THE POTENTIAL OF TASKS FOR MATHEMATICAL LEARNING AND ITS USE IN INSTRUCTION –PERSPECTIVES OF EXPERTS FROM GERMANY AND TAIWAN

Anke Lindmeier[1], Ting-Ying Wang[2], Feng-Jui Hsieh[2], Anika Dreher[3]

[1]FSU Jena, Germany, [2]NTNU Taipei, Taiwan, [3]PH Freiburg, Germany

The potential of tasks to stimulate students' mathematical thinking and the adequate use of this potential in instruction are prominent indicators for instructional quality. Since the assessment of a task's potential depends on the aims of instruction, it may be argued that corresponding perspectives vary between cultural contexts. However, so far, this has not been systematically investigated in cross-cultural comparisons. In this study, we investigate whether Western (German; N=17) and East Asian (Taiwanese; N=19) professors of mathematics education have different perspectives on the potential of word problems for students' learning and the use of this potential in instruction by means of vignettes from a cross-cultural research project. We illustrate how differences reflect cultural aspects of mathematics instruction.

THEORETICAL BACKGROUND

The potential of tasks for students' mathematical learning and the use of this potential in teaching (the *potential of tasks and its use*) have been shown to be crucial factors for students' learning. Across cultures, there is a consensus that competent teachers are able to identify tasks with high learning potential, and, in addition, implement them in a way that uses this potential (e.g., Stein & Lane, 1996). However, it is well known that Western and East Asian perspectives on mathematics teaching and learning are different in many aspects (Leung, 2001). Hence, it is questionable whether research focusing on the evaluation of a task's potential and its use can be cross-culturally valid (Clarke, 2013) and it is thus important in our inter-cultural research community to seek corresponding evidence. Consequently, this research report investigates how professors of mathematics education (experts) from Taiwan and Germany (representing an East Asian and a Western perspective) evaluate the potential of tasks and its use in instructional situations. We focus on a very common kind of task that is used in mathematics instruction across grades and cultures: word problems with links to real-life situations.

Word problems, their learning potential and use in Germany and Taiwan

Generally, mathematical tasks are considered to have a high potential for students' learning, if they are focused on the instructional content, aligned with the teaching aims, and suited to stimulate students to work mathematically. Word problems, in particular, often have features that are considered to promote learning, such as their potential to provoke multiple solutions or require explanations (Stein & Lane, 1996).

2022. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.). *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 139-146). PME.

However, word problems are used with many different functions (e.g., Verschaffel et al., 2020). For example, word problems can be used to practice mathematical procedures, to discover new mathematical concepts, or to engage in mathematical modeling. Hence, it is an interesting question whether a certain word problem's potential for supporting students' learning may be evaluated differently. As it is known that the use of word problems varies between cultural contexts (e.g., Chang et al., 2020), this question is especially relevant for cross-cultural comparative research.

Mathematics teaching in Germany and Taiwan has typically different priorities such as meaningful learning vs. high procedural fluency (Leung, 2001), which may impact the perspective on word problems and their potential for learning. The German curriculum is literacy-oriented and clearly stands in a Western tradition. Engaging in mathematical modeling processes is hence an important practice (Chang et al., 2020). There is a focus on using real-life situations to encourage students to draw on their world-knowledge to understand them and validate solutions against the situation (Verschaffel et al., 2020). In Taiwan, word problems are used with a strong focus on the application of foundational knowledge and procedures (Chang et al., 2020; Pratt et al., 1999). Consequently, Taiwanese students were consistently found to outperform Western students in comparative studies where word problems were used for assessment, benefiting from a sound knowledge base and flexible use of procedures, that may result from high perseverance in studying (Leung, 2001).

