



# A commentary on the Special Issue “Innovations in measuring and fostering mathematical modelling competencies”

Peter Frejd<sup>1</sup> · Pauline Vos<sup>2</sup>

Accepted: 27 October 2021 / Published online: 16 November 2021  
© The Author(s) 2021

## Abstract

This is a commentary on the ESM 2021 Special Issue on Innovations in Measuring and Fostering Mathematical Modelling Competencies. We have grouped the ten studies into three themes: competencies, fostering, and measuring. The first theme and the papers therein provide a platform to discuss the cognitivist backgrounds to the different conceptualizations of mathematical modelling competencies, based on the modelling cycle. We suggest theoretical widening through a competence continuum and enriching of the modelling cycle with overarching, analytic dimensions for creativity, tool use, metacognition, and so forth. The second theme and the papers therein showcase innovative ideas on fostering and on the definition and analysis thereof. These reveal the need for a social turn in modelling research in order to capture aspects of student collaboration and agency, as well as tensions in fostering when tasks are derived from real-world scenarios, but socio-mathematical norms come from the (pure) mathematics classroom. The third theme, measuring, and the papers therein offer insights into the challenges of positivist research that aims to develop innovative measurement instruments that are both reliable and valid, particularly in light of student group work, cultural background, and other socio-cultural aspects. Drawing on the three discussions, we go on to make recommendations for further research.

**Keywords** Cognitivism · Competencies · Fostering · Mathematical modelling · Measuring · Positivism · Social turn

---

✉ Peter Frejd  
peter.frejd@liu.se

Pauline Vos  
pauline.vos@uia.no

<sup>1</sup> Linköping University, Linköping, Sweden

<sup>2</sup> University of Agder, Kristiansand, Norway

## 1 Introduction

This is a commentary on the ESM Special Issue (SI) on *Innovations in Measuring and Fostering Mathematical Modelling Competencies*. The wider research field of this SI concerns mathematical modelling education, that is, education in solving real-world problems through the use of mathematics. The theme of this SI, however, is more specialized. It brings together two different verbs, *fostering* and *measuring*, and an object of research, *mathematical modelling competencies*. The underlying principle is namely that for students to become successful modellers, they need *mathematical modelling competencies*, which should be *fostered* through practical modelling work in mathematics classrooms, assisted by competent teachers. To study the effects of fostering and to compare different types of fostering, robust research is necessary, and this must be based on sound, theoretical conceptualizations of analytic constructs, and reliable and valid research instruments in order to *measure* mathematical modelling competencies. However, Schukajlow et al. (2018) have observed a *research gap*, that is, a lack of methodological studies on the development of reliable and valid research instruments for measuring modelling competencies. This SI aims to close this research gap. In this commentary, we will look critically at the ten presented studies, at how they innovate research and fill the observed research gap, and at the challenges addressed by the studies. Finally, we offer recommendations for further research.

We begin by introducing the theoretical angle of this commentary. In mathematics education research, that is, to say not focusing on modelling education, in the 1990s a shift in theoretical research stances occurred, which became known as the *social turn* (Lerman, 2000). The rationale was that many research observations could be better interpreted, in the sense of Max Weber's *Verstehen*, by including interacting contexts beyond the individual's mind. For example, researchers observed mathematical thinking impacted by socio-mathematical norms (Yackel & Cobb, 1996), mathematics classroom processes impacted by a *didactical contract* (Brousseau, 2002), curricula impacted by out-of-school requirements (Morgan & Sfard, 2016), and so forth. Consequently, researchers began to incorporate interpretive and reflective methods (e.g., ethnographic, semiotic, or discourse analysis) and socio-cultural theories (e.g., Bernstein, Bourdieu, Brousseau, Freire, Vygotsky, etc.). This *social turn* created a shift in journals such as ESM and JRME, as evidenced by decreasing numbers of papers based on psychometric methods and psychological theories (Inglis & Foster, 2018). However, within the body of research on mathematical modelling, this *social turn* is less conspicuous—as demonstrated in review studies by Geiger and Frejd (2015), Kaiser and Brand (2015), and Stillman (2019)—which may be due to a desire in the modelling community to include mathematicians with an interest in teaching (Houston et al., 2009). Given the weaker *social turn* in research on mathematical modelling education, a considerable part of research in this area has retained a more psychological or cognitivist focus on the attributes of individuals. This focus cannot easily include socio-cultural aspects of modelling, such as collaboration between students, although this is considered a vital aspect of mathematical modelling (Blum, 2002). In addition, cognitivism cannot easily deal with the tensions which typically emerge from modelling problems being situated in the socio-cultural environment outside of school while being solved *in* a mathematics classroom. We, the authors of this commentary, have been influenced by the social turn and have applied socio-cultural theories from Bourdieu, Sfard, and Vygotsky in our research on mathematical modelling (Frejd, 2010, 2020; Hernandez-Martinez & Vos, 2018, Vos & Frejd, 2020; Vos et al., 2020; Vos, 2020; Årleback & Frejd, 2013). We will therefore

