Leuders, T. & Loibl, K. (2021). Beyond subject specificity – student and teacher thinking as sources of specificity in teacher diagnostic judgments.

RISTAL 4 / 2021

Research in Subject-matter Teaching and Learning

Volume 4 – Special Issue edited by
Timo Leuders & Katharina Loibl

Beyond subject specificity – student and teacher thinking as sources of specificity in teacher diagnostic judgments

Timo Leuders & Katharina Loibl

Abstract

Teachers' diagnostic judgments are considered highly relevant for good teaching and successful learning. A question regularly asked is: Which aspects of diagnostic judgments are generic and which aspects are specific for a subject or even a topic within that subject? Several sources of specificity have to be considered in diagnostic judgment processes. In order to analyze systematically the aspect of specificity in research on diagnostic judgments, we propose a framework of diagnostic judgments, which explicates relevant components of thinking and behavior on the level of the student and the teacher. We use this framework to discuss systematically the question of content specificity.

Keywords
diagnostic judgments, subject-specificity, topic-specificity, cognitive framework

1 Introduction

Teachers' diagnostic judgments are considered highly relevant for good teaching and successful learning (Furtak et al, 2016; Helmke & Schrader, 1987; Herppich et al., 2018; Leuders, Dörfler, Leuders, & Philipp, 2018). Although it is consensus that a valid diagnosis of student learning (learning processes or achievement as learning outcome) are a prerequisite for pedagogical decisions and effective teaching, there are still many open questions, e.g. how teachers use information and knowledge when generating judgments, or how exactly the quality of teacher judgments unfolds in practice and influences learning processes.

With the focus on the use of information and the application of knowledge, one of the questions regularly asked – either explicitly or implicitly – is: Which aspects of diagnostic judgments are generic and which aspects are specific for a subject or even a topic within that subject (e.g., Lorenz & Artelt, 2009; Kolovou, Naumann, Hochweber, & Praetorius, 2021)? Such questions of specificity can relate to the diagnostic situation (e.g., the student behavior to be diagnosed) or to the teachers' knowledge necessary for accurate diagnostic judgments. Approaches to tackle such questions have been very diverse in methodology and scope. The following examples can be considered to be rather extreme endpoints of a spectrum:

In their meta-study, Südkamp, Kaiser, and Möller (2012) approach the question of specificity by analyzing the impact of certain moderators on judgment accuracy: The subject (e.g., mathematics, language) had no significant statistical influence on judgment accuracy. Also, judgment specificity (from overall rating on a scale to task-specific judgments) and domain specificity (judgments on one specific ability or overall achievement) were not associated with judgment accuracy. However, the congruence of the domain specificity between the achievement test and teachers' judgment was a significant predictor for judgment accuracy.

The general lack of moderating influence cannot be interpreted as a hint on a generic component of judgment accuracy, since this could also be caused by the selection and the design of the studies included. The question of the influence of specific teacher knowledge on judgments is hard to realize in meta-studies: While categories like "subject" or measures like "rank order accuracy" may be applied over many studies, specific teacher knowledge (such as PCK or even more topic-specific knowledge) cannot be expected to be measured equivalently over many studies. This may also be the ultimate reason why Südkamp et al. (2012) did not find suitable data for moderators indicating the relevance of specific teacher knowledge.

The second example is a study on diagnostic judgments with very high specificity with respect to content and knowledge. Leuders and Leuders (2015) analyze verbal judgments of pre-service teachers on two extended solutions of primary students on mathematical tasks. They demonstrate that the diagnostic judgments contain generic aspects (e.g., diligence in writing down) and content-specific aspects (e.g., mathematical strategies). Teachers who solved the tasks in advance on their own (thus, gaining specific knowledge on the task demands) later showed increased specificity in their judgments by referring to specific steps in the solution or to assumed underlying thinking processes. Such findings on the specificity of judgments can only be acquired in designs that assess diagnostic judgments with open formats that allow for capturing the degree of specificity chosen by the judging person and not determined by the research design.

