Your full name: Ann-Kristin Adleff

Affiliated authors with institutions: Natalie Ross (University of Hamburg)

Affiliation: University of Hamburg

Current position: PhD student

Title of your paper: Understanding instructional quality through the measurement of task characteristics

Abstract (300 words)

The quality of mathematics instruction is a complex, multi-dimensional construct. It is the interplay of different actors and factors inside and outside of the classroom, which determines the nature of a mathematics lesson. Mathematical tasks are often at the centre of both students' and teachers' attention before (lesson preparation), during (main carriers of student activity) and after the lesson (homework and other assignments) and are used to assess students' learning gains. In addition, they facilitate the students' access to the mathematical content. The present study based on the TEDS-Validate research project aims to better understand the relations between the instructional quality (assessed via classroom observations), and the quality of the tasks that the students work on during lessons.

Trained raters assessed the instructional quality in 28 classrooms in different regions of Germany using an observation protocol developed for the TEDS-Instruct study. The observation protocol comprises items concerning the surface and the deep structure of both generic and subject-specific dimensions of instructional quality. The potential of the approximately 2600 tasks taken from the observed classrooms was rated based on a newly developed classification scheme for mathematical tasks. First results suggest the validity and reliability of the ratings. In accordance with prior studies, we expect to find significant correlations between task quality and the cognitive activation in the classroom as well as subject-specific aspects of instructional quality.

Extended summary (1000 words, excluding reference list) introduction, theoretical background, methods, preliminary findings/findings, results, reference list.

Instructional quality – or teaching quality – is a question of great interest in current research in mathematics education. There is, however, no clear consensus when it comes to defining high quality teaching. According to Berliner (2005), quality teaching is both good and successful. In this context, the former refers to a normative and rather theoretical outlook on teaching. In contrast, the latter examines the outcomes of teaching from an empirical point of view. Effective (or successful) teaching is characterized by certain outcomes on the side of the students such as high learning gains, intrinsic learning motivation and a positive self-concept in mathematics (Scherer et al. 2016). Thus, "[based] on the educational effectiveness paradigm, teaching quality can be defined as instructional aspects influencing students' cognitive and affective learning outcomes" (Fischer et al. 2019). Following this perspective, in order to better understand instructional quality, one needs to identify and investigate those aspects on mathematics teaching that have an impact on students' learning.

> From a cognitive perspective, the detail and content of tasks have a significant effect on learning; from a cultural perspective, tasks shape the learners' experience of the subject and their understanding of the nature of mathematical activity; from a practical perspective, tasks are the bedrock of classroom life, the 'things to do' (Watson & Ohtani 2015, S. 3).

Evidence from the COACTIV project (*Professional Competence of Teachers, Cognitively Activating Instruction, and Development of Students' Mathematical Literacy)* suggests that task quality is an indicator for the level of cognitive activation and promotes students' learning as part of instructional quality (Baumert et al. 2010). The aim of the present study is to explore the relations of different aspects of instructional quality and the quality of the tasks implemented throughout the corresponding lessons.

The studies TEDS-Instruct and TEDS-Validate (*Teacher Education and Development Study*) focus on relations between mathematics teachers' professional competences and their students' learning gains (Kaiser et al. 2017). In this context, the quality of both the teaching and the tasks from the classrooms are estimated and interpreted as mediating variables between the teacher and the students. The instructional quality was assessed based on the classroom observation protocol developed in the context of the study TEDS-Instruct (Schlesinger et al. 2018). It comprises various items from three generic dimensions – that is classroom management, cognitive activation and student support (Klieme & Rakoczy 2008) – as well as subject-specific items of instructional quality. In addition to these measurements, all mathematical tasks set by the teachers throughout the observed lessons were collected resulting in a total sample of approximately 2600 tasks.

Regarding the scope of task analysis, Resnick (1975) distinguishes between rational and empirical task analysis. The former is based on the task itself as well as potential and idealized performances imaginable as successful approaches in solving the task. Empirical task analysis, on the other hand, integrates actual solution processes using transcripts of students' notes or self-reports. While there have been attempts to combine aspects of both approaches in order to contrast the (intended) task potential with the actual implementation in the classroom (e.g. Blömeke et al. 2006), the present study focuses on rational task analysis – hence, task potential only – in order to minimize intersections with classroom features that also impact the measurement of instructional quality.