Based on these differences, it can be assumed that there may be different perspectives in Germany and Taiwan on what constitutes a high potential of word problems for students' mathematical learning, and, consequently, how this potential should be used in mathematics instruction. Particularly, there are indications that word problems with real-life contexts are used with different aims in Germany and Taiwan: While in Taiwan such problems are primarily seen as opportunities to apply mathematical concepts and procedures to deepen mathematical understanding, in Germany they are seen as opportunities to learn mathematical modeling as a specific practice.

Eliciting culture-specific norms using of vignettes

To elicit and contrast perspectives on teaching quality across cultures, we follow approaches that use classroom vignettes to assess professional noticing (Dreher et al., 2021). Professional noticing with respect to teaching is described as a process of attending to aspects of classroom situations that are relevant for instructional quality (selective attention) and interpreting them by drawing on corresponding professional knowledge and other resources (knowledge-based reasoning) (Sherin, 2007). Typically, instruments to assess noticing use text- or video-based vignettes as representations of practice. A common "operational trick" in these approaches is to design or select vignettes in which something happens that does not meet the expectations of "good" teaching, i.e., they include a *breach of a norm* regarding some aspect of instructional quality (Dreher et al., 2021). The vignettes are shown together with a prompt to evaluate the depicted classroom situation and to give reasons for the

evaluation. A person's reaction to the critical incident serves then as the indicator for the noticing; the reasoning can be used to infer what knowledge and beliefs guided the noticing process.

Up to now, such vignettes have mainly been used to assess noticing. One could, however, also use them to investigate whether the noticing of experts from different cultures reflects differing norms regarding aspects of instructional quality. To do so, one would need vignettes that potentially show breaks of culture-specific norms. However, in comparative studies, such culturally sensitive instruments are usually avoided as much as possible in order not to jeopardize the comparability of the results. This does not solve the problem that seeking the highest possible comparability may be detrimental to the validity of the instruments precisely when conceptions of instructional quality differ across cultures (Clarke, 2013). To the best of our knowledge, this has not been systematically investigated for the instructional quality regarding task potential and its use, as the corresponding instruments were lacking.

RESEARCH QUESTIONS

Against this background, we ask: Do mathematics education experts from Taiwan and Germany have different perspectives on the potential of word problems and its use as represented in vignettes authored in Germany or Taiwan?

CONTEXT AND METHODS OF THE STUDY

The reported study is part of the binational research project "Teacher noticing in Taiwan and Germany" (TaiGer Noticing) aiming at investigating the role of culture-specific norms regarding aspects of instructional quality. To this end, we developed a set of text vignettes reflecting potentially culture-specific norms regarding aspects of instructional quality (Dreher et al., 2021). Due to the prominent role of tasks in mathematics teaching, one of these aspects is the potential of tasks and its use. To validate whether the developed vignettes reflect indeed norms regarding this aspect in the respective countries, all vignettes were evaluated by experts in Germany and Taiwan. This report uses the responses regarding two of the vignettes (task2, task4). Vignette task2 was developed in Germany and vignette task4 in Taiwan. Both included a breach of a norm from the perspective of the authoring national team members. Due to the sophisticated method of a concurrent vignette development process in the research project (Dreher et al., 2021), we could ensure that the resulting vignettes represent classroom situations that may occur in secondary mathematics education of both countries (ecological validity).

Instruments

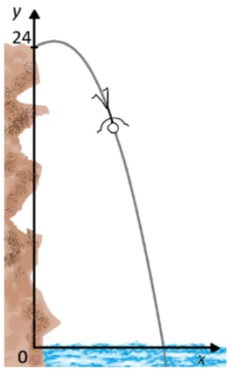
The two vignettes have a similar structure: First, a task that is considered to have a high potential for mathematical learning from the perspective of the authoring national team is presented. Second, a classroom situation is described (approx. 230 words of a fictitious transcript).

