of representations from the moving picture without being distracted by the large amount of context information. Adding to this potential difficulty, the representations used by the students and the teacher cannot be viewed at the same time.

Comic vignettes are close to texts regarding the aspect of temporality but can also be compared to video where the aspect of individuality is concerned. The comic classrooms have some basic furnishing and the comic characters allow to observe facial expressions, which could enhance the test-takers' immersion into the situations. However, in contrast to video, the individuality of the characters is reduced to a certain amount so that, for example, girls and boys cannot be distinguished and the comic characters look very similar in the six different vignettes. Analysing comic vignettes requires from the test-takers to connect graphical elements (comic storyboard, depicted representations) to the text in the speech bubbles in order to make sense of the classroom situation. The reduced amount of context information in contrast to the videos might, on the one hand, help to focus the analysis on the use of multiple representations. However, compared to the text vignettes, the representations used by students and teachers merge to a large extent with the comic background, which could, on the other hand, make the analysis more difficult.

4.1.5 The test design: a multiple matrix design with test booklets

In order to investigate the role of the different vignette formats on the test-takers' engagement with the vignettes and on their analysing regarding the use of representations, a multiple matrix design comprising of six test booklets was applied (see Table 1). Each booklet included the six classroom situations and always two situations were implemented in the same vignette format (text, comic or video).

classroom situation	booklet 1	booklet 2	booklet 3	booklet 4	booklet 5	booklet 6
1	T	T	C	C	V	V
2	C	C	V	V	T	T
3	V	V	T	T	C	C
4	T	V	V	C	C	T
5	C	T	T	V	V	C
6	V	C	C	T	T	V

Table 1: Multiple matrix booklet design (T ≙ text, C ≙ comic, V ≙ video)

The booklets contained explanations on how to start the videos on a laptop or tablet computer when the participants were asked to analyse a classroom situation in video format (see also section 4.1.7 on the administration of the test instrument). The links amongst the booklets can be seen in Table 1: always two booklets were linked to each other by sharing a cluster of three vignettes in the same format. Thus, a balanced distribution of the six situations in the three formats could be achieved.

4.1.6 Different question formats: open-ended and closed-ended

In order to examine the role of different question formats for the test-takers' analysis regarding the use of representations in the classroom vignettes, each of the six situations was followed by the same open-ended question and four rating-scale items (see Table 2). The open-ended question explicitly prompted the test-takers to evaluate the use of multiple representations in the classroom situations by taking into account both the teacher ("the teacher's response") and the students ("in order to help the students"). The participants were also asked to give reasons for their answers as the analysing results are considered to be an indicator for the competence of analysing under investigation. In particular, the analysing results are assumed to give insight if the test-takers are able to identify the unconnected conversions in the classroom situation caused by the teacher's change of representations and interpret them as potentially obstructing for the students' understanding.

In addition to the open-ended question, the participants rated four statements on the use of multiple representations for each classroom situation using a six-point Likert scale (1 = *I strongly disagree*; 6 = *I strongly agree*). In order to avoid that the participants are provided with correct evaluations of the classroom situations, all rating-scale statements describe on purpose the teachers' reaction as positive and supportive with respect to the students' learning with multiple representations (see Table 2).

Question format	Sample items
open-ended	*How appropriate is the teacher's response in order to help the students? Please evaluate regarding the use of representations and give reasons for your answer.*
rating-scale	*By using an additional representation, the teacher supports the students' understanding.* *The teacher's approach fosters a flexible use of representations.* *Through the help of the teacher, the students learn to use different representations.* *The explanations of the teacher connect the different representations in a learner-oriented way.*

Table 2: Open-ended question and rating-scale statements related to the use of multiple representations

With the aim to investigate how authentic the participants found a given classroom situation and how they perceived their motivation, immersion and resonance when dealing with it, they were asked to evaluate their engagement with each vignette. Therefore, the test-takers evaluated four statements according to a six-point Likert scale (1 = *I strongly disagree*; 6 = *I strongly agree*) after analysing a classroom situation regarding the use of representations (see Table 3).

Engagement (in terms of)	Sample items
authenticity	*The classroom situation appeared as authentic to me.*
immersion	*I felt part of the situation, as if I had been there in the classroom.*
motivation	*I found it motivating to deal with the classroom situation.*
resonance	*Dealing with the situation made me think of my own classroom experience.*

Table 3: Rating-scale statements related to the engagement with the vignettes (cf. Seidel et al., 2011)

4.1.7 Administration of the test instrument

The test booklets were administered as paper-and-pencil tests and each booklet started with a short introduction into the setting of the classroom vignettes (group work in the context of fraction learning in grade 6). This introduction was followed by a short explanation of the notion *representation* in the context of teaching and learning mathematics in order to ensure a similar understanding of all participants. The test booklets were randomly administered to the participants. Each test-taker was provided with earphones and a laptop or tablet computer with the two video vignettes included in his or her test booklet. It was possible to pause the videos or watch them again. Thus, each participant was enabled to analyse the six classroom situations according to his or her individual pace of work, which took around 60 to 80 minutes.

4.2 The sample

The sample of the study consisted of N = 298 pre-service and in-service teachers comprising four different stages of teacher professionalisation:

- n = 162 student teachers at the beginning of their professional education at university (66.9 % female; M_{age} = 21.6, SD_{age} = 2.4; $M_{semester}$ = 1.8, $SD_{semester}$ = 1.4)
- n = 84 pre-service teachers at the beginning of their 18-month induction phase at secondary schools (65.1 % female; M_{age} = 26.3, SD_{age} = 3.6)

- n = 30 pre-service teachers at the end of their 18-month induction phase at secondary schools (60.0 % female; M_{age} = 28.3, SD_{age} = 3.9)
- n = 22 in-service teachers (63.6 % female; M_{age} = 38.3, SD_{age} = 9.6; experience in teaching mathematics in secondary schools between two and forty years, M_{exp} = 11.1; SD_{exp} = 8.8)

All student teachers were enrolled in courses for teaching mathematics at secondary school level and came from different Universities of Education in the State of Baden-Wuerttemberg, Germany. They completed the test instrument described above during a course at their home university. The pre-service teachers completed the test instrument at the beginning of a seminar on mathematics education which was held weekly during their induction phase. The in-service teachers were contacted via e-mails to the headmasters of their schools. They completed the test instrument after classes at their schools in the presence of the author of this study.

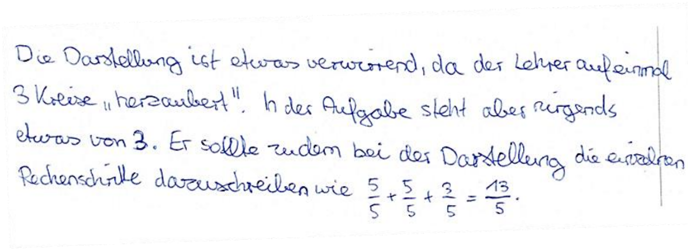
4.3 Data analysis

In the following sections, the coding of the participants' answers will be described in detail. In view of the research interest regarding the impact of different question formats on the participants' analysis of the classroom situations, the answers to both, open-ended and closed-ended questions were coded dichotomously in order to enhance the comparability of the analysing results. Thereby, this study can also build on research by Dreher & Kuntze (e.g., 2015a) who chose a similar dichotomous coding for the analysis of answers to open-ended questions.

4.3.1 Coding of the answers to the open-ended questions

The answers to the open-ended questions (cf. section 4.1.6) were coded by two independent raters reaching a good inter-rater reliability with $\kappa \geq 0.86$ (Cohen's kappa). The coding resulted in a maximum score of six points in order to make it comparable to the coding of the rating-scale items (cf. section 4.3.2). The top-down coding scheme was derived from the theoretical framework related to the use of multiple representations as outlined above (see chapter 2). As the six classroom situations of the test instrument show unexplained and unconnected conversions, the answers were coded with respect to the participants' corresponding analysis. An illustration of the coding and corresponding sample answers are shown below. All examples refer to classroom situation number five which was described in detail in section 4.1.3.

According to the theoretical framework of the study, *code 1* was assigned to answers indicating that a participant has identified the unconnected conversion introduced by the teacher's reaction to the students' question and has interpreted it as potentially obstructing for the students' understanding as the change of registers remains unconnected and unexplained (see Figure 14 and 15 for coding samples). Corresponding answers could also include positive aspects of the teacher's reaction, such as a positive evaluation of the register of representation used for the explanation or suggestions how the teacher could better support the students in making connections between the different representation registers (see Figure 16 for an example).

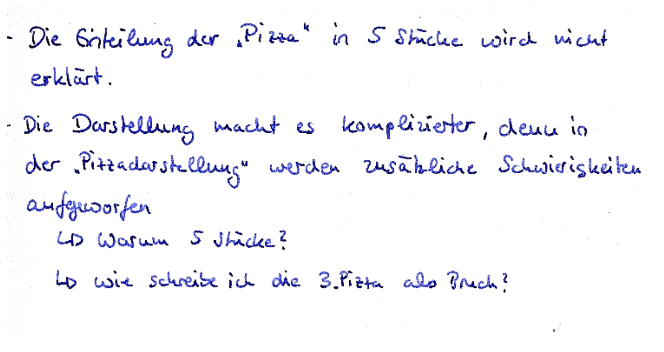

Die Darstellung ist etwas verwirrend, da der Lehrer aufeinmal 3 Kreise „herzaubert". In der Aufgabe steht aber nirgends etwas von 3. Er sollte zudem bei der Darstellung die einzelnen Rechenschritte dazuschreiben wie $\frac{5}{5}+\frac{5}{5}+\frac{3}{5}=\frac{13}{5}$.

The (graphical) representation is somewhat confusing because the teacher comes up with 3 circles "as if by magic". There is, however, no "3" in the task. He (the teacher) should also write down each step next to the (graphical) representation:

$$\frac{5}{5}+\frac{5}{5}+\frac{3}{5}=\frac{13}{5}$$

Figure 14: Code 1: coding sample # 1 (student teacher A)

Code 1 was assigned to the answer of student teacher A (Figure 14) for the following reasons: The student teacher has identified the unconnected conversion introduced by the teacher in the vignette ("teacher comes up with 3 circles as if by magic"; "there is however no 3 in the task") and has interpreted it as potentially obstructing for the students' understanding ("The (graphical) representation is somewhat confusing"). The alternative strategy suggested by the student teacher ("write down each step next to the (graphical) representation) shows that an improved connection between two registers of representations (graphical and operation in the fraction register) has been regarded as necessary to support the students' understanding in the classroom situation.



- Die Einteilung der „Pizza" in 5 Stücke wird nicht erklärt.
- Die Darstellung macht es komplizierter, denn in der „Pizzadarstellung" werden zusätzliche Schwierigkeiten aufgeworfen
  ↳ Warum 5 Stücke?
  ↳ wie schreibe ich die 3. Pizza als Bruch?

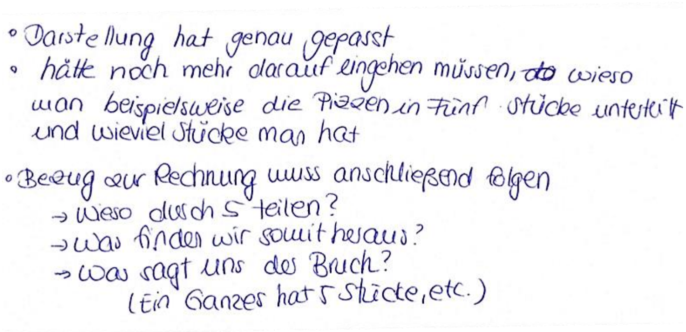
The division of the “pizza” into 5 pieces is not explained.

The (graphical) representation makes it more complicated, because the “pizza representation” causes additional difficulties

→ Why 5 slices?
→ How can I write the third pizza as a fraction?

Figure 15: Code 1: coding sample # 2 (pre-service teacher A)

The analysing result of pre-service teacher A (Figure 15) was also assigned code 1: The unconnected conversion introduced by the teacher in the vignette has been identified (“the pizza representation causes additional difficulties”; “Why five slices?”) and has been interpreted as potentially problematic for the students’ understanding (“The (graphical) representation makes it more complicated”). Pre-service teacher A also mentions the problem of changing from the pizza register back to the fraction register once a solution has been found (“How can I write the third pizza as a fraction?”) and has thus identified another unconnected conversion that could be problematic for the students’ understanding and problem-solving.



- Darstellung hat genau gepasst
- hätte noch mehr darauf eingehen müssen, ~~da~~ wieso man beispielsweise die Pizzen in Fünf Stücke unterteilt und wieviel Stücke man hat
- Bezug zur Rechnung muss anschließend folgen
  → wieso durch 5 teilen?
  → was finden wir somit heraus?
  → was sagt uns der Bruch?
  (Ein Ganzes hat 5 Stücke, etc.)

The (graphical) representation is perfect. However, the teacher should have explained better why, for example, the pizzas are subdivided into five pieces and how many pieces there are altogether. The connection to the calculation must follow:
→ Why divide by 5?
→ What do we find out by doing so?
→ What does the fraction tell us? (one whole pizza consists of 5 pieces, etc.)

Figure 16: Code 1: coding sample # 3 (student teacher B)

Code 1 was also given for the answer of student teacher B (Figure 16): The pizza representation is evaluated as positive in a first step of analysis (“The (graphical) representation is perfect”). However, the unconnected conversion has still been identified and has been interpreted as insufficiently explained to the students (“The teacher should have explained better why…”). It is explicitly mentioned that the pizza representation has to be better connected to the register the students have been using at the beginning of the classroom situation (“The connection to the calculation must follow”). The suggested strategy contains questions that could help students to better connect the two different registers of representations (pizza and calculation) by actively relating meaningful constituents to each other: “What does the fraction tell us? One whole (pizza) consists of five pieces”.

