Teachers’ assessment competence:

Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model

Stephanie Herppich¹, Anna-Katharina Praetorius², Natalie Förster³, Inga Glogger-Frey¹, Karina Karst⁴, Detlev Leutner⁵, Lars Behrmann³, Matthias Böhmer⁶, Stefan Ufer⁷, Julia Klug⁸, Andreas Hetmanek⁹, Annika Ohle¹⁰, Ines Böhmer¹¹, Constance Karing¹², Johanna Kaiser¹³, and Anna Südkamp¹⁰

¹University of Freiburg, ²German Institute for International Educational Research, ³University of Münster, ⁴University of Mannheim, ⁵University Duisburg-Essen, ⁶University of Luxembourg, ⁷University of Munich, ⁸University of Vienna, ⁹Technical University of Munich, ¹⁰TU Dortmund University, ¹¹University of Wuppertal, ¹²University of Jena, ¹³University of Kiel

Author Note


This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SU 580/3-1. Correspondence concerning this article should be addressed to Stephanie Herppich, Department of Educational Science, University of Freiburg, 79085 Freiburg, Germany.

E-mail: stephanie.herppich@ezw.uni-freiburg.de
Highlights

- A new conceptual model on teachers’ assessment competence is introduced.
- Competence, as its base, is a learnable situation-specific cognitive disposition.
- The model integrates research on assessment processes, practices, and products.
- We discuss how its elements become measurable in a variable assessment process.
- The model aims to inspire integrative research on assessment.
conceptual model of teachers' assessment competence

cognitive disposition comprised by several competencies

dispositions

evaluation

quantifiable performance

formation of hypotheses

collection of information

election of methods, collection of data, analysis of data

yes/no

decision 1

decision 2

decision x

situation 1

situation 2

situation x

assessment

assessment

situation 1

situation 2

situation x

yes/no

decision 1

decision 2

decision x

quantifiable performance

formation of hypotheses

collection of information

election of methods, collection of data, analysis of data

yes/no

decision 1

decision 2

decision x

situation 1

situation 2

situation x

additional information needed?

goal processing of given information about students

judgement

explicit approach needed?
Teachers’ assessment competence:

Integrating knowledge-, process- and product-oriented approaches into a competence-oriented conceptual model
Abstract

In this article, we present a new model of teachers’ assessment competence. The model is based on the educational competence concept, thus defining competences to be context-specific, learnable cognitive dispositions that are needed to successfully cope with specific situations. Integrating research on assessment processes, practices, and products, we specify the range of situations our model applies to, and discuss how its elements may become involved and measurable in a variable assessment process. The model aims to inspire future integrative research on the description, explanation, prediction, and promotion of teachers’ assessments in various situations.

Keywords: Teachers’ assessment competence; Educational assessment, Competence; Teacher education
Teachers’ assessment competence:
Integrating knowledge-, process- and product-oriented approaches into a competence-oriented conceptual model

Introduction

In various situations of their daily professional lives, teachers need to assess their students to make educational decisions, for example, when planning lessons, adapting the pace of instruction, choosing assignments, giving feedback and deciding on grades, placement, and tracking (e.g., Allal, 2013; Furtak et al., 2016; Glock, Krolak-Schwerdt, Pitchen Cate, 2015; Helwig, Anderson, & Tindal, 2001; Praetorius, Berner, Zeinz, Scheunpflug, & Dresel, 2013; Thiede et al., 2015; Vogt & Rogalla, 2009; Voss, Kunter & Baumert, 2011). To give some examples: First, teachers may need to decide whether they can start with a new aspect of a topic or whether the students need more practice. To make this decision, they need information about their students’ current understanding (Ruiz-Primo & Furtak, 2007). Second, teachers may have a student who does well in class discussions and homework but fails in tests. In this case, teachers should learn more about the reasons for these discrepancies to provide adequate support (Karing, Dörfler, & Artelt, 2013). Third, teachers may need to grade their students’ performance via a written exam. In this instance, they need to come up with a well-constructed test and to arrive at valid and reliable grades (Brookhart, 2011, 2013). Fourth, depending on the school system, teachers may need to recommend their students one of several secondary school tracks. To make a responsible decision, they need to consider comprehensive information about the students’ characteristics that are relevant for learning (Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006).

Accordingly, these assessments and their quality are highly important and assessment competence has been considered a foundation of teaching expertise for decades (e.g., Weinert, Schrader, & Helmke, 1990, speaking of diagnostic competence, which is used largely synonymous with assessment competence in German research). Competence is usually
understood as a complex ability construct that is relevant for mastering real-life situations (Koeppen, Hartig, Klieme, & Leutner, 2008). One major advantage of basing teacher quality research on competence constructs is that the complex interplay of different teacher characteristics and the environment is explicitly taken into account (Kunter et al., 2013). Whereas the concept of assessment competence is also emphasized in current research on (mathematics) teachers’ general professional competence (Hoth et al., 2016; see also Blömeke, Busse, Kaiser, König, & Suhl, 2016), it has not yet been conceptually modeled as a core competence of teachers in detail (cf. Philipp & Leuders, 2014). Such a conceptual model, however, would be an essential base to study teachers’ assessments within the framework of professional competence.

Several research approaches exist that deal with teachers’ assessments from different angles (e.g., assessment processes, Pit-ten Cate, Krolak-Schwerdt, & Glock, 2016; assessment practices, Black & Wiliam, 2009) but they (a) are largely unrelated to each other, (b) do not incorporate the ideas connected to the competence literature, and (c) are not comprehensive from a cognitive perspective and a measurement perspective. To close this gap our scientific network of 16 researchers from different backgrounds has developed a new conceptual model of teachers’ assessment competence (for an overview of the network’s results, see Author, 2017). Our model incorporates research on assessment products (i.e., judgments; e.g., Südkamp, Kaiser, & Möller, 2012), on assessment processes (e.g., Allal, 2013; Krolak-Schwerdt, Böhmer, & Gräsel, 2013), and practices (e.g., Black & Wiliam, 2009; MacMahon & Jones, 2015) as well as their quantification and thereby extends the focus of existing models (e.g., Black & Wiliam, 2009; Böhmer, Gräsel, Krolak-Schwerdt, Hörstermann, & Glock, in press; Karst, 2012; Klug, Bruder, Kelava, Spiel, & Schmitz, 2013; Südkamp et al., 2012; Xu & Brown, 2016). Furthermore, we based our model on a current conceptualization of professional competence. This conceptualization adds a focus on teachers’ knowledge, beliefs, and motivations (e.g., DeLuca, LaPointe-McEwan, & Luhanga, 2016a, b; Xu &
Brown, 2016) and assumes competences to be learnable and enacted measurably in classes of professional situations (Kaiser et al., 2017; Koeppen et al., 2008). Our approach thus allows for cognitively based explanations of teachers’ assessments via valid measurements in varying situations. Consequently, it also allows deriving ways to promote teachers’ assessments that draw on underlying dispositions and processes without disregarding the situatedness of assessments (Kaiser et al., 2017). These results may not be achieved based on models that take a broad account of the nature and subjects of assessment and its external conditions but consider the different personal variables, their relationships and measurement to a much lesser degree (e.g., DeLuca et al., 2016b; Xu & Brown, 2016). Our model is meant to inspire future integrative research that aims to describe, explain, predict and foster teachers’ assessments as an effective interaction with the environment.