include socio-cultural aspects of modelling in this commentary, for instance, regarding student collaboration and task situations.

We will discuss the ten papers by grouping them into three categories organized under the themes of the SI, namely, *competencies*, *fostering*, and *measuring*. Each category will include an introductory paragraph to provide a context for our commentary, and at the end of each section, we will present questions or make suggestions that will be discussed in Section 5.

## 2 Competencies

While the term *competence* is used in everyday language and in scientific research, it lacks a unanimous conceptualization. The term is defined in encyclopedias as “the set of demonstrable characteristics and skills that enable and improve the efficiency or performance of a job” (Competence, 2021). The term was developed from 1959 onwards in the research discipline known as cognitive psychology, in which the leading theory is known as cognitivism. Cognitivism was developed to counter behaviorism, which focused on observable behavior. By contrast, cognitivist researchers study, for example, how students think when reasoning mathematically, making errors, and so forth.

The German psychologist Franz Weinert (1930–2001) lays the foundations for conceptualizing the term *competence* (1999, as cited in Röhr-Sendlmeier & Käser, 2017). Five studies in this SI cite his work. He advised that competence be defined cognitively and that aspects such as affect or willingness be avoided (Weinert, 1999 as cited in Röhr-Sendlmeier & Käser, 2017), which represents a cognitivist approach to the study of how individuals think without the inclusion of “complicating factors like affect and cultural issues in an attempt to simplify the research tasks” (McLeod, 1989, p. 252). Below, we will elaborate on two papers (Cevikbas et al.; Lu & Kaiser) that discuss the tensions which arise when including or excluding certain aspects in or from *mathematical modelling competencies*.

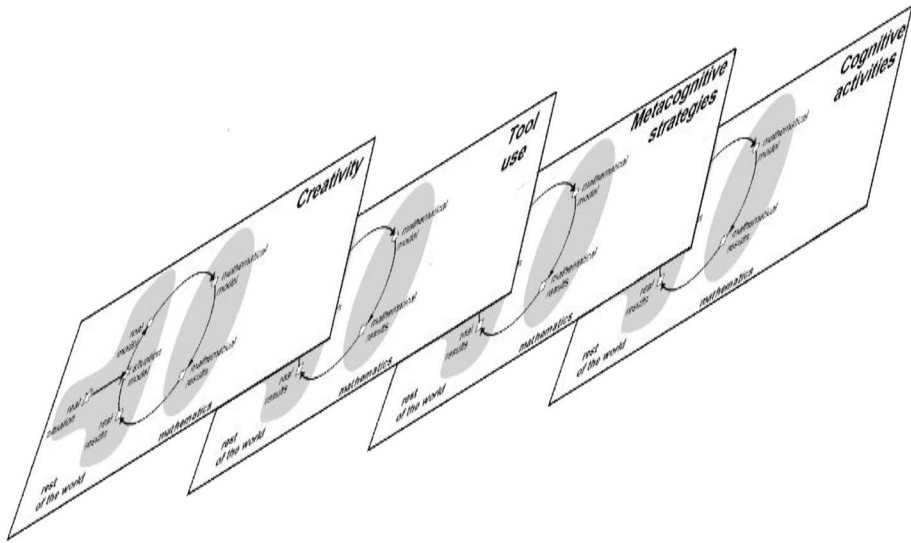
Cevikbas et al. present a comprehensive literature review of research on mathematical modelling competencies. They apply an analytic framework with three approaches to mathematical modelling competencies: (1) a *holistic* or top-down approach, which perceives modelling competency (singular) as one of several mathematical competencies; (2) an *analytic* or bottom-up approach, which assumes that a set of interconnected competencies and sub-competencies is needed for modelling, with optional additional aspects, such as willingness (Maaß, 2006), readiness to act (Kaiser, 2017) and metacognition (Stillman, 1998; Vorhölter, 2018), and *global competencies*, such as social and communication abilities not typical for modelling, but nevertheless necessary in it; and (3) *further* approaches, which encompass approaches not included in the other categories.