Code 0 was assigned to answers indicating that a participant has only concentrated on the representation used by the teacher without referring to the students’ representation (see Figure 17) or that a participant has identified the conversion without mentioning that the unexplained change of representations might be problematic for the students’ understanding (see Figures 18 - 20). Code 0 was also assigned to evaluations of the teacher’s reaction that did not show any argumentation related to the use of multiple representations (see Figure 21).

gut, anhand eines „Alltagsbeispieles“ erklärt. sehr verständlich für SuS, da sie es aus ihrem Alltag kennen und sich das gut vorstellen können.	Well explained with an “everyday example”. Easy to understand for students, because they know it from their everyday life and can easily imagine the situation.

Figure 17: Code 0: coding sample # 4 (student teacher C)

Code 0 was applied to the analysing result of student teacher C (Figure 17) for the following reasons: The student teacher gives a positive evaluation of the teachers’ explanation based on the idea that the explanation was related to the students’ everyday life (“Well explained with an everyday example”). However, the student teacher’s answer does not indicate any reference to multiple representations and the unconnected conversion introduced by the teacher in the vignette has not been identified.

Die Darstellung der Lehrperson finde ich gut. Schülerinnen und Schüler können sich Brüche mit Hife von Pizzas gut vorstellen. Die Abbildungen sind praxisnah und nicht so abstrakt.	I think the teacher’s representation is good. Pizzas can help students to get a clear idea of fractions. The drawings are close to everyday life and not too abstract.

Figure 18: Code 0: coding sample # 5 (student teacher D)

Code 0 was also assigned to the answer of student teacher D (Figure 18): Although the student teacher refers to pizzas as representations of fractions, the representation used by the teacher is not analysed in relation to the representation the students have been working with. Consequently, the unconnected conversion has not been not identified in this sample answer. The focus is on the more general idea that pizzas are helpful representations for learning fractions (“Pizzas can help students to get a clear idea of fractions”) and there is no indication of a situation-specific analysis regarding the change of representations.

Diese Darstellung eignet sich perfekt. Die SuS können am Ende die Teilstücke abzählen und bekommen so das Ergebnis. Auch der Übergang von Rechnung und Darstellung ist klar. Die SuS müssen nicht lange überlegen was sie wo hinzeichnen müssen, oder was sie wie ableiten müssen.	This representation is ideal. The students can count the slices at the end and will get the solution. The shift from the calculation to the (graphical) representation is also clear. The students don’t have to think long about what goes where or what should be divided how.

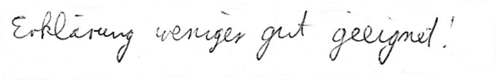
Figure 19: Code 0: coding sample # 6 (student teacher E)

Code 0 was assigned to the answer of student teacher E (Figure 19) for the following reasons: The pizza representation is positively evaluated (“This representation is ideal”). The reason given for this evaluation is that the change of representations will help students to find an answer more easily (“The students can count the slices at the end and will get the solution”). The conversion from the calculation to the (graphical) representation is identified and explicitly mentioned, however, it is interpreted as unproblematic for students’ understanding (“The shift from the calculation to the (graphical) representation is also clear”). Accordingly, the student teacher’s analysing result does not indicate that the identified conversion has been interpreted as potentially obstructing for students’ understanding.

* Hilfestellung „Pizza“ ist sehr gut (eigene Erfahrung) → SuS sind motiviert	“pizza” is a very good support (own experience) → motivating for students
* eine bildliche Darstellung der Pizzastücke ist sehr hilfreich	a graphical representation of the pizza slices is very helpful
• Ich finde die Hilfestellung sehr gut	I find the teacher’s support very good.

Figure 20: Code 0: coding sample # 7 (in-service teacher A)

Code 0 was also assigned to the analysing result of in-service teacher A (Figure 20): The support provided by the teacher in the vignette is evaluated as positive and helpful (“pizza is a very good support”). The in-service teacher’s own experience that working with pizza representations can enhance students’ motivation is provided as a reason for this evaluation. However, there is no situation-specific analysis of the classroom situation related to the unconnected conversion introduced by the teacher in the classroom vignette.



Explanation not really useful!

Figure 21: Code 0: coding sample # 8 (in-service teacher B)

Code 0 was also assigned to the answer of in-service teacher B (Figure 21): Although there was a rejecting evaluation of the teacher’s explanation in the shown classroom situation, there is no argumentation which indicates that this evaluation refers to the use of multiple representations.

4.3.2 Coding of the answers to the closed-ended questions

The coding of the answers to the rating-scale items (see section 4.1.6 for sample items) was also derived from the theoretical framework on the use of multiple representations as outlined above (see chapter 2): Disagreement with a provided statement (scoring 1, 2 or 3 on the Likert scale) was regarded as an indicator for a critical or rejecting evaluation of the teacher’s reaction to the students’ question and led to the assignment of code 1. Code 0 was given in the case of agreement with a provided statement (scoring 4, 5 or 6 on the Likert scale) as it was seen as indicator for an approving evaluation of the teacher’s reaction in the corresponding classroom situation.

In view of the research interest regarding the impact of different question formats on the participants’ analysis of the classroom situations, it was of particular importance to reach comparability with the dichotomous coding of the answers to the open-ended questions (see section 4.3.1). Within each vignette, the results for the rating-scale items were consequently aggregated into a dichotomous code. Code 1 was assigned to the analysis of a classroom situation if the answers to the four rating-scale items indicated only negative evaluations of the teachers’ reaction or both, rejecting and approving evaluations. Code 0 was distributed if all answers to the four related rating-scale items showed approval of the teachers' reaction in the corresponding classroom situation. This dichotomous coding reflects the coding procedure for the open-ended format and the comparability between open-ended and closed-ended items could thus be enhanced. As in the case of the coding described for the answers to the open-ended questions, the dichotomous coding for the rating-scale items described above resulted in a maximum score of six points.

4.3.3 Software-based data analysis

In this study, the data was analysed using the software SPSS Statistics 23 and the item response modeling software ConQuest 2.0 (Wu et al., 2007). When analyses of variance (ANOVAs) were conducted, the effect sizes were computed in omega squared (ω^2). This measure is described as less liable to bias than, for example, Pearson's correlation coefficient *r* and is therefore regarded to be more accurate in the case of ANOVAs (Field, 2013). Values of .01, .06 and .14 represent the benchmarks for small, medium and large effects, respectively (Field, 2013, p. 474). Although it is generally conservative, Bonferroni was used as procedure in *post hoc* tests for multiple comparisons since it provides the best control over type I error rates (Field, 2013). As recommended by Field (2013, p. 459), the Games-Howell procedure was always run in addition to Bonferroni in case population variances differ and also to account for the unequal group sizes of the subsamples in the study.

5 Results

In the following sections, the results of this study will be presented in the order of the research questions described in chapter 3.

5.1 The participants' engagement with the classroom vignettes

With regard to research question one, the participants' evaluations regarding their engagement with the vignettes were explored in a first step of analysis. Mean values between 4.1 (*SD* = 1.29, resonance to video vignettes) and 4.9 (*SD* = 0.86, authenticity of text vignettes) indicate on average positive ratings with respect to the authenticity of the vignettes and the participants' perceived immersion, motivation and resonance while dealing with the classroom situations (Figure 22).

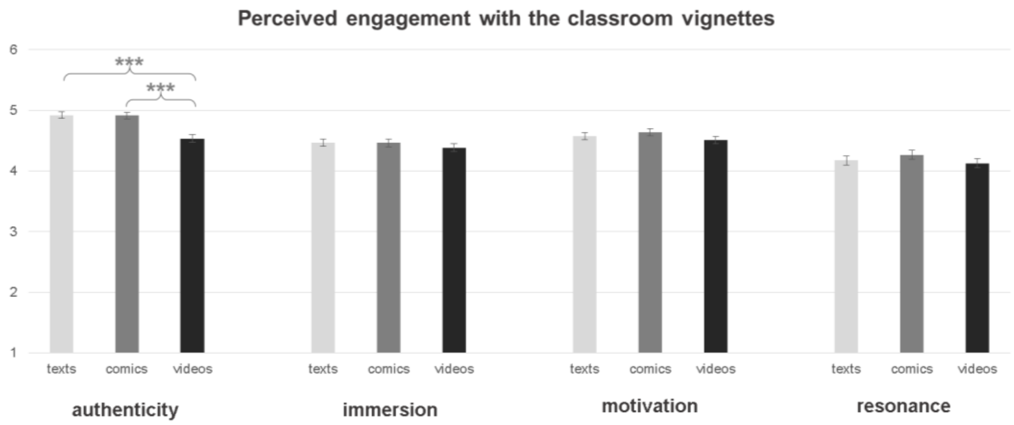


Figure 22: Participants' evaluations regarding their perceived engagement with the vignettes (means and standard errors, 1 = *I strongly disagree.* / 6 = *I strongly agree.*)

In order to investigate the possible role of the vignette format for the participants' perceived engagement (research question two), an analysis of variance (ANOVA) was conducted. It revealed a small but significant effect of the vignette format on the perceived authenticity (F (2, 878) = 15.22, p < .001, ω^2 = .03), indicating that video-based vignettes were on average rated as less authentic than texts and comic-based vignettes. No significant differences were found between the ratings of texts, comics and videos with respect to the participants' perceived immersion, motivation and resonance (see Figure 22).

In a next step of analysis, the participants' different levels of professionalisation were additionally taken into account. The conducted analysis of variance (ANOVA) showed a small but significant main effect of the level of professionalisation on the perceived *authenticity* of the classroom vignettes (F (3, 869) = 6.28, p < .001, ω^2 = .02). Bonferroni *post hoc* tests revealed that in-service teachers perceived the classroom vignettes significantly more authentic than student teachers (p < .001) and pre-service teachers at the beginning of their induction phase (p < .05). The perceived authenticity increased thus with the level of professionalisation. Figure 23 shows that the positive but lower ratings regarding the authenticity of video vignettes do not apply to in-service teachers who rated video-based vignettes as most authentic. With respect to the perceived authenticity, no significant interaction between the participants' level of professionalisation and the vignette format could be found.

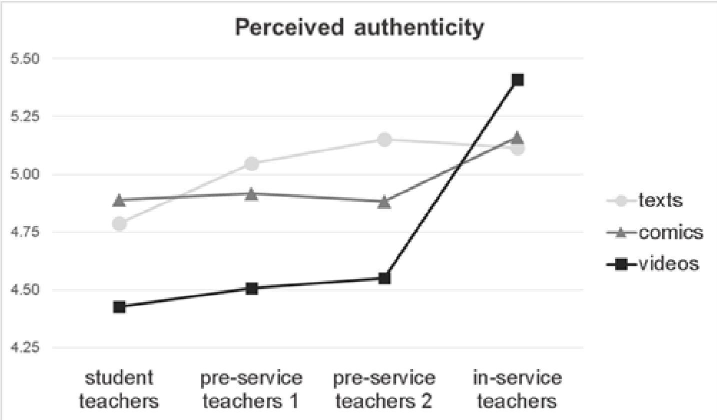


Figure 23: Perceived authenticity of the classroom vignettes (mean values, 1 = *I strongly disagree.* / 6 = *I strongly agree.*)

Data analysis with respect to *immersion* revealed no significant effect of the vignette format on how much the participants' felt part of the presented classroom situation. There was, however, a small but significant main effect of the level of professionalisation on the participants' perceived immersion (F (3, 871) = 4.33, p < .01, ω^2 = .01). Bonferroni *post hoc* tests showed that the in-service teachers perceived their immersion into the classroom situations significantly higher than the student teachers and pre-service teachers at the beginning of their induction phase at school (both ps < .05). The perceived immersion increased thus with the level of professionalisation (see Figure 24). Pre-service teachers felt most immersed when dealing with texts, while in-service teachers felt more part of a classroom situation when dealing with video-based vignettes.

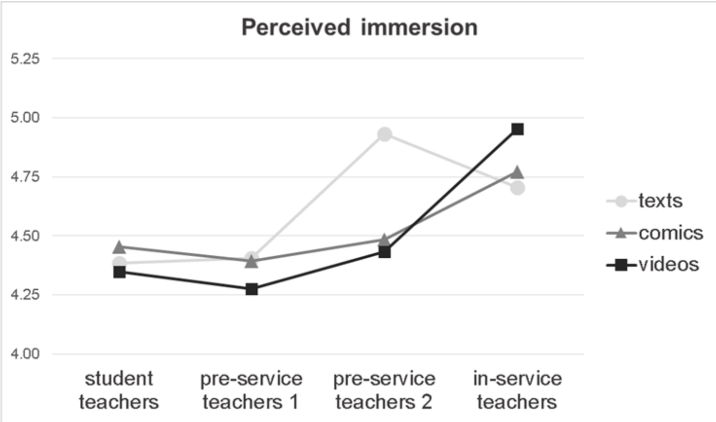


Figure 24: Participants' perceived immersion when dealing with the classroom vignettes (mean values, 1 = *I strongly disagree.* / 6 = *I strongly agree.*)

Comparable results were found regarding the perceived *resonance*: While the vignette format did not significantly affect the participants' connections to their own teaching experience, a small but significant main effect could be found in the case of the participants' level of professionalisation (F (3, 868) = 7.46, p < .001, ω^2 = .02). Bonferroni *post hoc* tests revealed that the in-service teachers perceived the resonance to own teaching experiences significantly higher than the student teachers and pre-service teachers at the beginning of their induction phase (both ps < .001). The perceived resonance increased thus with the level of professionalisation (see Figure 25). While student teachers and pre-service teachers could best connect to own teaching experiences when dealing with comics or texts, in-service teachers perceived highest resonance when dealing with video vignettes.