In the remainder of this article we start by defining the subject, scope and contextual framing of our model. We then introduce a current concept of competence and illustrate how it is conducive to modeling teachers’ assessment competence. Moreover, we summarize the lines of research that need to be combined in a comprehensive model of teachers’ assessment competence. Furthermore, we introduce the model, explain how it draws on existing research, and explicate its assumptions. Finally, we comment on the implications of our model and discuss the challenges for empirical research based on the model.

**Subject, Scope, and Context for a Model of Teachers’ Assessment Competence**

We started the development of our model by defining the *subject of teachers’ assessment competence*. Thus, based on existing definitions, we defined *educational assessment* as the process of assessing school students with respect to those characteristics that are relevant for learning, in order to inform educational decisions (cf. DeLuca, Chavez, & Cao, 2013; DeLuca et al., 2016b; Klauer, 1978; Schrader, 1989).

This definition rests on several assumptions about educational assessments. First, the objective of educational assessments is to inform educational decisions and actions (DeLuca...
et al., 2016b; Popham, 2011). More specifically, educational assessments should primarily help teachers to improve learning (via formative assessment) and to optimize decisions around certification, placement, or selection of students (via summative assessment; Harlen, 2005). Accordingly, educational assessments take place prior to an educational decision and are subordinate to it (cf. Schrader, 1989). This definition entails that teachers’ assessments are considered to be conceptually different from teachers’ educational decisions and actions (e.g., feedback, communication of assessment results). In this regard, our approach is more specific than other approaches to teachers’ assessments (e.g., Black & Wiliam, 2009; DeLuca et al., 2016b; Klug et al., 2013). Although educational assessments are closely intertwined with decisions and actions in the daily professional practice of teachers, we assume that it is more fruitful to separate them conceptually, and consequently to model teachers’ assessment competence independently from other forms of teacher competence, such as instructional competence (cf. Weinert et al., 1990). This separation is particularly relevant for the promotion of teachers’ assessment competence, because assessing students and deciding between various alternative ways of instruction plausibly entails different cognitive processes (cf. Feldon, 2007; van Merriënboer, Clark, & de Crook, 2002). The alternative to separating both processes would be to model an undifferentiated comprehensive “adaptive instructional competence”. This competence, however, would be difficult to model theoretically and particularly difficult to teach and learn (cf. van Ophuysen, 2010).

The following example illustrates the theoretical difference we see between assessments and educational decisions: During a quiz, a teacher notices that one of her students lacks knowledge that is necessary to understand the new topic. As a consequence of her assessment, for example, she could decide to help the student by covering the lacking knowledge herself. Alternatively, she could decide to ask a fellow student to do so. This instructional decision – support by the teacher or by a fellow student – is based on the judgment lacking knowledge. At the same time, this decision is influenced by other factors.
like the teacher’s available time, her knowledge about instructional options, or by the fellow student’s competences. The teacher might arrive at an adequate judgment (assessment) but make an inappropriate decision (instruction).

Second, our definition explicates who is assessed and which constructs are assessed. The subject of assessment are all characteristics of a school student that are relevant for learning. In this vein, our approach to assessment is specific, as we neither focus on the assessment of student’s characteristics in higher education (Sadler, 2013), nor on the evaluation of tasks (Karst, 2012; McElvany et al., 2009) or the evaluation of instruction (Helmke & Helmke, 2015; cf. Fernández-Ballesteros et al., 2001 for the difference between assessment and evaluation).

Third, our definition implies that we are not only interested in teachers’ assessments of students’ achievement, although assessing achievement is central for formative- (e.g., to provide adaptive instruction) and summative purposes (e.g., for grades and certificates). Nevertheless, we assume the assessment of other characteristics, like students’ emotions, motivation, and behavior relating to learning as being equally as crucial (Givvin, Stipek, Salmon, & MacGyvers, 2001; Sbarra & Pianta, 2001). For example, the assessment of collaborative behavior might trigger specific instructional interventions (e.g., when students start to quarrel; Kändler; Wiedmann; Rummel, & Spada, 2015).

Finally, we understand assessment - the subject of teachers’ assessment competence - as a process (DeLuca et al., 2013). For the sake of terminological clarity, we will use assessment in this article to refer to both teachers’ cognition and behavior. We will use judgment processes to refer to teachers’ cognitive processes and assessment practices to refer to teachers’ behaviors associated with assessment. On a general level, assessment practices involve the development, selection, and application of methods to collect, analyze, and interpret data about students’ characteristics such as tests, interviews, and observation (Ingenkamp & Lissmann, 2008).
Teachers’ assessments as professional processes should be reliable and valid (e.g., Brookhart, 2011). An orientation towards scientific quality is particularly important if an assessment-informed educational decision has far-reaching consequences for a student (e.g., in the case of certification or tracking decisions; Lintorf, Behrmann, & van Ophuysen, 2016; McMahon & Jones, 2015). Thus, to determine the quality of teachers’ assessments, we need to consider their reliability and validity in our model of teachers’ assessment competence.

Our model is not only based on a specific definition of educational assessment. Its development took place against a Central European, that is, Austrian, Luxembourguian, and, particularly, German educational background. This context has shaped the primary focus of the model. Most importantly, in contrast to other countries, teachers’ employment status and school funding in these systems is hardly, or not at all, affected by students’ assessment results (cf., Eurydice Network, 2016; Schneider & Bodensohn, 2017). Accordingly, teachers may feel less tension between the need for assessment for optimizing learning and for student accountability, on the one hand, and the need for assessment for school accountability purposes on the other hand (cf. Brookhart, 2011; DeLuca et al., 2016b; Xu & Brown, 2016). Against this background, focusing on this tension seemed less useful to start with.