The two examples demonstrate that several sources of specificity have to be considered in diagnostic judgment processes and that the type of research largely determines, which questions on specificity of diagnostic judgments can be answered. In order to analyze systematically the aspect of specificity in research on diagnostic judgments, we propose a framework of diagnostic judgments, which explicates relevant components of thinking and behavior on the level of the **student** and the **teacher** (developed further from Loibl, Leuders & Dörfler, 2020). In the following, we use this framework to discuss systematically the question of **content specificity**.

2 Specificity of diagnostic judgments within a cognitive framework

Diagnostic judgments are social judgments of teachers on students, and consequently are connected to cognition on two levels, the level of students' thinking and behavior and the level of teachers' diagnostic thinking on their students' thinking and diagnostic behavior.

The **student level** (which is the object of diagnostic judgments) can be considered as comprising several components: the student **behavior** is the result of processing information (**thinking**) in the **situation**, e.g., when solving a task presented to the student. This process is influenced by students' **dispositions** (e.g., knowledge, motivation, self-efficacy).
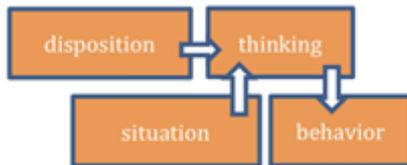
Fig. 1: Structure of situation, disposition, thinking, and behavior on the student level

Since students' thinking processes and their dispositions are not overt, teachers' diagnostic thinking usually processes the manifest behavior of students (as a constituent of the diagnostic situation).

On the **teacher level**, a diagnostic judgment (regarded as **behavior**) is the result of processing information (diagnostic **thinking**) on student behavior in a diagnostic **situation** (see Fig. 2). In addition to this information, the teacher may have knowledge on the students' dispositions and on student thinking and their relation to student behavior (orange dotted lines in Fig. 2). This knowledge is part of the teachers' **dispositions**. Diagnostic judgements can be categorized as **assessments** or **predictions**: For assessments, the teacher infers from manifest student behavior (e.g., test performance) to the latent categories of student thinking or dispositions (e.g., skills or strategies). For predictions, the teacher infers from knowledge about latent features (dispositions, thinking) or observation of manifest student features (behavior) to future manifest student behavior. This also holds for any judgments on tasks. Here, a teacher predicts the solution behavior of a generic student. While the diagnostic situation and the diagnostic behavior represent the externally "visible" side of diagnostic judgments, the the teachers' dispositions and the cognitive processes during diagnostic judgment ("teacher thinking") pertain to the internal "latent" side (see Fig. 2).

Taking the student and the teacher level together in an integrated framework[1] (see Fig. 2), one can observe that the structure is analogous for both levels: **Behavior** is the result of processing information (**thinking**) available in the **situation** on the basis of **dispositions**. The integrated framework provides a theoretical basis for research approaches that consider the student level while seeking to explain teachers' diagnostic judgments as information processing.

---

[1] The teacher level is introduced as cognitive modeling of diagnostic judgements (DiaCoM) in Loibl, Leuders, & Dörfler (2020).
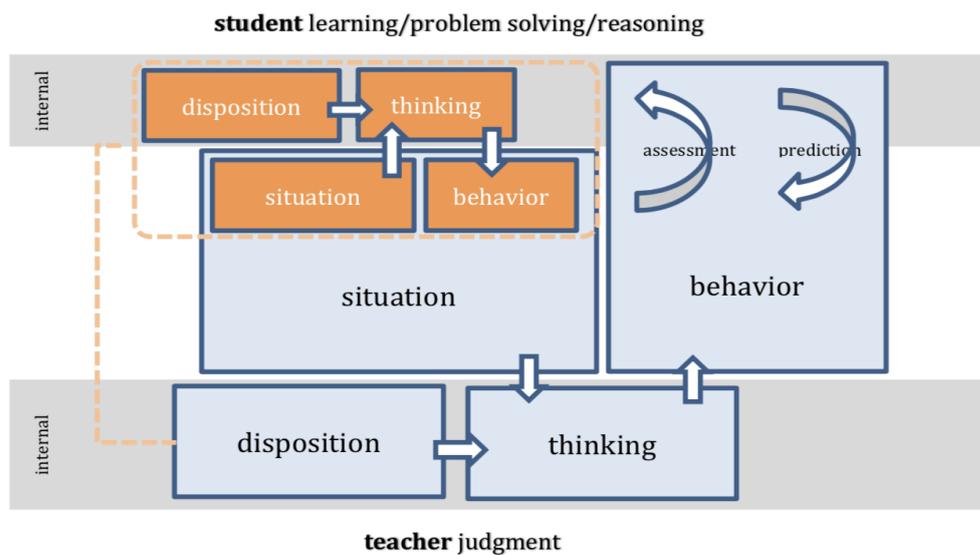
Fig. 2: Integrated framework on teachers' diagnostic judgment