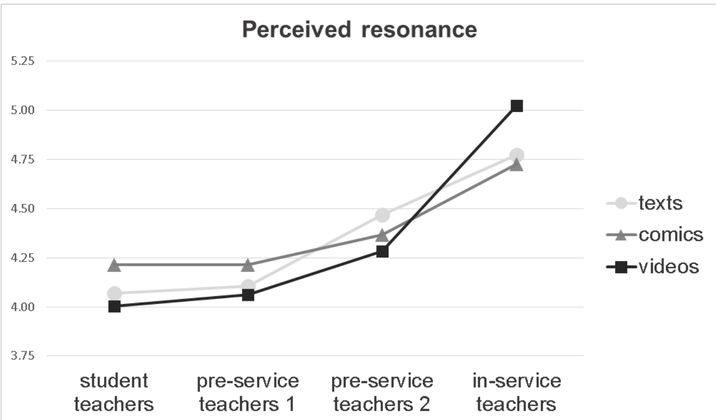


Figure 25: Participants' perceived resonance when dealing with the classroom vignettes (mean values, 1 = *I strongly disagree.* / 6 = *I strongly agree.*)

With respect to *motivation*, neither the level of professionalisation nor the vignette format showed a significant effect on how motivated the participants felt when dealing with the classroom vignettes. While student teachers and pre-service teachers felt least motivated when dealing with video vignettes, in-service teachers rated video as the most motivating vignette format (see Figure 26).

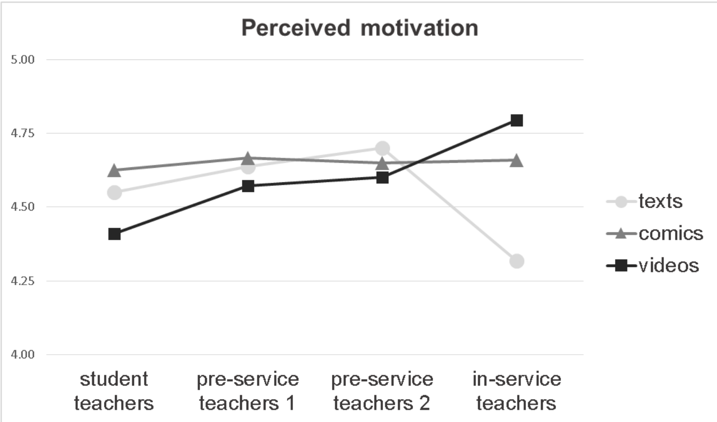


Figure 26: Participants' perceived motivation when dealing with the classroom vignettes (mean values, 1 = *I strongly disagree.* / 6 = *I strongly agree.*)

5.2 The competence of analysing as unidimensional construct

Measuring teachers' analysing regarding the use of multiple representations in the framework of competence assessment requires investigating whether the theoretical conceptualisation as a competence construct holds up empirically. In this context, the principle of *unidimensionality* states that the different items of an assessment should reflect one and the same underlying competence dimension (e.g., Segars, 1997; Blömeke, Gustafsson & Shavelson, 2015). Therefore, it has to be examined whether each item of the vignette-based test contributes in a meaningful way to the competence of analysing being investigated, thus forming a unidimensional construct. Since the *Rasch model* is based on the assumption that useful measurement involves the examination of only one human attribute at a time, it provides a mathematical framework against which data can be compared with respect to unidimensionality (Bond & Fox, 2015). For this purpose, residual-based fit statistics can be used to determine how well each item fits within the underlying test construct and whether the requirement for unidimensionality holds up empirically. The residuals represent the differences between the Rasch model's theoretical expectation of item performance and the performance actually encountered for that item in the data matrix (Bond & Fox, 2015). The usually reported fit statistics focus on two aspects of fit (infit and outfit), each of which is routinely described in both an unstandardised form (mean squares, indicating 'how *much* misfit') and a standardised form ($t\infty$ or Z, indicating how *likely* the misfit is). The unstandardised form is the average value of the squared residuals for any item, its expected value is 1. In the standardised version of the fit statistics, the mean square value is transformed to produce a statistic with a distribution just like t or Z, its expected value is 0. According to Bond & Fox (2015, p. 270), mean squares between 0.75 and 1.3 and *t*-values between -2.0 and +2.0 indicate a sufficient compatibility with the Rasch model and its requirement for unidimensionality.

Addressing the fourth research question, a Rasch model was accordingly applied to the data in order to investigate if the items of the vignette-based test reflect one and the same underlying competence dimension. Taking into account that each of the six classroom situations was responded to in three different vignette formats (text, comic, video) and two different question formats (open-ended and closed-ended), the analysis of item fit was conducted for those 36 items (*6 vignettes* x *3 vignette formats* x 2 *response formats*). This first analysis of the data revealed good fit values for all 36 items (see Table 4), indicating that each item contributes in a meaningful way to the competence of analysing being investigated and that the requirement for unidimensionality holds up empirically.

Rasch item fit statistics	mean square values	*t* - values
outfit (unweighted fit statistics)	$0.86 \leq MNSQ \leq 1.16$	$-0.8 \leq t \leq 0.9$
infit (weighted fit statistics)	$0.93 \leq wMNSQ \leq 1.09$	$-0.7 \leq t \leq 0.9$

Table 4: Item fit statistics for the Rasch analysis with 36 items

5.3 The participants' analysing of the classroom vignettes

In order to address research question four, the participants' answers were explored with respect to the frequency of the distributed codes for their analysing results (0/1). The findings show that code 0 was applied more frequently than code 1, indicating that in most of the classroom situations the participants were not able to identify and interpret the unconnected conversions as potentially obstructing for the students’ understanding (Figure 27).

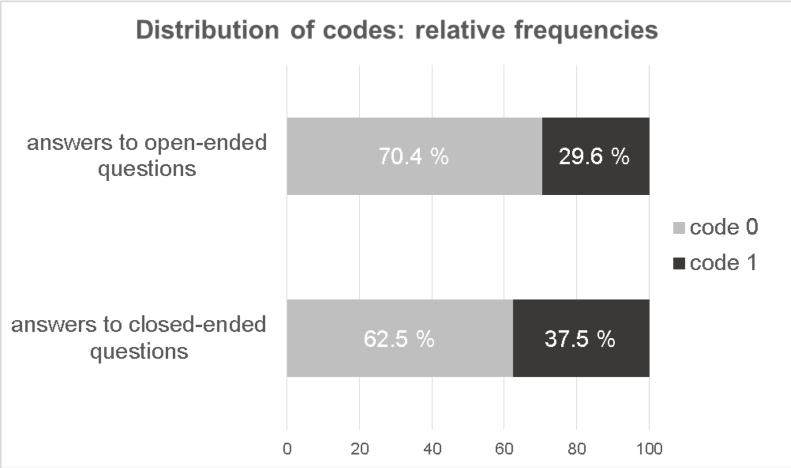


Figure 27: The participants’ analysing of the classroom situations (distribution of codes)

A comparison of the answers to the open-ended and closed-ended questions revealed that more correct analysing results were obtained in the case of the closed-ended format (37.5 %) in contrast to the open-ended format (29.6 %).

Regarding research question five, the participants’ different levels of professionalisation were additionally taken into account. The results indicate that the pre-service teachers as well as the in-service teachers reached on average higher scores than the student teachers (see Figure 28). This effect was small but significant between student teachers and pre-service teachers at the beginning of their induction phase with respect to both, the open-ended (F (3, 288) = 6.31, p < .001, ω^2 = .05) and closed-ended questions (F (3, 294) = 2.20, p < .05, ω^2 = .02). Again, a systematic difference between the analysing results of the two question formats could be observed with regard to all four levels of professionalisation: More correct analysing results were obtained in the case of responses to the closed-ended questions (see Figure 28).

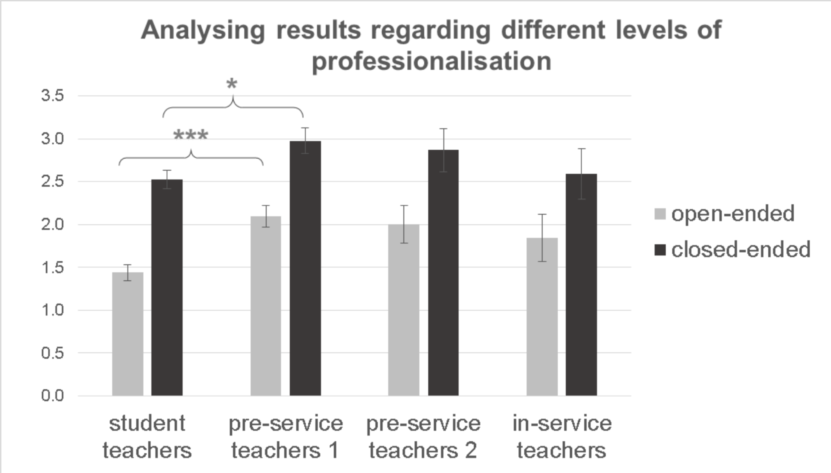


Figure 28: Analysing results regarding different levels of professionalisation (mean values and standard errors; maximum score = 6.0)

5.4 The role of the vignette format and the question format for the participants' analysing

In the following, Rasch models will be conducted to analyse the data. Therefore, the term *ability* is used to describe person abilities as estimated in Rasch models (Bond & Fox, 2015).

Specific item characteristics, such as the vignette format (text, comic, video) or the question format (open-ended, closed-ended) can play a role for the solution of an item as they can increase or decrease the demands that are required in order to solve it (Hartig, 2008). For this reason, the role of specific item characteristics should be taken into account when test scores are computed. IRT (item response theory) models are supposed to be highly applicable in this context, since they account for both, the ability of the test-takers as well as specific situational demands caused by item characteristics (Hartig et al., 2012; Blömeke, Gustafsson & Shavelson, 2015). The unique strength of the *Rasch model*, a one-parameter IRT model, can be seen in its requirement that the outcome of any interaction between person and item can be determined by the ability of the person and the difficulty of the item (Bond & Fox, 2015). Both, item difficulty and person ability are estimated in relation to each other on a common logit scale. The difficulty of an item is calculated from the total number of persons who succeeded on it, while the person ability is calculated from the total number of items to which the person responded successfully. From the difference between a person's ability and an item's difficulty, the probability of success can be computed (Rasch, 1960; Bond & Fox, 2015). While other IRT models are exploratory models requiring that the model fit the data in order to account for as much variance in the data set as possible, the Rasch model is a *confirmatory model* and requires that *the data fit the model* (Bond & Fox, 2015). Against that framework, the data of an empirical study can be tested for the presence of disturbances that may influence the estimation of item or person parameters. The Rasch model is consequently particularly useful to investigate possible relations between the participants' analysing on the one side and the vignette format and question format on the other side (research questions six and seven).

Following the argumentation above, a Rasch model with 18 items was applied to the data in a further step of analysis (*6 vignettes* x *3 vignette formats*). As the findings reported above indicate a possible effect of the question format on the participants' analysing, the data set related to the open-ended questions was analysed separately from the data set containing the answers to the closed-ended questions. First, the Rasch analysis and results with regard to the open-ended questions will be reported.

The Rasch analysis for the open-ended questions revealed good infit and outfit values for all 18 items, indicating that they sufficiently fit the Rasch model ($0.98 \leq wMNSQ \leq 1.07$, $-0.4 \leq t \leq 0.5$; $0.94 \leq MNSQ \leq 1.23$, $-0.3 \leq t \leq 1.2$). Since the participants responded to each vignette either in the format text, comic or video, two thirds of the response data is missing by design. According to Mislevy (2016), missing responses by design can be ignored in item response theory (IRT) under the condition of random assignment in multi-matrix designs, which is fulfilled by the design of this study (see section 4.1.5 and 4.1.8). However, as most traditional item statistics are only meaningful for complete designs, reliability indices (e.g., Cronbach's alpha) are not computed in ConQuest when more than 10 % of the response data is missing (Wu et al., 2007) and can consequently not be reported here. The *EAP/PV* reliability provided by ConQuest can be obtained by dividing the variance of the individual expected a posteriori ability estimates by the estimated total variance of the latent ability (Wu et al., 2007). It appeared to be rather low (0.266), which might be due to the amount of missing data by design, but could also be caused by the relatively small number of test items (Bond & Fox, 2015). Another reason might be that answering the open-ended questions of the vignette-based test was quite difficult for the participants of the study, as has already been reported in the distribution of codes for the analysis of the classroom situations (see section 5.3 and Figure 27 above). This can also be seen in the Wright map (Figure 29), a graphical representation of the estimated person and item parameters. The logit scale, which is the measurement unit common to both person ability and item difficulty, is displayed on the left of the map (Bond & Fox, 2015). In Figure 29, each X represents 0.4 cases, the number of the items are displayed on the right side of the map. Each item and person is located along the common logit scale according to its estimated value. Persons with more positive values are more able and items with more positive values are more difficult. Person locations are plotted so that a person has a 50 % probability of succeeding on an item at the same location.