In detail, vignette tasks2 builds on a task "cliff-jumping" (topic of quadratic functions, Figure 1, left). It requires students to understand a real-life situation (presented as graphically supported text), make an educated guess about the solution based on the real-life context, and determine the solution with the help of a given mathematical model. The German authors saw the potential of the task for learning especially in its clear focus on the known difficulties of students to understand and interpret the connection between the real-life situation and mathematical models. They would expect teachers to use the educated guesses or the given visual representation to validate mathematical solutions and support students' modeling processes.

Vignette task4 builds on a task "student camp" (topic of systems of linear equations, Figure 1, right). It requires students to understand a real-life situation (presented as text) and set up a system of linear equations to determine the solution. The Taiwanese authors saw a specific potential for students' learning of this tasks, as it is suited to discuss pros and cons of different possibilities to assign variables: Assigning x and y to be the numbers of groups of students leads, for example, to a simpler calculation than assigning x and y the numbers of students in congruence to the unknowns in the word problem. The Taiwan team members would hence expect the teacher to discuss how different ways of variable assignment lead to systems of equations with different characteristics so that students acquire abilities to use different strategies flexibly for effective solutions.

Cliff divers jump from high rocks into the sea. In order to determine whether it is safe to jump, it is important to know where they will hit the water (entry point).
Imagine that the rock is approximately vertical and 24 m high.

1. How far is the entry point from the cliff? Please make an educated guess.
2. An optimal jump trajectory can be described by the function $f(x) = -(x-1)^2 + 25$.
   How far is the entry point from the cliff in this case?



In a student camp with a total of 60 secondary and primary students, students are grouped with others of the same school type. If every group consists either of 9 secondary students or of 4 primary students, then there is a total of 10 groups.
How many secondary school students and how many primary students are in the student camp?

Figure 1: Task "cliff-jumping" (vignette task2, Germany); Task "student camp" (vignette task4, Taiwan).

The classroom situations represented by the vignettes task2 and task4 were designed to depict non-optimal use of the potentials of the tasks from the perspective of the authoring team (breach of a norm). In the vignette task2, the teacher works in an interactive manner with the students but makes no advantage of the task's potential to focus on mathematical modeling processes. In the vignette task4, the teacher presents two different ways of assigning variables (x, y groups of students; x, y numbers of students) and labels the first one as resulting in a simpler calculation, but does not use the potential of the task to discuss the pros and cons of the different ways of assigning

variables. During the development of the vignettes, the team members from Germany as well as the members from Taiwan already experienced that seeing the specific tasks' learning potentials and, subsequently, their non-optimal use can be difficult for the members of the other culture.

Sample and procedures

Participants were recruited from professors of mathematics education who were active in mathematics education research and in preparing future secondary mathematics teachers. As we aimed for a sample of 15 experts in each country and assumed a participation rate of at least 50%, in Germany, a random sample of 30 professors out of the full list of persons meeting these criteria was contacted. In Taiwan, these criteria yielded a list of only 32 professors and thus all of them were contacted. In total, a sample of $n_1 = 19$ Taiwanese professors (6 female, 13 male) from 10 universities and a sample of $n_2 = 17$ German professors (7 female, 10 male) from 13 universities worked on the vignettes (completion rates were TW 59%, GER 56%). To capture the experts' perspectives on the tasks' potentials and their use, the experts were given the following open-ended prompt: "Please evaluate the teacher's use of the task in this situation and give reasons for your answer."

Both vignettes were administered to experts in both countries online in their native language (German resp. Chinese). Responses were translated into English as the common language within the research team and analyzed with respect to two main aspects: 1) Did the experts evaluate the teachers' use of task as inadequate? And if so: 2) What were their reasons? We coded whether the experts saw a breach of the same norm as the authors. In addition, we expected that experts may see further reasons why the task implementation can be criticized, so we extracted further reasons inductively from the answers. More than one reason could be assigned to an answer.