In the following, we apply the integrated framework to analyze the sources of specificity that can be distinguished (and investigated empirically) with respect to diagnostic judgments.

2.1 Specificity of the content of the diagnostic situation

The object of any diagnostic judgment is students' cognition (often termed "student thinking") and its manifestations as students' behavior. More generally, one could say that any diagnostic judgments refer to a 'student model' (orange part of Fig. 2), which can be represented explicitly or implicitly in the teachers' mind and which can be more or less elaborated. The specificity of diagnostic judgments is therefore strongly linked to the specific features of the student model and the components that are objects of the diagnostic judgment: Does the judgment refer to students' dispositions? Does it include assumptions on student thinking? Does it incorporate aspects of the tasks (i.e., the situation)? Which type of student behavior is used to inform the judgment?

The components of the student model used in diagnostic judgments can be of varying specificity. Often, studies distinguish between task-specific, topic- or domain-specific, and subject-specific judgments. With **task-specific judgments**, we refer to predicting or assessing the solution for individual tasks, the self-efficacy or the confidence with respect to individual tasks, etc.. This can be done either task by task (e.g., McElvany et al., 2009) or by comparing specific tasks (e.g., Ostermann, Leuders & Nückles, 2018). **Domain- or topic-specific judgments** predict or assess the student behavior with respect to a group of tasks or a task type targeting a specific topic (e.g., importance of plants in ecosystems, Hoppe, Renkl, & Rieß, 2020) or ability (e.g., reading comprehension, Kolovou et al., 2021). Even broader, **subject-specific judgments** predict or assess the gen-

eral or typical behavior in a broad range of situations pertaining to a subject (e.g., mathematical literacy). All these levels are found subsumed under the term content specificity.

Lorenz and Artelt (2009) as well as Praetorius Karst, Dickhäuser, and Lipowsky (2011) found that teachers' judgement accuracy for different domains or topics within a subject (e.g., reading and writing) was correlated, but this correlation could not be found for judgments in different subjects (e.g., language and mathematics). Kolovou et al. (2021) also provided evidence that within a subject (language or mathematics) teachers' judgment accuracy with respect to different domains or topics (e.g., reading and listening; algebra and geometry) is strongly associated. However, their analyses also revealed that the facets of diagnostic competence regarding the domains or topics within a subject are still psychometrically separable, suggesting a narrower specificity.

This type of specificity is usually explained by domain specific knowledge (e.g., PCK) of teachers, although this dependence is not tested empirically in the cited studies. Other studies do not compare judgments between existing groups (usually teachers with different expertise) but generate differences in knowledge by controlled interventions. These studies deliver evidence for such an explanation. One example is the study by Hoppe et al. (2020) which showed that training diagnostic judgements on students' misconceptions in one biology topic did only enhance judgments in that topic, but did not lead to better judgments on students' misconceptions in another biology topic. So, one can assume that in this case there is no "generic component" of judgment accuracy.