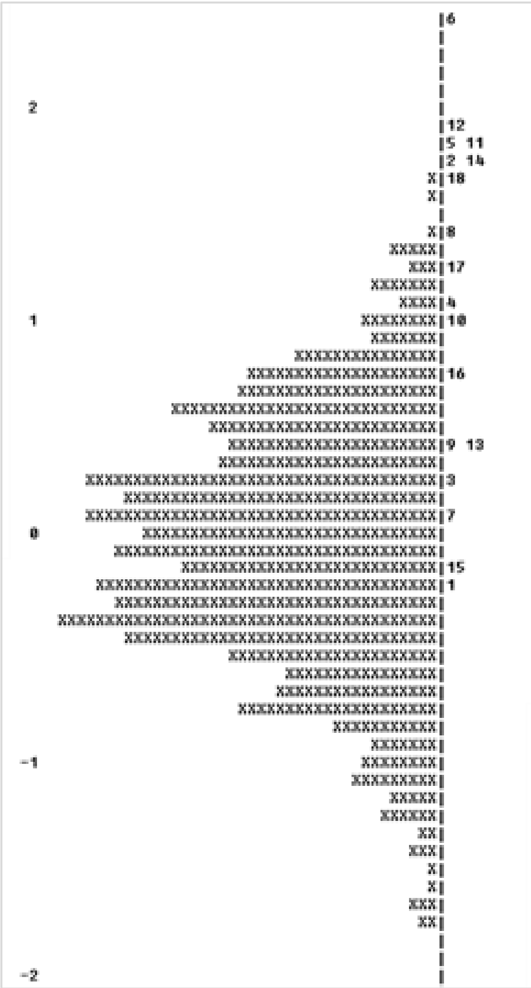


Figure 29: Wright map (for open-ended items), each X represents 0.4 cases, 1-6 ≙ text vignettes; 2-12 ≙ comic vignettes; 13-18 ≙ video vignettes

The Wright map in Figure 29 indicates that the item distribution is top heavy compared to the person distribution and that low-performers are not well targeted by the open-ended items. The items at the top of the map were thus too difficult for most of the participants (e.g., item 5, which represents the sample vignette described in detail in chapter 4.1.3). Many participants could not even succeed on the easiest items at the bottom of the map, such as items 1 and 15.

The computed item difficulties allow, however, to answer research question six, addressing possible effects of the vignette format on the participants' analysis of the classroom situations. Figure 30 shows the mean values of the item difficulties computed in the Rasch analysis described above.

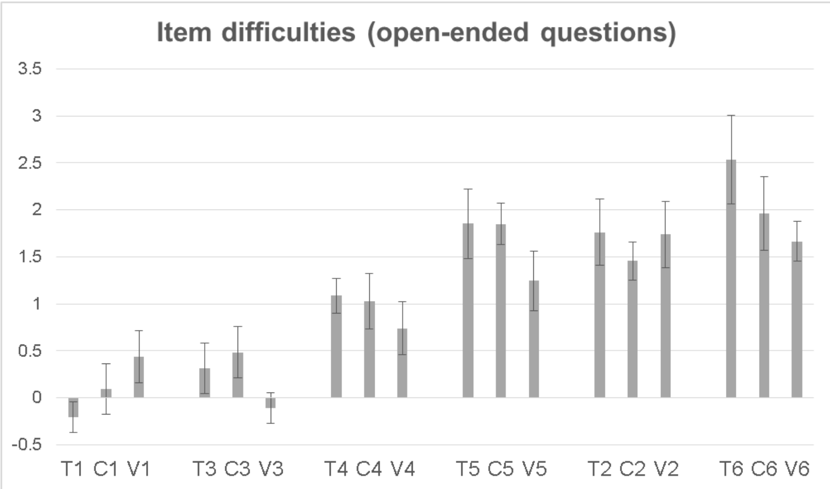


Figure 30: Item difficulties (in logits; mean values and standard errors); T ≙ text, C ≙ comic, V ≙ video; 1 - 6 ≙ number of classroom situation

The findings show that with respect to their empirical difficulty, the 18 open-ended items can be grouped by the classroom situations they represent, rather than by the different vignette formats. Low item difficulties for classroom situations 1, 3 and 4 indicate that they were more often successfully analysed regarding the use of multiple representations and therefore 'easier' than classroom situations 5, 2 and 6.

Since the logit scale is an interval-level measurement scale in which all logit units are of the same size, person and item parameters estimated by the Rasch model can be considered as interval measures (Bond & Fox, 2015). Consequently, an analysis of variance (ANOVA) could be conducted to further investigate if the different vignette formats (text, comic, video) impact on the difficulties of the items. The results of the ANOVA showed no significant effect of the vignette format on the item difficulties (F (2, 15) = 0.16, $p > .05$, $\omega^2 = .01$), indicating that the participants' analysing regarding the use of multiple representations was not systematically influenced by the vignette format a classroom situation was presented in.

There is, however, a large and highly significant effect regarding the content of the classroom situations on the item difficulties (F (5, 12) = 20.48, $p < .001$, $\omega^2 = .87$). Bonferroni *post hoc* tests show, for example, that classroom situations 1 and 3 have significantly lower item difficulties than classroom situations 5, 2 and 6 (all ps < .01). Classroom situations 5 and 6 show the highest item difficulties, indicating that most participants failed to analyse these situation successfully with respect to the use of multiple representations (see Figure 30).

In a next step, the Rasch analysis and findings with regard to the closed-ended questions will be reported. Comparable to the Rasch analysis for the answers to the open-ended questions, the analysis for the answers to the closed-ended questions revealed good infit and outfit values for all 18 items ($0.93 \leq wMNSQ \leq 1.12$, $-0.7 \leq t \leq 0.9$; $0.81 \leq MNSQ \leq 1.24$, $-1.1 \leq t \leq 1.3$). The *EAP/PV* reliability provided by ConQuest (0.402) was slightly higher than the reliability obtained for the answers to the open-ended questions. This could be due to the fact that the closed-ended questions were easier to answer for the participants of the study, as has already been indicated by the results reported above. Accordingly, the graphical representation of person abilities and item difficulties in the corresponding Wright map (see Figure 31) shows that the item distribution is more balanced compared to the person distribution and that low performers are consequently better targeted by the closed-ended items.

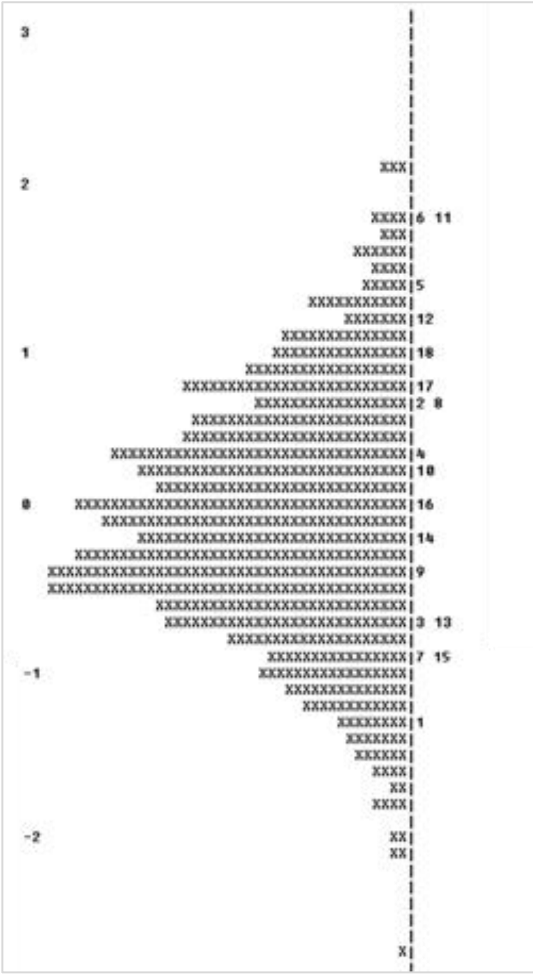


Figure 31: Wright map (for closed-ended items), each X represents 0.5 cases, 1-6 ≙ text vignettes; 2-12 ≙ comic vignettes; 13-18 ≙ video vignettes

Compared to the Wright map for the open-ended items (Figure 30), the most difficult items (item 6 and 11) lie now within the range of the person abilities and there are less participants who were not able to succeed on the easiest item (item 1).

Figure 32 shows the item difficulties computed in the Rasch analysis described above. Comparable to the findings for the open-ended questions, the items can be grouped by classroom situations and situations 1, 3 and 4 reveal lower item difficulties than situations 2, 6 and 5, respectively. With respect to research question six, an analysis of variance (ANOVA) was conducted to investigate the possible effects of the different vignette formats (text, comic, video) on the empirical item difficulties in the case of the closed-ended question format.

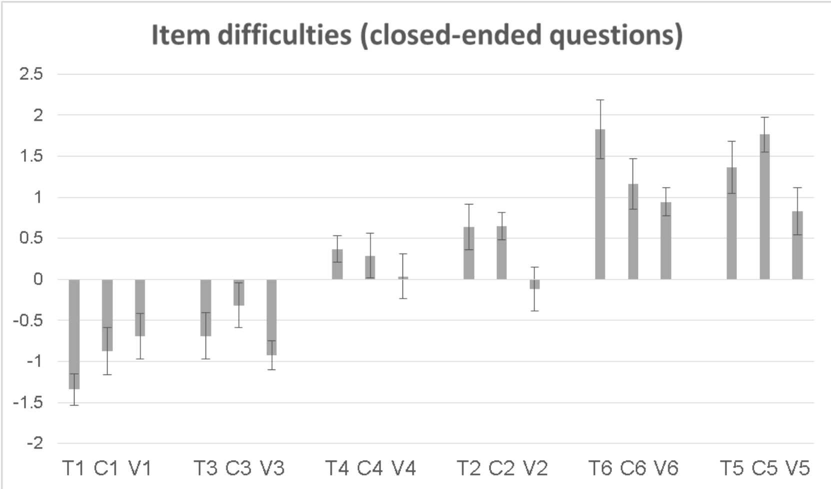


Figure 32: Item difficulties (in logits; mean values and standard errors); T ≙ text, C ≙ comic, V ≙ video; 1 - 6 ≙ number of classroom situation

The corresponding results of the ANOVA revealed no significant main effect (F (2, 15) = 0.32, p > .05, ω^2 = .01), indicating that the participants' analysis of the classroom situations was not systematically influenced by the format in which they were presented. Again there was, however, a large and highly significant effect regarding the content of the classroom situations on the item difficulties (F (5, 12) = 19.09, p < .001, ω^2 = .86). Comparable to the findings for the open-ended formats, Bonferroni *post hoc* tests show that classroom situations 1 and 3 have significantly lower item difficulties than classroom situations 5 and 6 (all *p*s < .05). Again, classroom situations 5 and 6 show the highest item difficulties, indicating that most participants failed to analyse these situation successfully with respect to the use of multiple representations (see Figure 32).

The next step of data analysis addresses research question seven. The results reported above have repeatedly given rise to the assumption that the question format (open-ended, closed-ended) has an effect on the participants' analysis. In a first approach towards this research interest, a Rasch analysis with 36 items was conducted (*6 vignettes* x 3 *vignette formats* x 2 *question formats*) in order to allow the comparability of all item difficulties. Again, traditional item statistics were not computed by ConQuest due to the amount of response data missing by design. The *EAP/PV* reliability is slightly higher (0.465) than the reliability in the case of the closed-ended question format. The findings for the item difficulties reveal that each of the 36 items has a higher item difficulty in the open-ended question format (see

Figure 33). The differences between closed-ended items and corresponding open-ended items range between 0.16 logits (for comic vignette 5) and 1.87 logits (for video vignette 2). An analysis of variance (ANOVA) showed that the question format had a large significant effect on the item difficulties (F (1, 34) = 8.79, $p < .01$, $\omega^2 = .16$), indicating that the open-ended questions (M = 1.12, SD = 0.80) were on average more difficult to analyse than the closed-ended questions (M = 0.26, SD = 0.92). The item difficulties of the open-ended and closed-ended items are, however, highly related to each other (r = .926, p = .000), as can also be seen in Figure 33.