The Educational Competence Concept as a Theoretical Foundation

In their recent review on the assessment of competences in higher education, Blömeke et al. (2015) contrasted two current perspectives on competence. The first one is the holistic perspective (e.g., McClelland, 1973; Shavelson, 2010). Here, competence is equated with successful behavior in domain-specific real-life situations. The second one is the analytic perspective. In this perspective, competence is seen as a disposition that underlies a person’s behavior in domain-specific situations (e.g., Koeppen et al., 2008). This perspective was developed in educational research with the aim of fostering complex abilities and resulting behavior. To attain this goal, researchers describe, explain, and foster the characteristics underlying performance (i.e., the competence) to be able to describe, explain, and foster
adaptive behavior. Blömeke et al. (2015) took a step to reconciling these two perspectives. They suggested that the role of underlying dispositions was acknowledged from the holistic perspective and that the necessity to measure performance was recognized from the analytic perspective. Moreover, they proposed that disposition and performance were connected by processes like perception and interpretation of the specific assessment situation (see the literature on professional vision, e.g., van Es & Sherin, 2002), as well as decision-making in this situation. These processes are assumed to mediate the relationship between competence and performance (cf. Kaiser et al., 2017).

When developing our model, we followed Blömeke et al.’s (2015) suggestions. To inspire research on describing, explaining, and fostering teachers’ assessments, our primary focus was on the analytic perspective (see Knievel, Lindmeier, & Heinze, 2015; Rott, Leuders, & Stahl, 2015; Schneider & Bodensohn, 2017, for references to this perspective in models on teachers’ competences). With Koeppen et al. (2008, p. 62), we define competences1 “…as context-specific cognitive dispositions that are acquired and needed to successfully cope with certain situations or tasks in specific domains.” Several characteristics of this definition are relevant for our model. First, competences are not considered generic abilities (e.g., intelligence) but cognitive dispositions that are specifically related to certain real-life tasks. In a model of teachers’ assessment competence we are, thus, interested in those components of teachers’ knowledge that are necessary to assess students. Understanding competence as a disposition entails that it is attributed with consistency and stability (Funder, 2012). Consequently, a teacher is assumed to cope with several assessment situations in a

1 We use the term competence (plural: competences) to refer to a superordinate construct. We use the term competency (plural: competencies) as a synonym for constituent of a competence (Blömeke et al., 2015).
similar way (i.e., relatively consistently) over an extended period of time (i.e., relatively stable). At the same time, competences are deemed learnable by dealing with relevant situations (Blömeke et al., 2015; Koeppen et al., 2008). This means that teachers need to gather experience with the assessment of students in situations that are typical for the teachers’ profession to become competent in assessing students (cf. the idea of situated learning, e.g., Renkl, Mandl, & Gruber, 1996; Resnick, 1989). In addition, this means that competence may be fostered by interventions. These assumptions allow for hypotheses about beneficial approaches to teacher education as well as professional development (cf. DeLuca et al., 2013; Xu & Brown, 2016) based on teachers’ dispositions and pave the way for the development of interventions.

Inherent to this concept of competence is the idea that a competence can analytically be divided into single constituents, that is, competencies. These competencies are deemed measurable, and values on these measures are assumed to predict performance in relevant situations. Individuals can differ on their competencies and, thus, in their performances. In taking up Blömeke et al.’s (2015) suggestion, actual performance in different situations also needs to be measured to draw inferences about the realized competencies (cf. Kaiser et al., 2017).

Finally, following another pivotal suggestion by Blömeke et al. (2015), we consider teachers’ judgment processes to mediate the relationship between cognitive dispositions and performance in assessment situations, as expressed in assessment practices and judgments. Accordingly, this approach to competence integrates a cognitive stance with a focus on assessment as situated practice (Kaiser et al., 2017).

Prior Conceptualizations of Teachers’ Assessments

As elaborated above, being based on current educational conceptualizations of competence, a new model of teachers’ assessment competence should incorporate assumptions about teachers’ dispositions and cognitive processes that determine assessment
behavior in classes of professional assessment situations and that can be inferred from measured behavior in situations. Parts of these aspects have been the focus of research before, yet, they have scarcely been analyzed in combination and have not been integrated into one common model. Below, we give an overview of the most relevant lines of research.

First, research on assessment literacy has provided important insights into aspects of teachers’ knowledge that are relevant for assessing students (e.g., Brookhart, 2011; DeLuca, 2016a, b). Popham (2011, p. 267) defined assessment literacy as “an individual’s understandings of the fundamental assessment concepts and procedures”. The assessment concepts and procedures are usually drawn from standards for educational assessment (e.g., American Federation of Teachers, National Council on Measurement in Education, & National Education Association (AFT et al.), 1990; cf. DeLuca et al., 2016b). In addition, models of assessment literacy inform us how teachers’ other dispositions, such as their cognitive setup, beliefs, and affects, might influence the transformation of knowledge into action (e.g., Brookhart, 2011; Xu & Brown, 2016). Current models, moreover, view assessment literacy as a situated practice that is socially and culturally shaped, in line with a sociocultural perspective on learning (DeLuca et al., 2016a; Willis, Adie, & Klenowski, 2013; Xu & Brown, 2016). These models emphasize the relevance of context for teachers’ professional actions. Due to their broad sociocultural and descriptive account, however, they do not attempt to specify the range of situations they are relevant to. Similarly, they do not relate dispositions, cognitive processes, and performance in assessments to each other for individual teachers and do not focus on the quantification of the underlying competence. Considering the aim of studying teachers’ assessments as an effective interaction with the environment, these aspects can be added by a competence-based approach.

2 For discussions about advantages and disadvantages of cognitive approaches and situated approaches to learning see Sfard (1998; cf. Kaiser et al., 2017).
Furthermore, one line of research has contributed to the description of *assessment practices* and the evaluation of their quality. Related studies refer to practices of formative or summative assessment. Studies on formative assessment have provided insights into assessment practices that teachers can apply to foster students’ learning (e.g., Furtak et al., 2016; Herppich, Wittwer, Nückles, & Renkl, 2014). For example, Herppich et al. (2014) found that students in one-on-one tutoring learned more when tutors tended to guide their student’s learning more often with questions and hints and explained less. Studies sometimes also report on internal factors (e.g., teachers’ beliefs) and external factors (e.g., school context) that may influence the implementation of formative assessment practices (Izci, 2016). However, the concept of formative assessment is still very loosely defined and is applied to a multitude of practices. Accordingly, advice on quality assessment practices is necessarily very tentative (Bennett, 2011; Kingston & Nash, 2011).

Studies on summative assessment practices mostly inform us about their reliability and validity and about ways to promote their dependability. In these studies, teachers’ assessments are either compared with norms from measurement theory (Brookhart, 2013) or with procedures and criteria, sometimes prescribed by official standards (Black et al., 2010, 2011; Brookhart, 2013; Harlen, 2005; McMahon & Jones, 2015, Sadler, 2013; Wyatt-Smith & Klenowski, 2013). Neither research on formative nor on summative assessment practices, however, is generally concerned with the knowledge teachers need to carry out assessment practices appropriately (see Allal, 2013; Herman, Osmundson, Dai, Ringstaff, & Timms, 2015; Wyatt-Smith & Klenowski, 2013 for exceptions).