Finally, an even more specific type of judgment can be inspected at the task-level. In their meta-analysis, Hoge and Colardaci (1989) found higher judgment accuracy for item-specific judgments than for global ratings of students. However, this finding was not replicated by Südkamp and colleagues (2012), who found no significant effect of the specificity of the judgment in their meta-analysis. In a study by Karing, Matthäi, and Artelt (2011) the accuracy of global judgments was even higher than the accuracy of task-specific judgments. Karst, Dotzel & Dickhäuser (2018) showed that the accuracy of global or task-specific judgments varied depending on the accuracy measure (rank order vs. level component, cf. Spinath, 2005). Thus, the effects of specificity on judgement accuracy remain inconclusive. In addition, little is known on the different cognitive processes underlying global or task-specific judgments.

A complementary source of specificity refers to the student dimension, more specifically the level of aggregation of student behavior (cf. Karst, Schoreit, & Lipowsky, 2014): Does the diagnostic judgment refer to a specific (single) student, to a group of students (often the teacher's own class) or to "all students", which is equivalent to a generic student. We do not elaborate on the consequences of the student specificity due to the focus of this article on content specificity.

2.2 Types of diagnostic behavior and their varying specificity

The last section discussed the specificity of the diagnostic situation, especially the specificity on the content level. In addition, the diagnostic behavior of the teachers, also var-

ies in type and specificity. We distinguish between predictions and assessment and discuss the specificity with regard to content both in general and with a focus on the methodology.

Predictions are inferences from latent (dispositions, thinking) or manifest student features (behavior, e.g., answers to previous tasks) on their – not yet manifest – behavior in a given diagnostic situation. Predictions can vary with regard to all levels of content specificity: A teacher may predict how a student will solve a specific task or perform a test on a specific topic.

The disposition of the student is topic or domain specific (e.g., specific knowledge, interest) or unspecific (e.g., IQ). Student's thinking unfolds during the solution of specific tasks. A teacher's prediction can refer to such a task or a type of task. However, there is still a lack of evidence on which features teachers actually rely for their predictions. For example, Karing et al. (2011) surmise that in single task judgments, teachers integrate information on a specific task and knowledge about a specific student.

Research usually evaluates teachers' predictions with respect to the accuracy, when compared to objective measures (e.g., performance on a test). In the meta-analysis by Südkamp et al. (2012), judgment accuracy was higher, when the level of specificity of the prediction matched the level of specificity of the objective measure (e.g., teachers judging an ability which is also measured by the test). In most studies, the specificity of the prediction is determined by the measurement: Teachers are asked to predict the solution of a student on one task or on a broader test (with or without knowing the specific test items). Furthermore, task-specific judgments can be evaluated on an aggregated level - for instance, by calculating and comparing average hit rates as measures of accuracy (McElvany et al., 2009, Praetorius et al., 2011; Binder et al., 2019). While this seems to be a methodological question only, it seems reasonable to assume different thinking processes underlying the diagnostic judgment when teachers are asked how many tasks a student will solve (likely relying on students' dispositions) or indicate for each task whether a student will solve it correctly (possibly by reasoning on student's thinking processes).

Assessments are inferences from a manifest student behavior - usually answers to a specific task or set of tasks - to a student's latent disposition. Examples for assessments are categorization judgments (e.g., of a given student solution to a task as an error type) or grading activities (e.g., assignments of an achievement level to a student's performance on several tasks).

Even when only a single task is focused on in a diagnostic judgment, there can be varying grades of specificity with respect to the complexity and richness of a judgment. When only simple judgments are asked for (e.g., a decision on the correctness or a prediction of a solution rate) the specificity of the judgment is predetermined by the method with which it is captured. On the other hand, judgements given in open-answer formats may result in varying extents of specificity depending on the judgment given by a teacher (cf. example by Leuders & Leuders (2015) in the introduction).