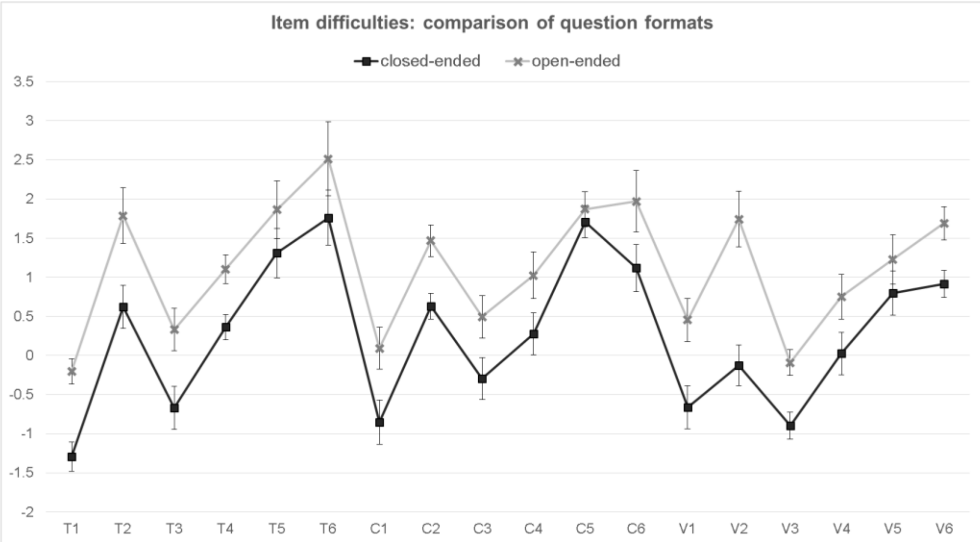


Figure 33: Item difficulties (in logits; mean values and standard errors); T ≙ text, C ≙ comic, V ≙video; 1 - 6 ≙ number of classroom situation

The common Rasch-based estimation of all 36 items, as conducted above, provides the opportunity to examine the association between empirical item difficulties and item characteristics (content, vignette format, question format) in one analysis. For this purpose, it has to be taken into account that the item difficulties are interval measures, while the number of classroom vignettes (1-6), vignette format (text, comic, video) and question format (open-ended, closed-ended) are variables on a categorical level of measurement. Consequently, η (*eta*) has to be used as a measure of association. Eta can be computed in SPSS and its values range from 0 to 1, with 0 indicating no association between the variables and values close to 1 indicating a high degree of association. The eta values displayed in Table 5 show the association between the item difficulties as dependent variable and the item characteristics as independent variables and confirm the findings reported above: The highest degree of association is revealed between the item difficulties and the classroom situations that were analysed (η = .83). Whereas also the question format (open vs. closed) appears to influence the situational demands in the vignette-based test to a large extent (η = .45), the association between item difficulties and vignette format (text, comic, video) is comparatively low (η = .15).

	classroom situation (1 - 6)	**question format** (open, closed)	**vignette format** (text, comic, video)
item difficulties	η = .83	η = .45	η = .15

Table 5: Association (eta values) between item difficulties and item characteristics

The next step of analysis involved the specification of a Rasch model that accounts for these associations in order to obtain a final estimation of the person abilities regarding the competence of analysing under investigation. According to the large significant effect regarding the content of the classroom situations on the item difficulties and the fact, that the vignette format did not play a systematic role for the participants' analysis, the six classroom situations were taken as one item each, regardless if they were responded to in the format text, comic or video. In order to take into account the large significant effect of the question format on the item difficulties, the answers to each of the six vignettes were included in both, open-ended and closed-ended question format. Therefore, the described Rasch model includes 12 items (*6 classroom situations* x *2 question formats*) in this analysis.

The existence of both open-ended and closed-ended items in the model should however be examined more closely. First, items sharing the same classroom situation as common stimulus are supposed to have higher dependencies amongst each other (Hartig & Höhler, 2009) which could potentially lead to a biased model estimation. Second, the different demands involved in answering open-ended and closed-ended questions could not only lead to significant differences in item difficulties, but could also cause items to measure different latent traits (Rauch & Hartig, 2010; Hohensinn & Kubinger, 2011). To control for such unintended multidimensionality, a Rasch model with one competence dimension as described above is compared to a Rasch model that groups the different question formats as subdimensions of the competence under investigation (cf. Hartig & Höhler, 2009). Figure 34 shows a schematic illustration of both models.

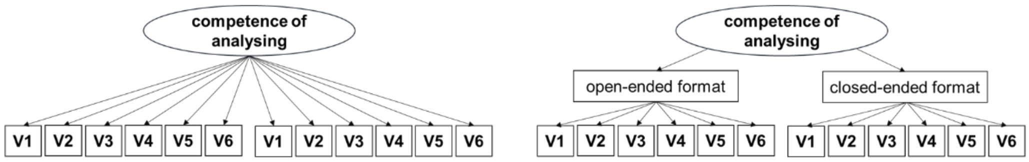


Figure 34: Schematic illustration of the two models: unidimensional model (left side) and model with two subdimensions (right side); V1-V6 ≙ classroom situations

As the models are nested, they can be compared with respect to their deviance, a statistic that indicates how well a model fits the data in comparison to another model (Wu et al., 2007). The model with two subdimensions has a slightly lower deviance than the model with one dimension (see Table 6). Comparing the value of this difference (4.49) to a chi-square distribution with 2 degrees of freedom reveals that there is no significant difference between the two models ($p > .05$).

	model with one dimension	**model with subdimensions**
deviance	4014.85	4010.36
# of estimated parameters	13	15
infit	$0.98 \leq wMNSQ \leq 1.04$ $-0.5 \leq t \leq 1.0$	$0.98 \leq wMNSQ \leq 1.03$ $-0.6 \leq \mathrm{t} \leq 0.5$
outfit	$0.97 \leq MNSQ \leq 1.04$ $-0.4 \leq t \leq 0.5$	$0.96 \leq MNSQ \leq 1.05$ $-0.4 \leq t \leq 0.6$
EAP/PV reliability	0.465	open-ended items: 0.384 closed-ended items: 0.455

Table 6: Comparison of models: model estimation statistics

Hence, it can be concluded that although the question format has a substantial influence on the difficulty of the items, open-ended and closed-ended items do still measure the same underlying competence construct. This is also supported by the item fit statistics of the unidimensional model that showed a good compatibility with the Rasch model (see Table 6), indicating that all 12 items contribute in a meaningful way to the competence construct being investigated. Since the findings for the model with subdimensions reveal also a high latent correlation between the two subdimensions for open-ended and closed-ended items (r = 0.701), the model with one competence dimension can be chosen as the preferred model.

After the stepwise analysis of the data set with respect to its compatibility to the Rasch model, the estimated person abilities will be reported. They are derived from the Rasch model with 12 items described above (see Figure 34 and Table 6 on the left side, respectively) and are calculated on the base of both open-ended and closed-ended items. The findings for the person parameters reveal that pre-service teachers and in-service teachers performed better on the test than student teachers (see Figure 35).

Lowest parameter values were found for student teachers (M = - 0.19, SD = 0.92) and highest parameter values for pre-service teachers at the beginning of their induction phase at school (M = 0.29, SD = 0.87). An analysis of variance (ANOVA) showed that there was a small but significant main effect of the participants' professionalisation level on the obtained person parameters (F (3, 294) = 5.37, p < .01, ω^2 = .05). *Post hoc* tests revealed that both groups of pre-service teachers performed significantly better on the test than student teachers (p < .01 or p < .05). Figure 35 shows also that the in-service teachers were less successful in analysing the classroom situations regarding the use of multiple representations than the pre-service teachers. This difference was, however, not significant.

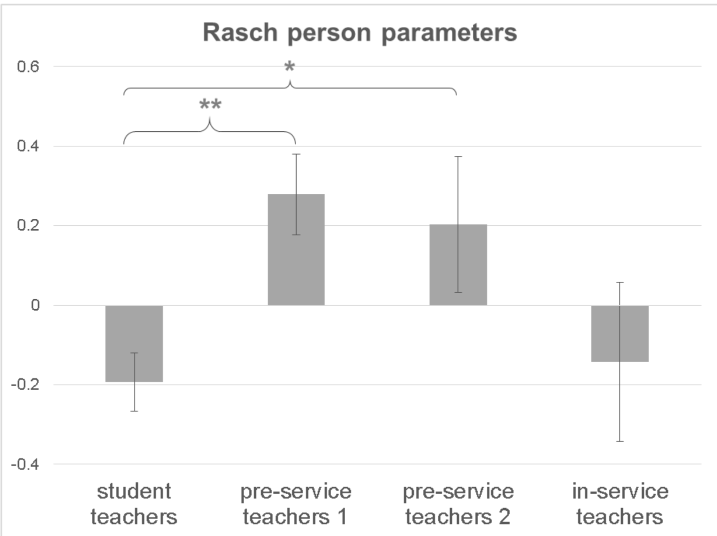


Figure 35: Rasch person parameters (mean values and standard errors in logits) grouped by the participants' different levels of professionalisation

6 Discussion and conclusions

The aim of this study is to conceptualise and assess teachers' competence of analysing the use of multiple representations in mathematics classroom situations. With this objective, a vignette-based test instrument was developed and it was investigated how different vignette formats (text, comic, video) and different question formats (open-ended, closed-ended) are related to teachers' analysing regarding the use of multiple representations. Rasch models were used to control for the quality of the competence measures.

In the following, the findings of this study will be interpreted and discussed with respect to the research questions. After considering limitations, conclusions regarding the theoretical and empirical contributions of this study will be drawn and implications for further research will be provided.

6.1 The participants' engagement with the classroom vignettes

The first and second research questions aimed at gaining insight into the participants' engagement with the classroom vignettes and how the different vignette formats (text, comic, video) are related to it. The participants' engagement with the vignettes in the test instrument was investigated with regard to the authenticity they attributed to the vignettes and considering the perceived immersion, motivation and resonance while dealing with the classroom situations. Positive mean values indicated that the participants were on average sufficiently engaged with the vignettes implemented in the test instrument. It can thus be concluded that the vignettes did not contain any characteristics impeding the participants' immersion, motivation and resonance or the perceived authenticity.

The format of the vignettes (text, comic, video) was on average not systematically related to the participants' immersion, motivation and resonance they perceived while dealing with the vignettes. However, the authenticity of the video vignettes was, although positive, on average rated significantly lower than the authenticity of the other vignette formats. This might be due to specific characteristics of the video vignettes: The high individuality in the videos might, for example, decrease the perceived authenticity when the classroom surroundings differ widely from those familiar to a participant. Taking into account the participants' different levels of professionalisation revealed a more differentiated picture of the findings: The perceived authenticity increased with the level of professionalisation and the in-service teachers rated the video vignettes as most authentic. Comparable results were found in the case of the participants' immersion: The immersion showed no systematic association with the vignette format (text, comic, video) but it increased with the participants' level of professionalisation. Accordingly, experienced practitioners felt most part of the classroom situations when dealing with the classroom vignettes. The level of professionalisation was also positively related to the participants' resonance, indicating that in-service teachers and pre-service teachers after their induction phase at schools could best connect to own classroom experience. The evaluations of the perceived motivation were positive throughout the different vignette formats and there were also no significant differences related to the participants' level of professionalisation.

With regard to the first and second research questions of this study, it can be summarised that the participants were on average sufficiently engaged with the vignettes of the test instrument. This contributes to the validity of the test instrument and can be considered as an important prerequisite for the quality of the measures. The format of the vignettes (text, comic, video) was not systematically related to the participants' immersion, motivation and resonance. The video vignettes were on average rated positive, but less authentic than comic and text vignettes. Based on these findings, it can be concluded that the texts, comics and video vignettes implemented in the test instrument of this study were comparably suitable to engage the participants in the classroom situations.

As the ratings for the authenticity of the vignettes and the participants' immersion and resonance increased with their level of professionalisation, teaching experience appears to play an important role for the teachers' engagement with the classroom vignettes. However, since also the student teachers and the pre-service teachers at the beginning of their induction phase gave positive ratings regarding these aspects, it can be concluded that all levels of professionalisation can be sufficiently engaged in the analysis of classroom vignettes.

These findings are not only in line with results reported by Seidel et al. (2011) and Herbst, Aaron and Erickson (2013), but can also add to them. Whereas Seidel et al. (2011) showed that video vignettes are particularly suitable to engage teachers into classroom situations, Herbst, Aaron and Erickson (2013) reported that both, video and animations have proved to be comparably useful in this respect. The results of the study presented here can contribute to the field of vignette-based competence assessment by showing that vignettes in the format video, text and comic are comparably suitable to engage teachers in classroom situations and that all three formats can be used for teachers at different levels of professionalisation.

6.2 The competence of analysing as unidimensional construct

The third research question aimed at finding out whether the competence of analysing regarding the use of multiple representations can be modelled as a unidimensional construct. A corresponding Rasch analysis revealed that all 36 items of the test (6 *vignettes* x 3 *vignette formats* x 2 *question formats*) contributed in a meaningful way to the underlying competence under investigation. The Rasch analysis can thus be seen as a control regarding the quality of the measures (Bond & Fox, 2015). With respect to the third research question of this study, it can be concluded that the theoretical conceptualisation of the competence of analysing holds up empirically: The competence of analysing under investigation, defined as teachers' *ability to connect classroom observations with relevant criterion knowledge regarding the use of multiple representations, so that unconnected conversions are identified and interpreted with respect to their role as potential learning obstacles* (section 2.6), could be measured as a unidimensional construct by means of the six classroom vignettes implemented in the test instrument.

6.3 The participants' analysing of the classroom vignettes

The fourth research question addressed the participants' analysing of the classroom vignettes. Based on the competence definition provided in section 2.6, teachers' competence of analysing the use of multiple representations can be inferred from an analysing result indicating that the unconnected conversion introduced by the teacher in the vignette has been identified and interpreted as potentially obstructing for students' understanding. The findings show that the participants of the study had on average a relatively low competence of analysing, since in most of the classroom situations they were not able to identify and interpret the unconnected conversions as potentially obstructing for the students' understanding. Taking into account the participants' different levels of professionalisation revealed that pre-service and in-service teachers reached on average higher test scores than student teachers. This is in line with findings by Dreher & Kuntze (2015a), who report on average low theme-specific noticing for the participants of their study but also that in-service teachers noticed the potentially obstructing conversions in the presented text vignettes more often than student teachers. Whereas Dreher & Kuntze (2015a) compared student teachers from different semesters to in-service teachers, the study presented here aimed at gaining insight into teachers' competence of analysing on four different levels of professionalisation. The findings reveal that pre-service teachers at the beginning of their induction phase reached the highest test scores and have thus shown the highest competence of analysing regarding the use of multiple representations. This difference was significant in comparison to the test scores of the student teachers at the beginning of their university studies. Pre-service teachers after their induction phase as well as in-service teachers reached comparatively lower scores. The differences were, however, not significant.