Other lines of research deal with teachers’ *judgment processes*. First, research on the quality of summative assessments usually does not strictly separate between cognition and behavior (Allal, 2013; Harlen, 2005; Sadler, 2013; Wyatt-Smith & Klenowski, 2013). Accordingly, conclusions from these studies also hold for the underlying cognitive processes (information selection, integration, and inferences, Borko et al., 2008). Second, research on
teachers’ data-based decision making is similarly interested in teachers’ cognitive processes and practices when they collect, analyze, and interpret data, use them to improve teaching and learning, and when teachers evaluate their data-based decision making (e.g., Lai & Schildkamp, 2016; Marsh & Farrell, 2015). The focus of this research is on ways to improve teachers’ data use and, partly, on underlying dispositions, not on the valid measurement of processes, practices, and results of data use.

Third, research from the social-psychological tradition studies the nature of teachers’ judgment processes with varying purposes (e.g., forming a first impression of a student, Dünnebier, Gräsel, & Krolak-Schwerdt, 2009; making tracking decisions, Böhmer et al., in press). This research draws on dual process models of decision making (Ferreira, 2006; Fiske & Neuberg, 1990) and findings about social-cognitive heuristics (e.g., anchoring effects, Dünnebier et al., 2009). It tells us how teachers arrive at those judgments that inform their educational decisions and how these processes are influenced by external factors, such as accountability for the decision (Böhmer et al., in press). The appropriateness of a process, however, is often inferred theoretically and not via an empirical criterion (Krolak-Schwerdt et al., 2013). Thus, the criterion validity of different judgment processes remains unclear. The knowledge teachers need to perform a particular process is usually not examined, either. Finally, in none of the lines of research on judgment processes is the quality of the resulting judgment as a product of assessment evaluated.

The quality of the resulting judgment, however, is studied in research on the accuracy of teachers’ judgments (Südkamp et al., 2012). Here, a teacher’s judgment is compared with a student’s result on a standardized achievement test (see Brookhart, 2013; Hoge & Coladarci, 1989; Südkamp et al., 2012, for summaries) or self-report questionnaire (e.g., Karing et al., 2013). From this tradition we learn how to determine a core aspect of the quality of teachers’ assessments that may allow for inferences about these teachers’ assessment competence. Moreover, this is the only line of research that is not only concerned with teachers’
assessments of students’ understanding or achievement but also with the assessment of other characteristics relevant for learning, such as a student’s self-concept or school anxiety (Praetorius et al., 2013; Spinath, 2005). These studies, however, are not concerned with teachers’ judgment processes or assessment practices. Moreover, it remains unclear how judgments of single student characteristics relate to complex judgments in daily instructional contexts and to teachers’ educational decisions (e.g., Südkamp, Praetorius, & Spinath, this issue; van Ophuysen, 2010).

A Conceptual Model of Teachers’ Assessment Competence

Based on the educational concept of competence, our model integrates the previously described lines of research on assessment. We define teachers’ assessment competence as a measurable cognitive disposition that is acquired by dealing with assessment demands in relevant educational situations and that enables teachers to master these demands quantifiably in a range of similar situations in a relatively stable and relatively consistent way. This definition is reflected in our conceptual model, which we present in two parts. The first part is a structural model that explicates how we assume all elements of the model to be related to each other (Figure 1). In the second part, we zoom in on a process model (Figure 2), following Blömeke et al.’s (2015) call to include processes in competence models. This part of the model shows our hypotheses about the steps (i.e., cognitive judgment processes and assessment practices) that connect competence with judgments as the products of assessments.

The Structural Model

Here, we elaborate on the elements that constitute the model, sparing details about the assessment process and its quantification, which will be presented in the second part of the model description.

---- Insert Figure 1 about here ----
**Assessment situations**

Following our definition, an assessment-competent teacher should be able to master a wide range of assessment related situations relevant to the teaching profession. A first crucial step to specifying a model of teachers’ assessment competence is, thus, to determine the types of situations that need to be mastered (e.g., Shavelson, 2010). This step has not yet been tackled systematically and it is to a great extent normative (cf. Keating, 2003; Klieme et al., 2003, Rychen & Salganik, 2003). In line with a situated approach to teachers’ assessment competence, the relevant situations may therefore differ between different educational systems or cultures (cf. Willis et al., 2013). Here, we propose a first generic approach to classifying assessment situations in varying contexts. Based on our definition of educational assessment (see the section on Subject, Scope, and Context for a Model of Teachers’ Assessment Competence), we assume that situations, which can be mastered with high assessment competence, have the following characteristics in common: (1) They start prior to educational decisions and actions and thereby allow assessments to be subordinate to educational decisions. (2) They include information for an educational decision, (3) involve assessment practices that refer to school students, and (4) they address characteristics that are relevant to learning. To measure assessment competence in a valid way, it will be necessary to find a representative set of situations that meet these criteria (Blömeke et al., 2015; Shavelson, 2010). We propose to sample situations that represent the range defined by the following continuous variables (cf. Karst, 2012; Karst, Dotzel, Dickhäuser, this issue): (1) the purpose of assessment, varying between formative and summative (e.g., Bennett, 2010), (2) to what extent the assessment situation can be planned by the teacher (e.g., Burns, 1984; Shavelson, 2006), (3) to what extent the assessment and the informed decisions have far reaching consequences for the student (entailing varying degrees of responsibility and accountability of the teacher, cf. Böhmer et al., in press; Volante & Earl, 2013), and (4) the
complexity of the assessment, that is, whether the teacher needs to consider one or several characteristics of the student in the assessment (Karst et al., 2016, this issue).

It should be noted that situations, which are sampled from the range defined by these criteria, might not only differ with respect to the assessment-related characteristics, but also with respect to particular boundary conditions that can influence the assessment (Esser, 1994, cf. Hempel, 1965). This view is in line with sociocultural perspectives on teachers’ assessments (e.g., Brown & Harris, 2016). Boundary conditions can facilitate or complicate any assessment without changing the type of situation. Referring to Esser (1994), such conditions may be (1) available resources that a teacher can draw on, such as teaching time (i.e., opportunities). These opportunities, however, may be (2) restricted, for example, by competing tasks. Assessment may be influenced by (3) specific institutional rules, that is, standards, conventions, regulations within a class, school, educational system, or subject (cf. Wyatt-Smith & Klenowski, 2013, for different assessment conventions in mathematics vs. English). Finally, the type of assessment situation, the opportunities, restrictions, and institutional rules are indicated (4) by significant symbols, most often by human language. To act appropriately, the teacher needs to correctly perceive and interpret the symbols to identify, for example, whether he or she is in a high- or low-stakes assessment situation (cf. Wyatt-Smith & Klenowski, 2013 for differences in assessment relevant language in mathematics vs. English; see Brown & Harris, 2016 for a thorough discussion of several conditions).