RESULTS

In this research report, we summarize the coding as follows (Table 1): First, we give the number of expert responses showing no negative evaluation of the classroom situation depicted in the vignette (no breach). We count the number of responses where experts saw the intended breach of a norm. In the remaining responses, the experts only gave other reasons for their negative evaluation. To answer our research question, we focus here on the perspectives of the majority of experts in each culture on the given vignettes. With this approach, we highlight what can be considered a norm within each culture (perspective shared by a majority).

	N	Task2			Task4		
		No breach	Intended breach of a norm	Only other reasons	No breach	Intended breach of a norm	Only other reasons
GER experts	17	4	**9**	4	2	4	11
TW experts	19	2	3	14	1	**11**	7

Table 1: Summary of Coding.

We present these findings for each vignette, highlight the differences, and illustrate them with sample answers, as far as this is possible within the space limits of this report. Regarding vignette task2, the majority of the German experts saw the breach of a norm as intended and evaluated the vignette negatively as the teacher did not make optimal use of the opportunity to focus on modeling processes (see GER1_8). In the Taiwanese sample, only 3 out of the 19 experts saw the intended breach of a norm. Some German experts, as well as Taiwanese experts, criticized the dealing with the algebraic demands or the appropriateness of the task implementation in respect to practical concerns, for example, whether the classroom discussion should better be complemented by written notes. Unlike any German expert, six experts from Taiwan were concerned about the structure of the teaching sequence from a content perspective, for example, whether it is appropriate to mix up questions of quadratic functions and quadratic equations or whether the teacher managed to focus on flexible use of different solution strategies (see TW27).

GER1_8: T focuses obviously on solving the quadratic equation, while the modeling aspects contained in the task are hardly or not at all addressed. The following questions are therefore not clarified: - Mark in the illustration what is to be calculated. - How did you come up with your educated guesses? Can the illustration be used to justify which educated guess is particularly realistic? - Why is the approach of S1 correct? - What is described by the solution -4? What is the difference between the real-life situation and the descriptive function?

TW27: [...] The key message that the problem was to solve a quadratic equation with one variable and that there is not only one solution strategy was not delivered.

Regarding the vignette task4, the majority of the Taiwanese experts saw the breach of a norm as intended and criticized that the teacher did not make optimal use of the opportunity to discuss the pros and cons of variable assignment (see TW28). In the German sample, only 4 out of the 17 experts saw the intended breach of a norm. As other reasons for a negative evaluation, Taiwanese experts, as well as German experts, mentioned that the teacher does not build enough on students' thinking or that s/he works out relevant steps instead of the students. Unlike the Taiwanese experts, 8 German experts saw a lack of focus on the equivalence of the two systems of equations that resulted from different variable assignments (see GER2_13). As above, we found hence a kind of reasoning within the German responses that we did not see in the Taiwanese responses.

TW28: 1. The last line of teacher T's statements ran too fast. It was obvious that some students expressed their preference for the second method, the teacher insisted that everyone uniformly learned the first method, and the lesson immediately progressed to solving the problem without spending time on discussing how to choose "groups" to set the unknowns. 2. Some students preferred the second method, maybe because they could only set

the unknowns based on what the problem asked. Although the first method was easy to solve, the students did not know how to choose which variables in the problem were appropriate to set the unknowns. The teacher must spend time discussing with the students how to set the unknowns rather than skipping and proceeding to solve the system of equations.

GER2_13 The teacher discusses the two models exclusively under the aspect of computational simplicity. The central phenomenon of equivalent modeling of a situation and the interesting insight that both models are algebraically identical is not addressed. In addition, the problem arises that the two systems do not emerge through one of the usual ways of transposing an equation, but through substitution. This is obscured by the identical naming [note: the teacher uses x, y in both systems of equations with different meanings] and is not discussed further.