With respect to research questions four and five, it can be summarised that the participants of the study showed on average a low competence of analysing regarding the use of multiple representations. Since pre-service teachers at the beginning of their induction phase at schools showed significantly higher test scores than student teachers at the beginning of their university studies, it can be concluded that the competence of analysing can be successfully fostered in this phase of teacher education. However, the fact that also the pre-service teachers' test scores were comparatively low indicates a lack of specific university courses regarding teachers' analysing of the use of multiple representations. Since the competence of analysing the use of multiple representations is regarded as an important prerequisite in order to master corresponding demands in the mathematics classroom, the comparatively low scores of in-service teachers and of pre-service teachers after their induction phase are particularly alarming. Although dealing with multiple representations is an essential part of their daily work in the mathematics classroom, they often failed to identify and interpret unconnected conversions as potentially obstructing for the students' understanding in the vignette-based test. Given the high importance of multiple representations for the learning of mathematics, these findings call for more specific teacher development programmes also during the induction phase and in the case of experienced practitioners.

The design of this study made it possible to compare teachers' analysing results not only with respect to their level of professionalisation but also regarding the way the analysing results were articulated in different question formats (open-ended, closed-ended). The findings reveal systematic differences in teachers' analysing related to the question format: more correct analysing results were obtained in the case of the closed-ended format (37.5 %) in contrast to the open-ended format (29.6 %). It can thus be concluded that the participants' analysing regarding the use of multiple representations was not only related to their level of professional development but also to the format, in which the analysing results were articulated.

6.4 The role of the vignette format and the question format for the participants' analysing

Research questions six and seven aimed at investigating the role of the vignette format and the question format for the participants' analysing regarding the use of multiple representations. In order to find answers to these questions, Rasch models were conducted to estimate the difficulties of the items in relation to the person abilities, meaning the empirically estimated competence scores.

With respect to research question six, a comparison of the item difficulties did not reveal any significant differences between the difficulties of the items related to text, comic and video vignettes. The items could be grouped according to the six classroom situations rather than with regard to the three vignette formats. It can thus be concluded that the specific characteristics of the different vignette formats as described in section 4.1.4 were not systematically related to the participants' analysing regarding the use of multiple representations. These findings are in line with results reported by Herbst, Kosko and Erickson (2013), showing that pre-service teachers' analysing of a classroom situation was not related to the implemented vignette formats (video and animation). The findings of the study presented here can add to these results, as the multiple matrix design made it possible to compare six classroom situations in three vignette formats (text, comic, video). From the findings it can be concluded that videos, texts and comic-based vignettes were comparably suitable to elicit teachers' analysing regarding the use of multiple representations in mathematics classroom situations.

Research question seven addressed the relation between the teachers' analysing regarding the use of multiple representations and the question format in which the student teachers, pre-service and in-service teachers articulated their analysing results. The comparison of the empirical item difficulties revealed that the question format had a large and significant effect on the difficulties of the items: Each of the 36 items (6 *vignettes* x 3 *vignette formats* x 2 *question formats*) had a higher difficulty in the open-ended question format. The corresponding items in the two formats were, however, highly related to each other. The computation of the association between the item characteristics (content of the vignette, vignette format, question format) and the participants' analysing regarding the use of multiple representations indicated that the content of the classroom vignettes played the most important role for the difficulties of the items. The item difficulties were also strongly influenced by the question format, whereas the association between item difficulties and vignette format was comparatively low.

Based on the findings above, a Rasch model was conducted, which accounted for the strong influence of the six classroom situations and considered both, open-ended and closed-ended items as implemented in the vignette-based test instrument of this study. The mean values of the estimated person parameters (cf. Figure 35) can be compared to the mean values of the participants' raw scores (cf. Figure 28): Pre-service teachers at the beginning of their induction phase at schools reached the highest scores and have consequently shown best their competence of analysing the use of multiple representations. Negative and also lowest parameters were revealed for the student teachers at the beginning of teacher education. The difference to the pre-service teachers, who had already finished their university education, was highly significant. Also the pre-service teachers after their induction phase reached significantly higher person parameters than the student teachers. However, the in-service teachers' estimated parameters were negative and comparatively low. This is due to the fact that the Rasch estimation included both, open-ended and closed-ended items and the student teachers high test scores on the closed-ended items (Figure 28) were also reflected in their person parameters (Figure 35). However, it can be concluded that the high degree of similarity between the Rasch estimated person parameters and the participants' raw scores regarding the four levels of professionalisation can be taken as further indicator for the quality of the measures and contributes to the validity of the test instrument.

6.5 Limitations of the study

Although the research questions could be answered, the study has some limitations which suggest to interpret the results with care. First, the sample of the study is not representative for mathematics student teachers, pre-service teachers or in-service teachers in the state of Baden-Wuerttemberg. Although the participants were, for example, recruited from different Universities of Education and from different secondary schools, the generalisability of the results might be limited. Second, the sizes of the subsamples differ to a large extent. It was tried to compensate for unequal group sizes by choosing appropriate *post hoc* procedures for multiple comparison, but the results for the small groups should still be interpreted with care. Third, the multiple matrix design leads to a large amount of missing response data. Although IRT-based estimations are assumed to compensate for missing data by design, traditional reliability indices could not be computed. Fourth, all classroom situations were situated in the content domain of fractions. Although multiple representations play an essential role for the learning of fractions and this focus was necessary in order to enhance the compatibility of the classroom situations, the findings might be restricted to analysing the use of multiple representations in the case of fraction learning. Fifth, low-performers were not well targeted due to the high difficulty of the test items. This could also be a reason for the low EAP/PV-reliability-indices. Sixth, the video vignettes were staged videos, meaning they were especially recorded for the test instrument of this study and based on the plots of the text vignettes. Although the positive evaluations regarding the authenticity of the vignettes support their quality, they were less complex than video clips that are recorded during instruction.

6.6 Implications of the findings and need for further research

In the following section, implications from the main findings of the study will be derived and needs for further research will be outlined.

(1) It could be shown that vignettes in three formats (text, comic and video) were comparably suitable to engage student teachers, pre-service teachers and in-service teachers with the classroom situations of the test instrument. The vignettes were thus not only approved by experts but also, and even more important, by the participants of the study. Since the participants’ engagement with the vignettes can be regarded as an essential prerequisite for their analysing of the classroom situations, these findings can be regarded as an important quality check. The findings can thus contribute to the field of competence assessment: When vignette-based test instruments are developed, the specific characteristics of possible vignette formats should be investigated in order to ensure that the participants’ engagement with classroom vignettes is not impeded, for example, by low immersion or the perception that the vignettes are not sufficiently authentic. It should also be taken into account that teachers at different levels of professionalisation might differ regarding their perception or preference of vignette formats.

(2) The second main finding of this study is that teachers’ competence of analysing regarding the use of multiple representations was on average low. The results of the study give insight into the competence of analysing with respect to four important levels of teacher education: the beginning of university education, the beginning and end of the induction phase and for in-service teachers. Accordingly, it could be shown that teacher education at university appears to foster the competence of analysing regarding the use of multiple representations. However, the scores of the pre-service teachers are not only relatively low but there is also a negative trend showing that the competence of analysing is not simply enhanced by more teaching experience in the induction phase or as in-service teacher. Given the high importance of the competence of analysing multiple representations for the mastery of corresponding professional demands in the mathematics classroom, this study follows Dreher & Kuntze (2015a, b) in their call for more specific teacher professionalisation and emphasises all levels of teacher education. Vignettes as implemented in the test instrument of this study could serve as rich learning opportunities

for corresponding teacher development programmes. The role of armchair analysing is of particular importance in this context: Teachers at all levels of professionalisation can be given the opportunity to analyse classroom situations without the pressure of acting in the midst of instruction. The opportunity to foster the identification of unconnected conversions and their interpretation as potentially obstructing for students' learning in an armchair setting could thus contribute to the mastery of corresponding professional requirements in the mathematics classroom.

Follow-up research should also investigate the competence of analysing in longitudinal studies in order to gain insight into its development in the course of teacher education. Also the impact of specific teacher development opportunities with the aim to foster the competence of analysing the use of multiple representations, as described above, could be examined in this context.

(3) The third main result of this study is that teachers' analysing regarding the use of multiple representations was not related to the different formats of the classroom vignettes implemented in the test instrument (text, comic, video). It can thus be concluded that all three vignette formats were comparably effective to measure the competence of analysing the use of multiple representations in mathematics classroom situations. With respect to different question formats, it could be shown that open-ended questions were systematically more difficult to answer than closed-ended questions. These findings can contribute to the field of vignette-based competence assessment: When test instruments are developed, it should be investigated which possible vignette formats and question formats are most suitable to reflect the specific professional requirements focused in a competence assessment. It should particularly be taken into account whether specific characteristics inherent to a certain vignette format are suitable to facilitate the analysis of a presented classroom situation or whether such characteristics could also impede analysing. However, certain characteristics of vignette formats, such as a high degree of temporality or high amount of context information, might also be implemented on purpose when they are part of the professional requirement under investigation. From the findings on different question formats, it can be concluded that it is an essential question whether teachers articulate their analysing results in open-ended or closed-ended items, as closed-ended items might lead to systematically lower item difficulties. These findings are in line with findings reported above (e.g., Hartig et al., 2012),

but can also add to them as the study presented here investigated open-ended and closed-ended items that share a common vignette as stimulus in order to make the articulated analysing results comparable. In this context, open-ended items appeared to be particularly suitable for the assessment of the competence of analysing the use of multiple representations: As teachers were expected to identify unconnected conversions in the classroom vignettes and to interpret them as potentially obstructing for student' understanding, open-ended items could provide better insight into their analysing results. The implementation of open-ended items is supported by the good inter-rater reliability that could be reached in this study. Follow-up research should add to the findings of this study by analysing teachers' answers to the open-ended questions in a more qualitative approach, for instance, in order to investigate which professional knowledge the participants draw on in their argumentation (cf. Dreher & Kuntze, 2015a) or what kind of alternative strategies they suggest for the teachers' reactions in the classroom vignettes. Therefore, open-ended formats might also be suitable for assessing teachers' competence of analysing in longitudinal studies as a qualitative approach might provide insight into how this important aspect of mathematics teachers' competence develops.

7 References

Acevedo Nistal, A., van Dooren, W., Clareboot, G., Elen, J. & Verschaffel, L. (2009). Conceptualising, investigating and stimulating representational flexibility in mathe-matical problem solving and learning: a critical review. *ZDM Mathematics Education, 41*(5), 627-636.

Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education, 33*, 131-152.

Ainsworth, S. (2006). A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*, 183-198.

Ainsworth, S., Bibby, P. & Wood, D. (1998). Analysing the costs and benefits of multi-representational learning environments. In M. W. v. Someren, P. Reimann, H. P. A. Boshuizen & T. de Jong (Eds.), *Learning with multiple representations* (pp. 120-134). Amsterdam: Pergamon.

Ainsworth, S., Bibby, P. & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences, 11*(1), 25-62.

Ball, D. L. (1993a). Halves, pieces, and twoths: Constructing representational contexts in teaching fractions. In T. Carpenter, E. Fennema & T. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 157-196). Hillsdale, NJ: Erlbaum.

Ball, D. L., Thames, M.H. & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407.

Baumert, J. & Kunter, M. (2013). The COACTIV Model of Teachers' Professional Competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers. Results from the COACTIV Project* (pp. 25-48). New York: Springer.

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

Blömeke, S. (2013): The Challenges of Measurement in Higher Education: IEA's Teacher Education and Development Study in Mathematics (TEDS-M). In S. Blömeke, O. Zlatkin-Troitschanskaia, C. Kuhn & J. Fege (Eds.), *Modeling and Measuring Competencies in Higher Education. Tasks and Challenges* (pp. 93-112). Rotterdam: Sense Publishers.

Blömeke, S., Gustafsson, J.-E. & Shalvelson, R. J. (2015). Beyond dichotomies. Competence viewed as continuum. *Zeitschrift für Psychologie, 223*(1), 3-13.

Blömeke, S. & Kaiser, G. (2014). Theoretical Framework, Study Design and Main Results of TEDS-M. In S. Blömeke, F.-J. Hsieh, G. Kaiser & W. Schmidt (Eds.), *International Perspectives on Teacher Knowledge, Beliefs and Opportunities to Learn.* Dordrecht: Springer.

Blömeke, S., Kaiser, G. & Lehmann, R. (Eds.) (2010). *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich* [Professional competences and learning opportunities of future secondary mathematics teachers, an international comparison]. Münster: Waxmann.

Bodemer, D. & Faust, U. (2006). External and mental referencing of multiple representations. *Computers in Human Behavior, 22*, 27-42.

Bond, T.G. & Fox, C.M. (2015). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences*. New York: Routledge.

Borko, H. (2016). Methodological contributions to video-based studies of classroom teaching and learning: a commentary. *ZDM Mathematics Education, 48*(1), 213-218.

Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers [Competencies, functions and educational actions of teachers]. In F.E. Weinert (Ed.), *Psychologie des Unterrichts und der Schule* (pp. 177-212). Göttingen: Hogrefe.

Brovelli, D., Bölsterli, K., Rehm, M. & Wilhelm, M. (2014). Using Vignette Testing to Measure Student Science Teachers' Professional Competencies. *American Journal of Educational Research, 2*(7), 555-558.