Abstracted from specific boundary conditions, research should sample a set of representative assessment situations from the range defined above, for example, via class observations or interviews (e.g., with stimulated recall or adaptations of the Critical Incident Technique, Flanagan, 1954) in representative samples of teachers.
**Dispositions.** We included two types of dispositions in our model, which we assume to influence teachers’ assessment performance (cf., Xu & Brown, 2016).

**Cognitive disposition.** As suggested by the definition of competence (cf. Koeppen et al. 2008), the cognitive disposition, which is specific to teachers’ actions in assessment situations, is an integral element of the conceptual model. In current conceptualizations of (mathematics) teachers’ professional competences in general (e.g., COACTIV: Baumert & Kunter, 2013; TEDS-M: Kaiser et al., 2017), the majority of teachers’ cognitive dispositions are structured referring to Shulman’s (1986) framework. The dispositions are categorized into content knowledge, that is, school-relevant knowledge about a subject, pedagogical content knowledge, that is, knowledge for teaching the subject, and pedagogical knowledge, the knowledge about teaching and learning independently from a particular subject (e.g., Kaiser et al., 2017). Aspects of teachers’ assessment competence, such as knowledge about assessment of student performance and knowledge about the analysis of student solutions, are subsumed under pedagogical knowledge or pedagogical content knowledge (Baumert & Kunter, 2013; Hoth, et al., 2016; Kaiser et al., 2017).

In contrast with these conceptualizations, we assume aspects of content knowledge, pedagogical content knowledge, and pedagogical knowledge to be *subsumed under teachers’ assessment competence*, which is in line with assumptions in research on assessment literacy (e.g., DeLuca et al., 2016a, b). Therefore, we draw on this line of research to hypothesize about constituents of the disposition. Referring to standard documents and assessment textbooks for teachers (e.g., AFT et al. 1990; Brookhart & Nitko, 2015) models of assessment literacy include general pedagogical knowledge about the concepts of formative and summative assessment (i.e., the assessment purpose). Similarly, these models include knowledge about different assessment methods, such as test development, test scoring, and test interpretation in relation to a reference standard (i.e., criterion-, ipsative, or norm-referenced, Xu & Brown, 2016). A competent teacher would also need to consider typical
judgment biases (e.g., halo effects) to come up with reliable and valid assessments (for more detailed operationalizations of relevant knowledge see DeLuca et al., 2016b; Xu & Brown, 2016). Xu and Brown (2016), moreover, consider teachers’ content knowledge and pedagogical content knowledge as important parts of teachers’ assessment knowledge base, which adds a content-specific aspect to the cognitive disposition. This could entail, for example, teacher’s knowledge about frequent normatively incorrect conceptions of students and how these are usually expressed (e.g., Resnick et al., 1989, for mathematics). Teachers may need this knowledge, for instance, to correctly interpret a students’ utterances in class, when the goal is to assess whether the students still have issues with understanding the current topic (cf. Borko et al., 2008; Lai & Schildkamp, 2016). Teachers may also need knowledge about ethical and legal issues associated with assessments (DeLuca et al., 2016; Xu & Brown, 2016) to prevent discrimination of students and assure correct data use (e.g. for certifications).

In contrast to other models that deal with teachers’ assessments (e.g., Black & Wiliam, 2009; Xu & Brown, 2016), we do not expect teachers’ knowledge about giving feedback or about communicating assessment results, in general, to be a constituent of teachers’ assessment competence because we do not consider teachers’ instructional decisions and actions to be a part of assessment.

To identify which of these aspects of knowledge are indeed needed to successfully cope with assessment related situations, coordinated empirical research is necessary (cf. Herman et al., 2015). In line with research on teachers’ data-based decision making, specific competencies can be assumed to become particularly important in specific steps of assessment (Lai & Schildkamp, 2016; Marsh & Farrell, 2015) and might thus best be inferred when teachers enact these steps.

**Other dispositions.** Although we follow the theoretical approach of Koeppen et al. (2008) in conceptualizing competence as a cognitive disposition, we recognize that other personal dispositions may influence teachers’ assessments in real-life situations (e.g., Izci,
Accordingly, our model assumes that whether and how the cognitive disposition becomes effective in varying situations is influenced by other dispositions of the teacher. The teacher’s other dispositions. Potentially relevant characteristics are teachers’ general information processing ability, motivation, self-concepts, as well as their beliefs and subjective theories (e.g., Izci, 2016; Kaiser, Helm, Retelsdorf, Südkamp, & Möller, 2012; Lai & Schildkamp, 2016; cf. Herppich & Wittwer, this issue). These dispositions might either moderate the relationship between the cognitive disposition and the teacher’s performance or they might influence performance independently from the cognitive disposition. If, for example, a teacher believes that intuitive decisions about teaching are more valid than decisions based on assessment data, this teacher will probably not systematically collect, analyze and interpret data, even if he or she has the necessary competence at their disposal (Lai & Schildkamp, 2016). In this case, the teachers’ belief statistically moderates the effect of competence on performance.

**The process model**

We now move further into our conceptual model (Figure 1) and zoom in on our process model (Figure 2). Afterwards, we consider the quantification of performance.

Assessment. Describing assessment as a process in a model of teachers’ assessment competence helps us to understand how teachers’ knowledge is translated into judgments. The relationships between knowledge, processes and judgments have only rarely been taken into account (see Wyatt-Smith & Klenowski, 2013 for an exception). To disentangle them, we need to theorize about assessment steps and decision points, which may lead to situationally more or less adequate judgments (see Figure 2). Knowing relevant assessment steps allows us
to identify aspects of knowledge that may be required to execute these steps in a competent way.

Our model of assessment is based on dual process models of social judgment formation (Ferreira et al., 2006; Fiske & Neuberg, 1990). In line with the assumptions of these models, research on teachers’ judgment processes showed that teachers form judgments about students along two information processing modes that constitute the endpoints of a continuum (Dünnebier, et al., 2009; Glock & Krolak-Schwerdt, 2014). The first mode is rather simple: It requires little cognitive effort as little information is considered. As this mode involves the activation and processing of heuristic knowledge about the person to be judged, it is assumed efficient and in some cases sufficiently accurate (Gigerenzer & Goldstein, 1996) but also susceptible to subjectivity and bias (Krolak-Schwerdt et al., 2013). The second mode is rather complex: It is associated with cognitive effort as given information about an individual is processed systematically (Ferreira et al. 2006; Fiske & Neuberg, 1990).