2.3 Specificity of teacher dispositions

Among the many teacher variables that potentially influence diagnostic judgments, teacher knowledge is regarded as one of the most important. Within their framework for pedagogical content knowledge (PCK), Ball, Thames, and Phelps (2008) characterize the knowledge components relevant for diagnostic judgments as knowledge about content and students (KCS). By definition, KCS is domain- or topic-specific knowledge (as are all components of PCK). KCS can be used, for instance, to identify possible causes for student errors (assessment) or to choose adequate tasks for testing an assumption on students' thinking. Such diagnostic activities refer to the student model (cf. orange part of Fig. 2) and therefore require specific knowledge regarding student cognition and behavior in the domains or topics to be diagnosed. Furthermore, von Aufschnaiter et al. (2015) distinguish between teachers' knowledge on students' knowledge, motivation, and learning and teachers' knowledge on diagnostic methods (e.g., quality criteria of diagnostic procedures). According to the authors, both facets contain subject-specific and subject-independent aspects.

For research on diagnostic judgments, a more differentiated view on the types of knowledge and their grades of specificity is needed. One of the most prominent questions is, which type of knowledge do teachers need to achieve accurate judgments? Or with regard to the cognitive processes involved (see 2.4), how do teachers process specific knowledge and information on tasks and/or student behavior to generate diagnostic judgments (regardless of the accuracy)? Research that investigates such questions regularly draws on designs that focus on specific domains or topics and systematically relate teacher knowledge and situational information.

While PCK shows only unsystematic and weak correlations with judgement accuracy (e.g., Binder et al., 2018; Rausch, Matthäi, & Artelt, 2015), specific knowledge about features that render tasks more or less difficult seems more important. This specific knowledge can have different layers with regard to the student model (orange part of Fig. 2):

When judging the difficulty of tasks, teachers apply knowledge about the demands of the tasks (i.e., the **situation** encountered by students). For instance, knowledge on the demands of text-picture integration when reading discontinuous texts correlated with judgment accuracy (McElvany et al., 2009); instruction on features that render fraction comparison tasks more or less difficult increased the judgment accuracy (Rieu, Loibl, Leuders, & Herppich, 2020).

For judgments on individual students, teachers need knowledge about the demands of the tasks and the **dispositions** of the students. For instance, Karing et al. (2011) distinguished between global judgments on students' reading ability and task-specific predictions of the correctness of student answers to text comprehension questions. The accuracy of both judgments differed in favor of the global judgment. The authors ascribed this finding to the cognitive demanding process of integrating information of students and task features underlying task-specific predictions. However, in their study they cannot reveal how this integration process is supposed to work.

Furthermore, teachers' diagnostic judgments can build upon their knowledge about **student thinking** - either by reconstructing the solution process (Morris, Hiebert, & Spitzer, 2009) or by applying knowledge on typical errors and misconceptions (Ostermann et al., 2018). In an experimental study, Ostermann et al. (2018) showed that content-specific knowledge on typical errors and misconceptions increased the judgment accuracy with respect to average solution rate and to rank order, while an unspecific sensitization for the general tendency to overestimate students' competences only improves the former. This finding supports the assumption that specific knowledge on task demands is actually used to judge the relative difficulty of tasks (Loibl et al., 2020).

2.4 Specificity of cognitive processes during diagnostic judgment

The last component of the cognitive framework in Fig. 2 to be discussed is diagnostic thinking: The generation of diagnostic judgments (diagnostic thinking) can be considered as processing (external) information from the diagnostic situation and (internal) knowledge and can be described and investigated via cognitive modeling (for a framework see Loibl et al. 2020). The pertaining cognitive processes can - on a general level - be considered as generic processes of human perception and reasoning: Typically judgment processes comprise steps of perceiving elements of the diagnostic situations (cues), of interpreting them by applying knowledge, and by more complex processes of weighing, integrating or decision making (cf. e.g., Greifeneder, Bless, & Fiedler, 2017). The specificity of diagnostic thinking enters through the specificities of the pieces of information that are processed: the specificity of the diagnostic situation and the types of knowledge.