Bruckmaier, G., Krauss, S. Blum, W. & Leiss, D. (2016). Measuring mathematics teachers' professional competence by using video clips (COACTIV video). *ZDM Mathematics Education, 48*(1), 111-124.

Bruner, J. S. (1966) *Toward a Theory of Instruction*. Cambridge, US: Harvard University Press.

Charalambous, C. Y. (2008). *Preservice teachers' mathematical knowledge for teaching and their performance in selected teaching practices: Exploring a complex relationship.* Doctoral dissertation. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/61673/chcharal_1.pdf?sequence=1&isAllowed=y

Charalambous, C. Y. & Pitta-Pantazi, D. (2007). Drawing on a theoretical model to study students' understandings of fractions. *Educational Studies in Mathematics, 64*, 293-316.

Clark, J.M. & Paivio, A. (1991). Dual Coding Theory and Education. *Educational Psychology Review, 3*, 149-210.

Cuoco, A.A. & Curcio, F.R. (Eds.) (2001). *The role of representation in school mathematics*. Boston, Virginia: NCTM.

Deliyianni, E., Gagatsis, A., Elia, I. & Panaoura, A. (2016). Representational Flexibility and Problem-Solving Ability in Fraction and Decimal Number Addition: A Structural Model. *International Journal of Science and Mathematics Education, 14*(2), 397-417.

Depaepe, F., Verschaffel, L. & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education, 34*, 12-25.

Doyle, W. (1977). Learning the classroom environment: An ecological analysis. *Journal of Teacher Education, 28*(6), 51-55.

Dreher, Anika (2015). Dealing with multiple representations in the mathematics classroom. Teachers' views, knowledge and their noticing. Doctoral dissertation. https://phbl-opus.phlb.de/files/61/ Thesis+Anika+Dreher.pdf

Dreher, A. & Kuntze, S. (2015a). Teachers' professional knowledge and noticing: The case of multiple representations in the mathematics classroom. *Educational Studies in Mathematics, 88*(1), 89-114.

Dreher, A. & Kuntze, S. (2015b). Teachers facing the dilemma of multiple representations being aid and obstacle for learning: evaluations of tasks and theme-specific noticing. *Journal für Mathematik-Didaktik, 36*(1), 23-44.

Dreher, A., Kuntze S. & Lerman, S. (2016). Why use multiple representations in the Mathematics classroom? Views of English and German Preservice Teachers. *International Journal of Science and Mathematics Education, 14*(2), 363-382.

Dreyfus, T. (2002). *Advanced mathematical thinking*. Dordrecht: Springer.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics, 61*, 103-131.

Fennema, E. & Franke, M. (1992). Teachers' knowledge and its impact. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). Reston, VA: National Council of Teachers of Mathematics.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. London: Sage.

Friesen, M., Dreher, A. & Kuntze, S. (2014). Pre-service Teachers' Noticing Focused on the Use of Representations in Classroom Videos. In S. Oesterle, C. Nicol, P. Liljedahl & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 6, p. 73). Vancouver, Canada: PME.

Friesen, M., Dreher, A. & Kuntze, S. (2015). Pre-Service Teachers' Growth in Analysing Classroom Videos. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Conference of the European Society for Research in Mathematics Education* (pp. 2783-2789). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Friesen, M. & Kuntze, S. (2016). Teacher students analyse texts, comics and video-based classroom vignettes regarding the use of representations – Does format matter? In C. Csíkos, A. Rausch & J. Szitányi (Eds.), *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 259-266). Szeged, Hungary: PME.

Friesen, M. & Kuntze, S. (in press). Competence Assessment with Representations of Practice in Text, Comic and Video Format. In O. Buchbinder & S. Kuntze (Eds.), *Mathematics Teachers Engaging with Representations of Practice*. ICME-13 Monograph.

Friesen, M., Kuntze, S. & Vogel, M. (2017). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht [Videos, texts or comics? The role of vignette format for measuring the competence of analysing the use of multiple representations in mathematics classrooms]. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Eds.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (pp. 153-177). Springer: Wiesbaden.

Goldin, G. (2008). Perspectives on representation in mathematical learning and problem solving. In L.D. English (Ed.), *Handbook of International Research in Mathematics Education*, Second Edition (pp. 176-201). Routledge: Taylor and Francis.

Goldin, G. & Shteingold, N. (2001). Systems of representation and the development of mathematical concepts. In A. A. Cuoco & F. R. Curcio (Eds.), *The role of representation in school mathematics* (pp. 1-23). Boston, Virginia: NCTM.

Goldman, R. (2007). Video representations and the perspectivity framework: epistemology, ethnography, evaluation, and ethics. In R. Goldman, R. Pea, B. Barron & S.J.

Derry (Eds.), *Video research in the learning sciences* (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum.

Haberkorn, K., Pohl, S. & Carstensen, C. (2016). Incorporating different response formats of competence tests in an IRT model. *Psychological Test and Assessment Modeling, 58*(2), 223-252

Hartig, J. (2008). Psychometric Models for the Assessment of Competencies. In J. Hartig, E. Klieme & D. Leutner (Eds.), *Assessment of Competencies in Educational Contexts* (pp. 69-90). Cambridge, USA: Hogrefe & Huber Publishers.

Hartig, J., Frey, A., Nold, G. & Klieme, E. (2012). An Application of Explanatory Item Response Modeling for Model-Based Proficiency Scaling. *Educational and Psychological Measurement*, *72*(4), 665-686.

Hartig, J. & Höhler, J. (2009). Multidimensional IRT models for the assessment of competencies. *Studies in Educational Evaluation, 35*, 57-63.

Herbst, P., Aaron, W. & Erickson, A. (2013). *How Preservice Teachers Respond to Representations of Practice: A Comparison of Animations and Video*. Paper presented at the 2013 Annual Meeting of the American Educational Research Association, San Francisco. http://hdl.handle.net/ 2027.42/97424

Herbst, P., Chazan, D., Chen, C., Chieu, V. M. & Weiss, M. (2011a). Using comics-based representations of teaching and technology to bring practice to teacher education courses. *ZDM Mathematics Education*, *43*(1), 91-103.

Herbst, P., Chazan, D., Kosko, K.W., Dimmel, J. & Erickson, A. (2016). Using multimedia questionnaires to study influences on the decisions mathematics teachers make in instructional situations. *ZDM Mathematics Education*, *48*(1), 167-183.

Herbst, P. & Kosko, K.W. (2014). Using representations of practice to elicit mathematics teachers' tacit knowledge of practice: a comparison of responses to animations and videos. *Journal of Mathematics Teacher Education*, *17*(6), 515-537.

Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371-406.

Hill, H. C., Schilling, S. G. & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, *105*, 11-30.

Hohensinn, C. & Kubinger, K. (2011): Applying Item Response Theory Methods to Examine the Impact of Different Response Formats. *Educational and Psychological Measurement, 71*(4), 732-746.

Hoth, J., Döhrmann, M., Kaiser, G., Busse, A., König, J. & Blömeke, S. (2016). Diagnostic competence of primary school mathematics teachers during classroom situations. *ZDM Mathematics Education, 48*(1), 41-54.

Janvier, C. (Ed.) (1987): *Problems of representation in the teaching and learning of mathematics*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Jacobs, V.R. & Empson, S.B. (2016). Responding to children's mathematical thinking in the moment: an emerging framework of teaching moves. *ZDM Mathematics Education, 48*(1), 185-197.

Jacobs, V.R., Lamb, L., Philipp, R. & Schappelle, B. (2011). Deciding How to Respond on the Basis of Children's Understanding. In M. G. Sherin, V. Jacobs & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 97-116). Routledge: New York.

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

Kersting, N. (2008). Using Video Clips of Mathematics Classroom Instruction as Item Prompts to Measure Teachers' Knowledge of Teaching Mathematics. *Educational and Psychological Measurement, 68*(5), 845-861.

Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R. & Stigler, J. W. (2012). Measuring usable knowledge: teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal, 49*(3), 568-589.

Kersting, N.B., Sutton, T., Kalinec-Craig, C. et al. (2016). Further exploration of the classroom video analysis (CVA) instrument as a measure of usable knowledge for teaching mathematics: taking a knowledge system perspective. *ZDM Mathematics Education, 48*(1), 97-109.

Kleinknecht, M. & Schneider, J. (2013). What do teachers think and how do they feel when they analyze videos of themselves teaching and of other teachers teaching? *Teaching and Teacher Education 33*, 13-23.

Klieme, E., Hartig, J. & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme & D. Leutner (Eds.), *Assessment of Competencies in Educational Contexts* (pp. 3-22). Cambridge, USA: Hogrefe & Huber Publishers.

Knievel, I., Lindmeier, A. M. & Heinze, A. (2015). Beyond knowledge: Measuring primary teachers' subject-specific competences in and for teaching mathematics with items based on video vignettes. *International Journal of Science and Mathematics Education, 13*(2), 309-329.

Koeppen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current Issues in Competence Modeling and Assessment. *Zeitschrift für Psychologie, 216*(2), 61-73.

Krauss, S. et al. (2013). Mathematics Teachers' Domain-Specific Professional Knowledge: Conceptualization and Test Construction in COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers. Results from the COACTIV Project* (pp.147-174). New York: Springer.

Kultusministerkonferenz (KMK) (2004). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss*. Darmstadt: betz-druck.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Eds.) (2013), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers. Results from the COACTIV Project*. New York: Springer.

Kuntze, S. (2012). Pedagogical content beliefs: global, content domain-related and situation-specific components. *Educational Studies in Mathematics, 79*(2), 273-292.

Kuntze, S., Dreher, A. & Friesen, M. (2015). Teachers' Resources in Analysing Mathematical Content and Classroom Situations – The Case of Using Multiple Representations. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Conference of the European Society for Research in Mathematics Education* (pp. 3213-3219). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Kuntze, S. & Friesen, M. (2016). Assessing pre-service teachers' competence of analysing learning support situations through a multi-format test instrument comprising of video, comic, and text vignettes. In S. Zehetmeier, B. Rösken-Winter, D. Potari & M. Ribeiro (Eds.), *Proceedings of ETC 3: ERME Topic Conference on Mathematics Teaching, Resources and Teacher Professional Development* (pp. 36-45). Berlin: Humboldt University.

Lande, E. & Mesa, V. (2016). Instructional decision-making and agency of community college mathematics faculty. *ZDM Mathematics Education, 48*(1), 199-212.

Leder, G.C., Pehkonen, E. & Törner, G. (Eds.) (2002): *Beliefs: A Hidden Variable in Mathematics Education?* Dordrecht: Kluwer Academic Publishers.

Leinhardt, G. (2001). Instructional explanations: A commonplace for teaching and location for contrast. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 333-357).Washington, DC, USA: American Educational Research Association.

Lesh, R., Post, T. & Behr, M. (1987). Representations and Translations among Representations in Mathematics Learning and Problem Solving. In C. Janvier (Ed.), *Problems of Representations in the Teaching and Learning of Mathematics* (pp. 33-40). Hillsdale, NJ: Lawrence Erlbaum.

Lindmeier, A. (2011). *Modeling and Measuring Knowledge and Competencies of Teachers. A Threefold Domain-Specific Structure Model for Mathematics*. Muenster: Waxmann.

Lindmeier, A. (2013). Video-vignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In U. Riegel & K. Macha (Eds.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* [Video-based competence research into subject-related didactics]. Fachdidaktische Forschungen (Vol. 4, pp. 45-62). Münster: Waxmann.

Mason, J. (2016). Perception, interpretation and decision-making: understanding gaps between competence and performance - a commentary. *ZDM Mathematics Education, 48*(1), 219-226.

Mayer, R.E. (2014). *The Cambridge Handbook of Multimedia Learning*. Cambridge, US: University Press.

Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth a thousand words?: Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*(3), 389-401.

Mislevy, R.J. (2016). Missing Responses in Item Response Modeling. In W.J. van der Linden (Ed.), *Handbook of Item Response Theory*. *Statistical Tools* (Vol. 2, pp. 171-195). Boca Raton: CRC Press.

Mitchell, R., Charalambous, C.Y. & Hill, H.C. (2014). Examining the task and knowledge demands needed to teach with representations. *Journal of Mathematics Teacher Education, 17,* 37-60.

Moreno, R. & Ortegano-Layne, L. (2008). Do classroom exemplars promote the application of principles in teacher education? A comparison of videos, animations, and narratives. *Educational Technology Research & Development, 56*(4), 449-465.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. http://www.nctm.org/uploadedFiles/Standards_and_Positions/Common_Core_State_Standards/Math_Standards.pdf

OECD (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing. http://dx.doi.org/10.1787/9789264190511-en

OECD (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science* (Revised edition of Vol. 1, February 2014), OECD Publishing. http://dx.doi.org/10.1787/9789264208780-en

Oser, F., Salzmann, P. & Heinzer, S. (2009). Measuring the competence-quality of vocational teachers: An advocatory approach. *Empirical Research in Vocational Education and Training, 1*, 65-83.

Padberg, F. (2012). *Didaktik der Bruchrechnung* [The didactics of fraction learning]. Heidelberg: Spektrum.

Pajares, F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research, 62*(3).

Palmer, D. J. & Stough, L.M. (2005). Identifying teacher expertise. An examination of researchers' decision-making. *Educational Psychologist, 40*(1), 13-25.

Petrou, M. & Goulding, M. (2011). Conceptualising teachers' mathematical knowledge in teaching. In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching* (pp. 9-25). Dordrecht: Springer.

Rasch, G. (1960). *Probabilistic models for some intelligence and achievement tests*. Copenhagen: Danish Institute for Educational Research.