To arrive at individual assessment steps, we augmented the cognitive dual process approach with the behavioral dimension of assessment practices (i.e., performance). Thereby, we considered models of psychological assessment (e.g., Fernández-Ballesteros et al., 2001) that are in line with models of teachers’ data-based decision making (e.g., Lai & Schildkamp, 2016). Examples for the endpoints of the continuum illustrate how we understand the different ways of assessment. Numbers in brackets refer to the steps indicated by these numbers in Figure 2. In one situation, a teacher may learn that a new student will join the class (cf. Krolak-Schwerdt et al., 2013). The teacher’s goal [1] could be to obtain a first impression about the student’s achievement and behavior to decide how this student might best be integrated into the class to optimize learning. This goal does not entail a high level of accountability. Such goals trigger simple heuristic ways of assessment (Krolak-Schwerdt et al., 2013). On the shortest assessment route in our model, the teacher now processes available information about this student [2]. For example, the teacher may have heard from colleagues
that the student achieved very good grades. This information may be sufficient to activate the category “number 1 in class”, which, in turn, heuristically leads to inferences about additional attributes (Krolak-Schwerdt et al., 2013; Nickerson, 1999). At this point, the teacher may (unconsciously) decide that no additional information is needed [3] and finally arrive at the judgment that the new student will do well, will be active and easily integrated in class ([9]; cf. Krolak-Schwerdt et al., 2013). Similarly, assessment of low cognitive complexity may be triggered when judgments cannot be planned but need to be accomplished fast and with low processing capacity, as is assumed necessary for many decisions during a lesson (Feldon, 2007).

Teachers will, however, execute a rather complex, iterative and multi-cyclic assessment (Klug et al., 2013; Lai & Schildkamp, 2015; Marsh & Farrell, 2015) when the judgment has far reaching consequences for the student (e.g., Böhmer et al., in press). At the end of primary school, teachers in some countries (e.g., Austria, Germany, Luxembourg, Switzerland, The Netherlands) recommend students to one of several secondary school tracks. These tracking decisions have considerable impact on students’ academic and professional developments (Trautwein et al., 2006). According to our model, the pertaining assessment starts when a teacher becomes responsible for making the tracking decision (e.g., The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK), 2015). The teacher needs to arrive at a reliable and valid judgment about each student’s achievement related characteristics, such as performance in individual subjects, learning trajectories and potential, as well as academic behavior and sometimes social behavior (KMK, 2015). At this point, available information does not suffice to arrive at a judgment [2, 3]. A teacher might start collecting information based on heuristic tendencies towards particular decisions for particular students. During several cycles of information collection, analysis, interpretation, and integration, these hypotheses may become more differentiated [4, 5]. In Germany, national standards stipulate the choice of assessment
methods and their application [7, 8]. Depending on the federal state, teachers may have to assess learning trajectories and performance via methods of formative and summative assessment (portfolios, tests, etc.), observe students continuously, and confer with the students’ parents (KMK, 2015). Research has shown that these kinds of assessment become particularly comprehensive if information about students is contradictory or inconclusive, for example, if students’ results on formal assessments fluctuate (Allal, 2013; Böhmer et al., in press). Furthermore, individual assessments may differ depending on the respective boundary conditions like the social context of a school (Allal, 2013). If the teacher is knowledgeable about the regulations, the implementation of methods, the analysis, interpretation and integration of the data and knows how to draw adequate inferences, as well as if the teachers’ other dispositions influence the assessment in a beneficial way, the resulting judgments may be valid and reliable ([9]; Wyatt-Smith & Klenowski, 2013).

These examples show how the nature of assessment depends on the type of assessment situation (e.g., Böhmer et al., in press). Whether teachers adequately execute the respective mode of assessment, however, crucially depends on whether they have the knowledge at their disposal that is necessary to perform a specific step (cf. Klug, Schultes, & Spiel, this issue). For more complex assessments, research on teachers’ data-based decision making has indicated that teachers need to know about different assessment methods and how to implement them to collect data. Moreover, they need to know about various sources of data (e.g., teacher made tests, standardized tests, formative assessment systems) to be able to apply them and make first interpretations (steps 6-8 in Figure 2; Marsh & Farrell, 2015). Many teachers seem to deal relatively well with the selection, implementation and analysis of assessment methods (Marsh & Farrell, 2015). Thus, the knowledge necessary to execute these steps might often be rather well-developed. Teachers, however, seem to need knowledge about key statistic concepts such as validity and reliability to draw valid inferences from the data (step 2). Moreover, they also need content knowledge and pedagogical content
knowledge to understand the subject specific meaning of a result (step 2; Lai & Schildkamp, 2016). For example, to assess what a student’s mistake means in terms of this student’s learning trajectory, a teacher needs to know what a correct answer would have been and what the mistake tells about the student’s conceptual understanding (e.g., Philipp & Leuders, 2014). Given that teachers seem to have difficulties with these steps of integrating and interpreting information, fostering this knowledge may be particularly important (Furtak, Bakemen, & Buell, this issue; Marsh & Farrell, 2015; Zeuch, Förster & Souvignier, 2017). If the training of these aspects of knowledge would lead to enhanced execution of these steps and better judgments, results would identify them as constituents of the competence (cf. Glogger-Frey, Deutscher, & Renkl, this issue).

Quantification of performance. Having explained the elements that constitute our model, we now turn to the question of how assessment competence may be inferred from teachers’ assessment practices and judgments in real-life situations (cf. Blömeke et al., 2015; Koeppen et al., 2008; Shavelson, 2010). To be able to infer competence in a dependable way, performance has to be quantified using valid criteria (cf. Blömeke et al., 2015; Shavelson, 2010). Next to the specification of assessment situations and model components, the identification of such criteria is, thus, crucial to study and promote assessment competence. For example, only if we know how to infer competence can we evaluate whether attempts to promote it were successful. In empirical studies, different criteria to measure teachers’ assessments have been used that may be suited to infer competence. These criteria apply to different stages of the assessment process.

Based on our process model (Figure 2) a first set of criteria refers to evaluating the quality of assessment steps that lead to a judgment. Several studies on assessment processes point to criteria that may indicate the underlying competence. As reported by Brookhart (2013) for the study of summative assessments, research has compared teachers’ grading practices with recommendations from measurement theory about reliability and validity of
judgments. Grading in this sense refers mainly to the selection, integration, and interpretation of given information to arrive at a judgment (steps 2 and partly 8 in the process model). These studies have, for example, analyzed how different teachers graded the same work of a student to establish reliability (e.g., Starch & Elliott, 1913). Other studies analyzed whether teachers selected only valid, achievement related information when choosing a grade (e.g., Rugg, 1918, cf. Böhmer et al., in press; Oudman, van de Pol, Bakker, & van Gog, this issue). More recently, studies have investigated whether teachers’ summative assessments meet assessment standards (Wyatt-Smith & Klenowski, 2013; D’Agostino and Welsh 2007 and Welsh and D’Agostino 2009 as cited in Brookhart, 2013). These studies, thus, applied a relatively comprehensive criterion to explore the validity of teachers’ assessments in regard to several steps of the assessment process.