The connection between diagnostic thinking and specificities of the situation can be of generic nature, e.g., when time restrictions or stress influence judgment processes (e.g., Becker, Spinath, Ditzen, & Dörfler, 2020; Rieu et al., 2020), or when beliefs or stereotypes generate judgment biases (e.g., Pit-ten-Cate, Krolak-Schwerdt, & Glocket, 2016). But also, specificities of the content (subject, domain, topic) can be highly relevant for the cognitive processes at work. For example, judging language proficiency relies on the perception and interpretation of volatile auditive information and therefore requires processes of storing and retrieving auditory memory.

3 Constituent papers

In how far the various discussed sources of specificity contribute to the formation of diagnostic judgments, must and can be investigated by empirical research. With view on the content specificity, subject specialist researchers can substantially contribute to this research area. Moreover, as generic and specific perspectives on teacher thinking can benefit from each other, interdisciplinary research appears indispensable.

This Special Issue includes four papers that illustrate such interdisciplinary approaches: Witzigmann et al. (2021) and Loibl and Leuders (2021) study teachers' assessments, while Brunner et al. (2021) and Schreiter et al. (2021) study predictions.

Witzigmann et al. (2021) focused on assessments of oral language samples in French as a foreign language. More precisely, they investigated the use of different cues (linguistic

features) as a basis for inferences from task-specific behavior on domain-specific dispositions (ability of oral language production) of individual students.

In a simulation study, Loibl and Leuders (2021) focused on inferences from task-specific behavior (responses to specific tasks) on topic-specific dispositions (misconceptions) of individual simulated students. They modelled teachers' diagnostic judgements by Bayesian reasoning and approximative heuristics.

Brunner et al. (2021) explored teachers' judgements of mathematical task difficulty with eye tracking. For their predictions, pre-service teachers had to draw on their topic-specific knowledge on typical student dispositions to infer typical behavior based on task-specific difficulties.

Similarly, Schreiter et al. (2021) also investigated judgments regarding the difficulty of mathematical tasks. That is, pre- and in-service teachers predicted task-specific behavior of individual generic students. Schreiter et al. distinguished between topic-specific difficulties (fractions) and unspecific difficulties (design of the task with regard to the extraneous cognitive load, such as split-attention).

References

Ball, D., Thames, M. H., & Phelps, G. (2008). Content Knowledge for Teaching. *Journal of Teacher Education, 59*(5), 389–407.

Becker, S., Spinath, B., Ditzen, B., & Dörfler, T. (2020). Der Einfluss von Stress auf Prozesse beim diagnostischen Urteilen – eine Eye Tracking-Studie mit mathematischen Textaufgaben. *Unterrichtswissenschaft*, *48*(4), 531-550.

Binder K., Krauss S., Hilbert S., Brunner M., Anders Y., Kunter M. (2018) Diagnostic Skills of Mathematics Teachers in the COACTIV Study. In: Leuders T., Philipp K., Leuders J. (eds) *Diagnostic Competence of Mathematics Teachers*. Springer, Cham.

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de Leon, V., Morrison, D., & Heredia, Sara C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*, *44*, 267e291

Greifeneder, R., Bless, H., & Fiedler, K. (2017). *Social cognition: How individuals construct social reality*. Psychology Press.

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

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

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgements of academic achievement: A review of literature. *Review of Educational Research*, *59*(3), 297-313.

Hoppe, T., Renkl, A., & Rieß, W. (2020). Förderung von unterrichtsbegleitendem Diagnostizieren von Schülervorstellungen durch Video und Textvignetten. *Unterrichtswissenschaft, 48*, 573-597.

Karing, C., Matthäi, J., & Artelt, C. (2011). Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I–Eine Frage der Spezifität? *Zeitschrift für Pädagogische Psychologie*, *25*(3), 159-172.

Karst, K., Dotzel, S., & Dickhäuser, O. (2018). Comparing global judgments and specific judgments of teachers about students' knowledge: Is the whole the sum of its parts?. *Teaching and teacher education, 76*, 194-203.