DISCUSSION

This study shows that despite the international consensus regarding the relevance of tasks' potential for mathematical learning and its use, the specific understanding may differ between cultures. First, our symmetric approach of designing vignettes within the national research teams in Germany and Taiwan differs significantly from typical approaches in cross-cultural research, as it is aimed at culturally sensitive vignettes. The presented study on two such vignettes with a sample of experts from each country explored whether the vignettes reflected indeed different culture-specific norms (and not only the particular view of the authors). By means of two vignettes focusing on word problems, we showed that perspectives of German and Taiwanese experts are different, but a) within each culture in line with the expectations of the research team members. Moreover, b) the differences in reasoning between the German and Taiwanese experts are in line with described cultural differences: In the case of task2, the concerns exclusively found in Taiwan resonate with the focus of East Asian mathematics education on the mathematical content and the product-oriented perspective on establishing flexible solving strategies. In the case of task 4, the unique German reasoning referred to a perceived potential of the task for the aim of a meaningful understanding of relations between different mathematical models of a situation rather than its potential to apply specific strategies of variable assignment.

The study also has some limitations. First, a study based on two vignettes regarding word problems in secondary algebra is, of course, not generalizable, but may rather serve as a proof of existence for cultural differences that call for further research. However, the overarching research project TaiGer Noticing could also uncover culture-specific norms of responding to students' thinking between Taiwan and Germany. Second, the brevity of this report allows only a first analysis based on the distinction between answers that reflect the intended breaches of norms and other reasons. An in-depth analysis of professional knowledge and other resources that shape the experts' evaluation is still missing and could substantiate our findings.

Despite these limitations, the study shows that the understanding of the potential of a task for mathematical learning and its adequate use may be inflicted by cultural differences. To illustrate possible consequences for research: If we would have used our data for assessing the noticing of the experts (note: the data was not collected for this purpose), the German experts would have largely missed the noticing target of task4, what was easy for their colleagues from Taiwan to notice, and, at the same time, the Taiwanese experts would have been outperformed by the German experts on task2. It should be discussed how these findings can inform future comparative studies, for example, of instructional quality or teacher noticing, where researchers always face the challenge of balancing the validity of instruments within cultures and their comparability across cultures.

Acknowledgements

This study is part of the project TaiGer Noticing which is funded by the DFG – German Research Foundation (grant numbers DR 1098/1-1 and LI 2616/2-1) and the Ministry of Science and Technology (MOST, grant number 106-2511-S-003-027-MY3).

References

Chang, Y.-P., Krawitz, J., Schukajlow, S., & Yang, K.-L. (2020). Comparing German and Taiwanese secondary school students' knowledge in solving mathematical modelling tasks requiring their assumptions. *ZDM*, *52*(1), 59–72.

Clarke, D. J. (2013). International comparative research into educational interaction: Constructing and concealing difference. In K. Tirri & E. Kuusisto (Eds.), *Interaction in educational settings* (pp. 5–22). Rotterdam: Sense Publishers.

Dreher, A., Lindmeier, A., Feltes, P., Wang, T.-Y., & Hsieh, F.-J. (2021). Do cultural norms influence how teacher noticing is studied in different cultural contexts? A focus on expert norms of responding to students' mathematical thinking. *ZDM*, *53*(1), 165–179. https://doi.org/10.1007/s11858-020-01197-z

Leung, F. K. S. (2001). In Search of an East Asian Identity in Mathematics Education. *Educational Studies in Mathematics*, *47*(1), 35–51.

Pratt, D. D., Kelly, M., & Wong, W. S. S. (1999). Chinese conceptions of "effective teaching" in Hong Kong: Towards culturally sensitive evaluation of teaching. *International Journal of Lifelong Education*, *18*(4), 241–258.

Sherin, M. (2007). The Development of Teachers' Professional Vision in Video Clubs. In *Video Research in the Learning Sciences*. Routledge.

Stein, M. K., & Lane, S. (1996). Instructional Tasks and the Development of Student Capacity to Think and Reason: An Analysis of the Relationship between Teaching and Learning in a Reform Mathematics Project. *Educational Research and Evaluation*, *2*(1), 50–80.

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM*, *52*(1), 1–16.