Social-psychological research on teachers’ judgment processes, moreover, studied whether teachers used an adequate processing mode, given a specific assessment situation. Böhmer et al. (in press), for example, found in a scenario study that experienced teachers used heuristic judgments when they should advise a colleague on tracking decisions. In contrast, as indicated by reading time and memory data, teachers’ judgments were more detail oriented when they should make tracking decisions themselves. This flexibility was attributed to expert knowledge about student characteristics and official rules for tracking decisions as foundations for adequate judgment processes. As going through the assessment process successfully is a prerequisite for a situation-adequate, sufficiently non-biased judgement, further criteria could be developed to measure how well teachers master other assessment steps, such as the formulation of hypotheses (steps 4-5).

These evaluation criteria for assessments, however, do not quantify the quality of assessment results, that is, of the judgments. In many studies, the correlations between teachers’ judgments of students’ characteristics and students’ outcomes in a standardized test have been used as an indicator of teachers’ judgment accuracy and, thus, assessment quality.
(e.g., Südkamp et al., 2012; Oudman et al., this issue). Teachers with a high ability to accurately judge student characteristics are considered more likely to make meaningful instructional decisions like selecting adequate classroom activities and materials (Vogt & Rogalla, 2009) or interacting adaptively with students (Pielmeier, Huber, & Seidel, this issue). Moreover, if teachers’ judgments replace high-stakes standardized tests, there is a social expectation that these judgments measure students’ achievement well, as inadequate judgments might have severe consequences for the students (cf. Brookhart, 2013; McMahon & Jones, 2015). Therefore, teachers’ judgments of students should be accurate (Furnari, Whittaker, Kinzie, & DeCoster, 2016; Kilday, Kinzie, Mashburn, & Whittaker, 2012; Mashburn, Hamre, Downer, & Pianta, 2006). Nevertheless, there are limitations to this criterion. It would be difficult, for instance, to evaluate teachers’ judgments that are meant to rely on multiple sources (e.g., school-tracking decisions, Böhmer et al., in press) by comparison with a single test result. Similarly, it would be challenging to evaluate on-the-fly assessments during teaching (Shavelson, 2006, cf. Alonzo & Kim, this issue) directly with a measure of judgment accuracy. Judgment accuracy should therefore not be the only criterion to quantify teachers’ assessment competence.

Sometimes, students’ achievement gains are used as a criterion for assessment competence (Bennett, 2011). As it may be used to evaluate how high-quality assessments turn into valuable outcomes, it can be a helpful distal criterion. Achievement, however, is influenced by many other factors, such as the teachers’ instruction and the students’ individual characteristics (van Ophuysen, 2010). To use it as a valid indicator for teachers’ assessment competence these other influences need to be controlled for, entailing, for example, that teachers’ instructional competence also needs to be modeled. Thus, we make a case for not using achievement as the only indicator of assessment competence.

Instead, we first propose that multiple criteria should be employed. These should refer to the assessment process as well as to its product, the judgment. In research, there have been
a few attempts to combine criteria (D’Agostino and Welsh 2007 and Welsh and D’Agostino 2009 as cited in Brookhart, 2013; Glock & Krolak-Schwerdt, 2014; Glock et al., 2016). Second, criteria should be chosen that apply to several types of assessment situations (Shavelson, 2010) as only then situations become comparable and determining a cross-situationally consistent competence is possible. This might be the case for some cognitive process measures (e.g., Krolak-Schwerdt et al., 2013), to some extent for measures of process validity (cf. Brookhart, 2013) and to a lesser degree for measures of judgment accuracy (Südkamp et al., 2012). Third, criteria should refer to realistic teaching situations and thus be relevant for teaching and learning because otherwise the measured competence would not inform educational decisions (Schrader, 1989). Some scenarios used in laboratory research on teachers’ judgments, for instance, might be more externally valid than others (see Pit-ten Cate et al., 2016 for a related discussion). Finally, the criteria should be economical, so that measurements do not become exorbitantly complex. According to these suggestions, we see it as an important task for future research to sample or to create a set of valid criteria to measure teachers’ assessment competence.

Conclusions

Drawing on current approaches to teachers’ professional competences (Blömeke et al., 2015; Kaiser et al., 2017; Koeppen et al., 2008) we developed a conceptual model of teachers’ assessment competence. Our model is comprehensive as it integrates various lines of assessment-related research and thereby extends existing views on teachers’ assessments in several ways. First, a competent teacher should be able to master a wide range of assessment situations. In line with other models of teachers’ competences (e.g., Kaiser et al., 2017; Knievel et al., 2015), we assume that inter-individual differences in performance across situations can largely be explained by differences in the cognitive disposition. Similarly, we assume that intra-individual differences between different types of situations may be explained by differences on competencies that are more or less important for a single type of
situation. For example, if a teacher has acquired profound knowledge of typical student misconceptions, this teacher may be able to assess the type and severity of student mistakes in class (Philipp & Leuders, 2014). With little psychometric knowledge, however, the teacher might perform less well when it comes to the development of a test (Marsh & Farrell, 2015). At the same time, we assume that assessment situations differ on boundary conditions (Esser, 1994) that may lead to varying performances by the same teacher with unchanged dispositions in the same type of situation. Thus, our model also assumes variations in performance due to the variability of context (cf. Willis et al., 2013, Xu & Brown, 2016).

Research is encouraged to examine the potential range of assessment situations and to quantify to what extent performance in similar situations is explained by underlying competence and by varying boundary conditions, respectively. Such research would shed light on the relevance of personal variables in relation to situational variables for the variability of assessments.

Second, our model integrates research on assessment products (e.g., Hoge & Coladarci, 1989; Südkamp et al., 2012) with research on assessment processes (e.g., Black & Wiliam, 2009; Lai & Schildkamp, 2016). The combination of both perspectives suggests that criteria to measure aspects of the assessment processes, such as indicators of process validity (e.g., Wyatt-Smith & Klenowski, 2013) and cognitive measures (e.g., Böhmer et al., in press), as well as criteria to measure judgments, such as judgment accuracy (cf. Brookhart, 2013), should be combined to comprehensively and validly infer assessment competence as both types of criteria have their strengths and shortcomings (see the section on Quantification of performance). Future research based on the model will need to identify an optimal mix of criteria.