Karst, K., Schoreit, E., & Lipowsky, F. (2014). Diagnostische Kompetenzen von Mathematiklehrern und ihr Vorhersagewert für die Lernentwicklung von Grundschulkindern. *Zeitschrift für Pädagogische Psychologie, 28,* 237-248.

Kolovou, D., Naumann, A., Hochweber, J., & Praetorius, A. K. (2021). Content-specificity of teachers' judgment accuracy regarding students' academic achievement. *Teaching and Teacher Education*, *100*, 103298.

Leuders, J., & Leuders, T. (2015). Assessing and supporting diagnostic skills in pre-service mathematics teacher education. Paper presented at the *Joint Meeting of PME 38 and PME-NA 36*, Vancouver, Canada.

Leuders, T., Dörfler, T., Leuders, J., & Philipp, K. (2017). Diagnostic Competence of Mathematics Teachers: Unpacking a Complex Construct. In T. Leuders, T. Dörfler, J. Leuders, & K. Philipp (Eds.), *Diagnostic Competence of Mathematics Teachers. Unpacking a Complex Construct in Teacher Education and Teacher Practice* (pp. 3-32). New York: Springer.

Loibl, K., Leuders, T., & Dörfler, T. (2020). A Framework for Explaining Teachers' Diagnostic Judgements by Cognitive Modeling (DiaCoM). *Teaching and Teacher Education, 91*, 103059

Lorenz, C., & Artelt, C. (2009). Fachspezifität und Stabilität diagnostischer Kompetenz von Grundschullehrkräften in den Fächern Deutsch und Mathematik. *Zeitschrift für Pädagogische Psychologie*, *23*(34), 211-222.

McElvany, N., Schroeder, S., Hachfeld, A., Baumert, J., Richter, T., Schnotz, W., Hortz, H., & Ullrich, M. (2009). Diagnostische Fähigkeiten von Lehrkräften: bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für pädagogische Psychologie, 23*(34), 223-235.

Morris, A. K., Hiebert, J., & Spitzer, S. M. (2009). Mathematical knowledge for teaching in planning and evaluating instruction: What can preservice teachers learn?. *Journal for research in mathematics education*, *40*(5), 491-529.

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

Pit-ten Cate, I. M., Krolak-Schwerdt, S., & Glock, S. (2016). Accuracy of teachers' tracking decisions: Short- and long-term effects of accountability. *European Journal of Psychology of Education, 31*(2), 225-243.

Praetorius, A. K., Karst, K., Dickhäuser, O., & Lipowsky, F. (2011). Wie gut schätzen Lehrer die Fähigkeitsselbstkonzepte ihrer Schüler ein? Zur diagnostischen Kompetenz von Lehrkräften. *Psychologie in Erziehung und Unterricht*, *58*(2), 81-91.

Rausch, T., Matthäi, J., & Artelt, C. (2015). Mit Wissen zu akkurateren Urteilen? Zum Zusammenhang von Wissensgrundlagen und Urteilsgüte im Bereich des Textverstehens. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 47(3), 147-158.

Rieu, A., Loibl, K., Leuders, T., & Herppich, S. (2020). Diagnostisches Urteilen als informationsverarbeitender Prozess–Wie nutzen Lehrkräfte ihr Wissen bei der Identifizierung und Gewichtung von Anforderungen in Aufgaben? *Unterrichtswissenschaft*, *48*(4), 503-529.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 19, 85-95.

Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgements of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, *104*(3), 743-762.

von Aufschnaiter, C., Cappell, J., Dübbelde, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßer, R. & Wolgast, A. (2015). Diagnostische Kompetenz: Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik, 61*(5), 738-757.

Katharina Loibl

is a professor for interdisciplinary research on learning and instruction at the University of Education Freiburg, Germany with a research focus on learning mechanisms and instructional designs. She is co-speaker of the research training group "DiaKom" within which this research was conducted.

Timo Leuders

is a professor for mathematics education at the University of Education in Freiburg, Germany with a research focus on teaching and learning in secondary education and teacher professionality. He is co-speaker of the research training group "DiaKom" within which this research was conducted.