A comprehensive analysis of mathematical task quality – or task potential – requires the consideration of various aspects. In order to allow for construct validity and reliability, a classification scheme consisting of different dimensions of mathematical task quality was developed based on prior work from COACTIV (Jordan et al. 2006) and PISA (Turner et al. 2015). The classification scheme includes categories to examine the tasks with regard to surface features, the underlying mathematical concepts and ideas, the procedural characteristics of doing mathematics as well as the cognitive and linguistic complexity. In addition, special task characteristics such as the openness of a task were included in the analysis. While surface features do not necessarily allow for consistent scales, all other characteristics were assessed using a four-point ordinal scale to enable computational comparability. The interrater reliability proved to be good or very good for the majority of the categories, but 20% of the tasks in the sample were still coded by two raters in order to control for rater effects.

First results indicate that the tasks from the observed lessons in different federal states of Germany offer little potential for cognitive activation and thus confirm the findings from the COACTIV project (Neubrand et al. 2013). To give an example, approximately 80 % of tasks in the sample required no or only little problem-solving activities. One reason for this may lie in their lack of openness. Considering the openness of the solution method, half of the tasks from the sample do not allow for more than one method by, for instance, stating the anticipated method or by requiring the students to recall declarative knowledge without applying it. Both categories show a medium level correlation (.386**), which makes sense as task openness is often considered to be a prerequisite for mathematical problem solving (Cai 2000).

Following the current matching of tasks and corresponding measurements for instructional quality, the relations between the task characteristics and the different dimensions of instructional quality will be analyzed. First results may be presented at the QUINT conference 2021. In accordance with prior studies, we expect to find significant correlations between task quality and the cognitive activation in the classroom as well as subject-specific aspects of instructional quality. As rational task analysis is limited to the task potential as opposed to the actual implementation in the classroom, we do not expect significant correlations between task potential and classroom management or student support.

References

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133 – 180.

Berliner, D. C. (2005). The Near Impossibility of Testing for Teacher Quality. *Journal of Teacher Education*, 56 (3), 205–213.

Blömeke, S., Risse, J., Müller, C., Eichler, D., & Schulz, W. (2006). Analyse der Qualität von Aufgaben aus didaktischer und fachlicher Sicht. Ein allgemeines Modell und seine exemplarische Umsetzung im Unterrichtsfach Mathematik. *Unterrichtswissenschaft*, 34(4), 330-357.

Cai, J. (2000). Mathematical thinking involved in US and Chinese students' solving of process-constrained and process-open problems. *Mathematical Thinking and Learning*, *2*(4), 309 – 340.

Fischer, J., Praetorius, A.-K., & Klieme, E. (2019). The impact of linguistic similarity on cross-cultural comparability of students' perceptions of teaching quality. *Educational Assessment, Evaluation and Accountability*, 31(2), 201–220.

Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M. et al. (2006). *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenklassifikation im COACTIV-Projekt*. Berlin: Max-Planck-Institut für Bildungsforschung.

Kaiser, G., Blömeke, S., Koenig, J., Busse, A., Doehrmann, M., & Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers—cognitive versus situated approaches. *Educational Studies in Mathematics*, 94(2), 161 – 182.

Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. Zeitschrift für Pädagogik, 54(2), 222–237.

Neubrand, M., Jordan, A., Krauss, S., Blum, W. & Löwen, K. (2013). Task Analysis in COACTIV: Examining the Potential for Cognitive Activation in German Mathematics Classrooms. In M. Kunter et al. (Eds.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers* (pp. 125 – 144). New York: Springer.

Scherer, R., Nilsen, T., & Jansen, M. (2016). Evaluating Individual Students' Perceptions of Instructional Quality: An Investigation of their Factor Structure, Measurement Invariance, and Relations to Educational Outcomes. *Frontiers in Psychology* 7, 110.

Schlesinger, L., Jentsch, A., Kaiser, G., König, J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM*, 50(3), 475 – 490.

Turner, R., Blum, W. & Niss, M. (2015). Using competencies to explain mathematical item demand: A work in progress. In K. Stacey & R. Turner (Hrsg.), Assessing mathematical literacy. The PISA experience (S. 85–115). Springer: Cham.

Watson, A. & Ohtani, M. (2015). Themes and Issues in Mathematics Education Concerning Task Design: Editorial Introduction. In A. Watson & M. Ohtani (Eds.), Task Design In Mathematics Education. An ICMI study 22 (pp. 3-14). Springer: Cham.