At the same time our model is specific because it was developed within Central European educational systems and is concerned with assessments of school students’ characteristics. Thus, its applicability to other educational contexts might be limited.
Nevertheless, we assume our competence based approach to be generally beneficial to describe, explain, predict, and foster educators’ assessments. Moreover, assessment situations certainly have common elements in Central Europe and North America or in schools and higher education, for example, such as formative and summative assessment purposes (Biggs & Tang, 2011) or steps in assessment processes (e.g., Ruiz-Primo & Furtak, 2007). These commonalities make several elements of our model relevant for other educational contexts (e.g., particular competencies and the basic ideas of the process model). Therefore, we hypothesize that the model might be transferable to other educational contexts with some adaptations.

Our model is also specific in that it considers assessments to be conceptually different from teachers’ educational decisions and actions such as giving feedback to students, parents, or administrators. Compared with other models (e.g., Marsh & Farrell, 2015; Xu & Brown, 2016), our model is not assumed to be a base for explaining a broad variety of teachers’ professional actions from the viewpoint of assessment. Instead, we intended to develop a model that inspires research to examine the important task of student assessment on a more fine-grained level (cf. Hoth et al., 2016). The complexity of teachers’ professional actions, overall, may be explained by a combination of such specific competence models (cf. Weinert et al., 1990). This approach should be beneficial for the study and promotion of teaching. Research has, for example, found that teachers struggle with the transition from interpreting and judging data to applying an appropriate instructional response based on the data (Marsh & Farrell, 2015). If we can define and validly measure the competences that are necessary to master the respective task (assessment or instruction) it will be possible to specifically train either the competencies that are necessary to arrive at a valid and reliable judgment (cf. van Es, & Sherin, 2002), or the competencies that are necessary to arrive at an adequate action, given data and judgment (Wittwer, Nückles, Landmann, & Renkl, 2010), or both.
Although modeling teachers’ assessment competence independently comes with major advantages, its separation from other competences of teachers, particularly from instructional competence (cf. Weinert et al., 1990), also bears challenges. We assume assessment and instruction to rely on at least partly different competencies (cf. Döhrmann, Kaiser, & Blömeke, 2012). Undoubtedly, however, they are closely intertwined in teachers’ actions. A particularly complex case may be the involvement of students in their own assessments (Black & Wiliam, 2009; DeLuca et al., 2016b). When teachers’ use data from peer- and self-assessments of students to inform decisions about how to improve learning (Black & Wiliam, 2009) or about students’ certification (Vögeli-Mantovani, 2011), they most probably mainly draw on their own assessment competence. However, when teachers teach assessment skills to students, parts of their own assessment related knowledge become content knowledge to be taught and they need to draw on knowledge about adequate instruction to foster learning, which we do not assume to be part of their assessment competence. To disentangle the competencies needed for the involvement of students in assessment, researchers, for example, could derive hypotheses about knowledge that is possibly relevant based on our model, and train these competencies to examine whether aspects of involvement change (cf. the section on Assessment).

Finally, we developed this conceptual model to show which different strands of research should be included and how they should be related to each other to study teachers’ assessment competence as an effective interaction with the environment. Based on our model, completely worked-out competence models need to be specified (cf. Koeppen et al., 2008). Given that the assessment steps of integrating and interpreting data (step 2 and partly step 8 in the process model) likely involve content knowledge and pedagogical content knowledge (Lai & Schildkamp, 2016; Marsh & Farrell, 2015), specifications may entail adaptations of the model to subjects. Furthermore, the specificity of such models will depend on their purpose. A model that is created to explain differences in the assessments of German elementary
school teachers may need to be less differentiated than a model that is created to study how German elementary school mathematics teachers’ assessment competence develops to efficiently foster this development. For each subject or cluster of similar subjects and each purpose, our model provides a blueprint structure, a template to sample relevant assessment situations, and a model of assessment as a process that can be tailored to given boundary conditions. Research needs to determine how encompassing a model can maximally be (e.g., in terms of school systems, students’ age, cf. Kaiser et al., 2017, or different subjects, cf. Philipp & Leuders, 2014; Wyatt-Smith & Klenowski, 2013) and how specific it needs to be to validly and reliably but also economically describe, explain, predict, and/or promote teachers’ assessment competence.
References

Allal, L. (2013). Teachers’ professional judgement in assessment: a cognitive act and a
20-34. doi:10.1080/0969594X.2012.736364

Alonzo, A. C., & Kim, J. (this issue). *Affordances of video-based professional development
for supporting physics teachers’ judgments about evidence of student thinking.*
Manuscript submitted for publication.

American Federation of Teachers, National Council on Measurement in Education, &
competence in educational assessment of students.* Washington: National Council on
Measurement in Education.

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). Boston, MA:
Springer US.

Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning (CBAL): A
preliminary theory of action for summative and formative assessment. *Measurement:
Interdisciplinary Research and Perspectives, 8*, 70-91.
doi:10.1080/15366367.2010.508686


Herppich, S., & Wittwer, J. (this issue). Preservice teachers' beliefs about students’ mathematical knowledge structure as a foundation for on-the-fly formative assessments Manuscript submitted for publication.


the process of assessing student achievement in the simulated classroom]. Zeitschrift für Pädagogische Psychologie, 26, 251-261. doi: 10.1024/1010-0652/A000076


development of national standards of education. An expert report]. Frankfurt am Main, Germany: Deutsches Institut für Internationale Pädagogische Forschung.


MODEL OF TEACHERS’ ASSESSMENT COMPETENCE


teachers. Theoretical and methodological advancements]. Münster, Germany: Waxmann.


Figure 1. Schematic diagram of the structural model of teachers’ assessment competence.

Adapted from *Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen* [Assessment competence of teachers. Theoretical and methodological advancements] (p. 81), by A. Südkamp and A.-K. Praetorius (Eds.), 2017, Münster, Germany: Waxmann. Copyright 2017 by Waxmann. Adapted with permission.
Figure 2. Schematic diagram of the process model of teachers’ assessments. Numbers indicate the steps as referred to in the article text. Adapted from Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen [Assessment competence of teachers. Theoretical and methodological advancements] (p. 82), by A. Südkamp and A.-K. Praetorius (Eds.), 2017, Münster, Germany: Waxmann. Copyright 2017 by Waxmann. Adapted with permission.
A conceptual model of teachers’ assessment competence is shown. The model includes a cognitive disposition comprised by several competencies. These competencies are assessed in different situations, leading to quantifiable performance. Decisions are made based on the assessment results in each situation, with decisions 1, 2, and x indicating multiple possible outcomes.
1. goal

2. processing of given information about students

3. additional information needed?
   - yes
   - no
     - no: judgment
     - yes: formation of hypotheses needed?

4. formation of hypotheses needed?
   - yes: formation of hypotheses
   - no: collection of information

5. collection of information

6. explicit approach needed?
   - yes: selection of methods, collection of data, analysis of data
   - no: no