Rauch, D. & Hartig, J. (2010). Multiple-choice versus open-ended response formats of reading test items: A two-dimensional IRT analysis. *Psychological Test and Assessment Modeling, 52*(4), 354-379.

Rehm, M., Bölsterli, K., Brovelli, D. & Wilhelm, M. (2014). Entwicklung von Unterrichtsvignetten [Development of classroom vignettes]. In D. Krüger, I. Parchmann & H. Schecker (Eds.), *Methoden in der naturwissenschaftsdidaktischen Forschung* [Research methods in science didactics] (pp. 213-225). Berlin, Heidelberg: Springer.

Renkl, A., Berthold, K., Große, C. S. & Schwonke, R. (2013). Making better use of multiple representations: How fostering metacognition can help. In Azevedo, R. & Aleven, V.

(Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 397-408). New York: Springer.

Roth, M. W. & McGinn, M.K. (1998): Inscriptions: Toward a theory of representing as social practice. *Review of Educational Research, 68*(1), 35-59.

Santagata, R. (2011). From teacher noticing to a framework for analyzing and improving classroom lessons. In M. G. Sherin, V. Jacobs & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 152-168). Routledge: New York.

Santagata, R. & Yeh, C. (2016). The role of perception, interpretation, and decision making in the development of beginning teachers' competence. *ZDM Mathematics Education, 48*(1), 153-165.

Santagata, R., Zannoni, C. & Stigler J.W. (2007). The role of lesson analysis in pre-service teacher education: an empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education, 10*(2), 123-140.

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). New York: Routledge.

Segars, A. H. (1997). Assessing the Unidimensionality of Measurement: A Paradigm and Illustration within the Context of Information Systems Research. *Omega 25*(1), 107-121

Seidel, T. & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen - Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen [How teachers perceive lessons - Assessing educational competencies by means of videos]. *Zeitschrift für Erziehungswissenschaft, 8*, 201-215.

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.

Shavelson, R. (2013). An Approach to Testing & Modeling Competence. In S. Blömeke, O. Zlatkin-Troitschanskaia, C. Kuhn & J. Fege (Eds.), *Modeling and Measuring Competencies in Higher Education. Tasks and Challenges* (pp. 29-43). Rotterdam: Sense Publishers.

Sherin, M. G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Using video in teacher education* (pp. 1-27). Bingley, UK: JAI Press.

Sherin, M.G., Jacobs, V.R. & Philipp, R.A. (2011). *Mathematics Teacher Noticing. Seeing Through Teachers' Eyes*. New York: Routledge.

Sherin, M.G., Russ, R. & Colestock, A. (2008). Professional vision in action: An exploratory study. *Issues in Teacher Education, 17*, 27-46.

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

Sherin, M.G. & van Es, E.A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*(3), 475-491.

Sherin, M.G. & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20-37.

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review, 57*, 1-22.

Sierpinska, A. (1992). On understanding the notion of function. In E. Dubinsky & G. Harel (Eds.), *The concept of function: Elements of pedagogy and epistemology*. Notes and Reports Series of the Mathematical Association of America (Vol. 25, pp. 25-58).

Spiro, R.J. & Jehng, J. (1990). Cognitive flexibility and hypertext: Theory and technology for the non- linear and multidimensional traversal of complex subject matter. In D. Nix & R. Spiro (Eds.), *Cognition, Education, and Multimedia*. Hillsdale, NJ: Erlbaum.

Stahnke, R., Schueler, S. & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: a systematic review of empirical mathematics education research. *ZDM Mathematics Education, 48*(1), 1-27.

Star, J.R. & Strickland, S.K. (2008). Learning to observe: Using video to improve preservice teachers' ability to notice. *Journal of Mathematics Teacher Education, 11*, 107-125.

Stürmer, K. & Seidel, T. (2015). Assessing Professional Vision in Teacher Candidates. Approaches to Validating the Observer Extended Research Tool. *Zeitschrift für Psychologie, 223*(1), 54-63.

Stylianou, D. (2010). Teachers' conceptions of representations in middle school mathematics. *Journal of Mathematics Teacher Education, 13*, 325-434.

Stylianou, D. A., Kenney, P. A., Silver, E. A. & Alacaci, C. (2000). Gaining insight into students' thinking through assessment tasks. *Mathematics Teaching in the Middle Grades, 6*, 136-144.

Syring, M., Bohl, T., Kleinknecht, M., Kuntze, S., Rehm, M. & Schneider, J. (2015). Videos oder Texte in der Lehrerbildung? Effekte unterschiedlicher Medien auf die kognitive Belastung und die motivational-emotionalen Prozesse beim Lernen mit Fällen [Videos or texts in teacher education? Effects of different media on cognitive load and motivational-emotional processes in case-based learning]. *Zeitschrift für Erziehungswissenschaft, 18*(4), 667-685.

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

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., et al. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam: IEA.

Tepner, O. & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens [Development of a test instrument for analysing pedagogical content knowledge]. In D. Krüger, I. Parchmann & H. Schecker (Eds.), *Methoden in der naturwissenschaftsdidaktischen Forschung* [Research methods in science didactics] (pp. 311–323). Berlin: Springer.

Törner, G. (2002). Mathematical beliefs - A search for a common ground. In G. Leder, E. Pehkonen & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 73-94). Dordrecht: Kluwer.

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

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*, 244-276.

Vogt, W. Paul (1999). *Dictionary of statistics and methodology*. Sage: Thousand Oaks, California.

Voss, T., Kunter, M. & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology*, *103*(4).

Weinert, F. E. (1999). *Concepts of Competence*. Munich: Max Planck Institute for Psychological Research.

Weinert, F. E. (2001a): Competencies and Key Competencies: Educational Perspective. In: Smelser, Neil J./Baltes, Paul B. (Eds.): *International Encyclopedia of the Social and Behavioral Sciences*. (Vol. 4, pp. 2433-2436). Amsterdam: Elsevier.

Weinert, F.E. (2001b). Concept of competence: a conceptual clarification. In D. Rychen & L. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45-65). Seattle: Hogrefe and Huber.

Willis, G.B. (2005). *Cognitive interviewing. A tool for improving questionnaire design*. Thousand Oaks, California: Sage Publications.

Wood, R. & Power, C. (1987). Aspects of the competence-performance distinction: Educational, psychological and measurement issues. *Journal of Curriculum Studies, 19*(5), 409-424.

Wu, M., Adams, R., Wilson, M. & Haldane, S. (2007). *Acer ConQuest Version 2.0. Generalised Item Response Modelling Software*. Victoria: Acer Press.

Yerushalmy, M. (1991). Student perceptions of aspects of algebraic function using multiple representation software. *Journal of Computer Assisted Learning*, *7*, 42-57.

8 Appendix

Zusammenfassung

Fragestellungen zur Konzeptualisierung und Messung professionsbezogener Kompetenzen von Lehrkräften stellen ein hochaktuelles Thema in der mathematikdidaktischen Forschung dar (z.B. Kunter et al., 2013; Kaiser et al., 2015). Trotz unterschiedlicher Ansätze in diesem Bereich besteht weitgehend Konsens darüber, dass Kompetenzen die persönlichen Voraussetzungen zur erfolgreichen Bewältigung berufsspezifischer situationaler Anforderungen beschreiben (Baumert & Kunter, 2013) und prinzipiell erlernbar und vermittelbar sind (Weinert, 2001b). Es gibt jedoch keine „Kompetenz“ per se, da die Beschreibung einer solchen stets einen relevanten berufsspezifischen Kontext voraussetzt (Hartig, 2008). Der Ausgangspunkt zur Beschreibung und Definition einer professionsbezogenen Kompetenz für Lehrkräfte sind folglich die beruflichen Anforderungen, die Lehrkräfte erfüllen müssen, um in Interaktion mit den Schülerinnen und Schülern die Lerngelegenheiten bereitzustellen, die verständnisvolle Lernprozesse ermöglichen (Lindmeier, 2011; Koeppen et al., 2008; Baumert & Kunter, 2013). Für Mathematiklehrkräfte wurde der *Umgang mit vielfältigen Darstellungen* als eine solche zentrale Anforderung beschrieben (Hill, Schilling & Ball, 2004; Ball, Thames & Phelps, 2008). Zahlreiche Studien weisen darauf hin, dass hierbei die *Wechsel zwischen unterschiedlichen Darstellungsformen* komplexe kognitive Prozesse erfordern und oftmals für viele Schülerinnen und Schüler zu einer Lernhürde werden (z.B. Ainsworth, Bibby & Wood, 1998; Duval, 2006; Ainsworth, 2006). Lehrkräfte benötigen daher spezifisches Wissen in diesem Bereich, um die Lernenden bei Darstellungswechseln unterstützen zu können (Duval, 2006; Mitchell, Charalambous & Hill, 2014; Dreher & Kuntze, 2015a, b). Lehrkräfte müssen jedoch auch in der Lage sein, Unterrichtssituationen zum Umgang mit Darstellungen zu *analysieren*, also Beobachtungen in Unterrichtsituationen mit diesem Wissen zu verknüpfen, um potentiell schwierige Darstellungswechsel erkennen zu können (Friesen, Dreher & Kuntze, 2015; Friesen & Kuntze, 2016). Es besteht weitgehend Konsens darüber, dass diese *Analyse von Unterrichtssituationen* eine wesentliche Voraussetzung dafür darstellt, dass Lehrkräfte überhaupt passende Lernangebote und Hilfestellungen zur Verfügung stellen können (z.B. Sherin, Jacobs & Philipps, 2011; Schoenfeld,

2011; Santagata & Yeh, 2016). Dennoch bleibt in aktuellen Konzeptualisierungen professionsbezogener Kompetenzen von Lehrkräften das Analysieren von Unterrichtssituationen im Hinblick auf *potentiell hinderliche Darstellungswechsel* weitgehend unberücksichtigt (z.B. Baumert & Kunter, 2013; Kaiser et al., 2015). Im Rahmen dieser Studie wird daher ein solches fachdidaktisches Analysieren von Unterrichtssituationen als wichtige professionsbezogene Kompetenz von Mathematiklehrkräften beschrieben. Da es bislang kaum empirische Studien gibt, in denen eine solche Kompetenz untersucht wurde, soll somit auch ein Beitrag zur Messung fachdidaktischer Analysekompetenz geleistet werden.

Um Kompetenzen von Lehrkräften unterrichtsnah zu erfassen, gelten vignettenbasierte Erhebungen als besonders geeignet (Kaiser et al., 2015; Blömeke, Gustafsson & Shavelson, 2015). Entsprechend wurde im Rahmen dieser Studie ein vignettenbasiertes Testinstrument mit sechs Unterrichtssituationen aus dem Bereich Bruchrechnung (Klasse 6) entwickelt, in welchen der *Umgang mit Darstellungswechseln* eine zentrale Rolle spielt. Bislang gibt es wenige Untersuchungen dazu, welche Rolle unterschiedliche *Vignettenformate* für die Auseinandersetzung mit Unterrichtsvignetten (z.B. für die wahrgenommene Authentizität) und die Analyse zum Umgang mit vielfältigen Darstellungen spielt, dasselbe gilt für unterschiedliche *Frageformate*. Da die spezifischen Eigenschaften unterschiedlicher Vignettenformate und Frageformate bei der Kompetenzmessung jedoch durchaus eine Auswirkung auf die Schwierigkeit der Items haben können (Hartig, 2008), sind Untersuchungen hierzu im Rahmen dieser Studie von besonderem Interesse. Um dem beschriebenen Forschungsinteresse nachzugehen, wurde jede der sechs Unterrichtssituationen im Testinstrument in drei Formaten (Text, Comic, Video) umgesetzt und offene sowie geschlossene Frageformate zur Analyse des Umgangs mit Darstellungen in den Unterrichtssituationen vorgelegt. Das beschriebene Testinstrument bearbeiteten N = 298 Lehramtsstudierende, Lehramtsanwärterinnen und Lehramtsanwärter sowie praktizierende Lehrkräfte. Die erhaltenen Daten wurden mit Raschmodellen analysiert, um die Qualität der vorgenommenen Kompetenzmessung zu prüfen (Bond & Fox, 2015).

Die Ergebnisse belegen eine gute Auseinandersetzung der Teilnehmerinnen und Teilnehmer mit den Vignetten in allen drei Formaten (Text, Comic, Video), wodurch eine wichtige Voraussetzung für die Analyse der vorgelegten Unterrichtssituation gegeben war. Es zeigte sich, dass fachdidaktische Analysekompetenz zum Umgang mit Darstellungen unabhängig von den eingesetzten Vignettenformaten (Text, Comic, Video) als eindimensionales Konstrukt modelliert werden kann. Während die drei unterschiedlichen Vignettenformate keinen systematischen Einfluss auf die Analyse der Teilnehmerinnen und Teilnehmer zum Umgang mit Darstellungen zeigten, wurde nachgewiesen, dass die Items aus den geschlossenen Formaten systematisch leichter zu beantworten waren. Die Analyseergebnisse der Teilnehmerinnen und Teilnehmer lassen auf eine eher niedrige Ausprägung fachdidaktischer Analysekompetenz zum Umgang mit vielfältigen Darstellungen schließen, da potentiell problematische Darstellungswechsel in den Unterrichtsvignetten häufig nicht erkannt wurden. Insgesamt konnte festgestellt werden, dass die drei Vignettenformate Text, Comic und Video vergleichbar zur Erhebung fachdidaktischer Analysekompetenz zum Umgang mit vielfältigen Darstellungen geeignet sind.