The first two approaches have in common an understanding of mathematical modelling as a cyclic process represented by *the modelling cycle* (Niss & Blum, 2020; Geiger & Frejd, 2015). This modelling cycle describes the various phases of modelling, from rendering the problem context mathematizable to interpreting and validating the mathematical results. We observe that both the holistic and analytic approaches adopt a primarily cognitivist perspective, and that both operationalize modelling competency/ies based on the modelling cycle. Consequently, whether research was initiated from a holistic or analytic approach cannot be clearly discerned from the empirical results. Cevikbas et al. show that current research on mathematical modelling competencies is largely conducted by German researchers using the analytic approach to study, among others, (sub-)competencies, instructional strategies, and measurement instruments. They call for the extension

of the theoretical conceptualizations of mathematical modelling competencies. We agree with this call and perceive several options, for instance, using the work by Blömeke et al. (2015), to which they refer. Blömeke et al. (2015) explain competence as an analytic construct existing on a continuum and ranging from competences that are observable and/or dichotomously measurable through clinical tests, to competences that are needed by experts in complex, holistic, real-life situations. These expert competencies are “tested” through real-life tasks by evaluating the useability of the solution against criteria from a client. This perspective of a continuum can be translated to mathematical modelling competences, ranging from the competencies of students as assessed by paper-and-pencil tests, in which cognitive components are “more easily” identified, to the modelling competences necessary for professional practice (Frejd & Bergsten, 2016, 2018). Expert modellers need mathematical, social, technological, and critical abilities, for example, regarding *communication* with clients; they need to select and employ *technology*, to perform *critical analysis* of the usefulness and effectiveness of the models developed, to discuss and present results, and so forth. Situated towards the middle of this continuum are Group Modelling Competencies (see, e.g., Watson et al., 1991), a construct that still needs further conceptualization and that applies to an important aspect of mathematical modelling (Blum, 2015). Thus, theorizing on mathematical modelling competencies can further build on Blömeke et al.’s (2015) continuum, by including modelling competencies beyond those that are observable and measurable. A second suggestion for theory enrichment follows the discussion of the next paper.

Lu and Kaiser add an interesting enrichment to the conceptualizations of mathematical modelling competencies. They measure mathematical modelling competencies while also analyzing student *creativity* as another dimension of modelling, thus continuing the important work by Wessels (2014). They conceptualize creativity in modelling as an overarching dimension to the cognitive activities in the modelling cycle. By positioning creativity in all phases of the modelling cycle, the standard cognitivist conceptualization of modelling is enriched. This conception of creativity as mediating cognitive activities is truly innovative. Similarly, Greefrath et al. (2011) explain that *tool use* overarches the modelling cycle. That is to say that in any phase of the modelling cycle, various (digital) tools can be used, for instance, in simulating the problem situation, in performing calculations, or in presenting results visually. When the analytic dimension of tool use is added, researchers can analyze how tools *mediate* the cognitive activities. Metacognitive strategies (Stillman, 1998; Vorhölter, 2018) also play a role in all phases of the modelling cycle, which then can be considered a further overarching analytic dimension to the cognitive activities depicted by the standard modelling cycle.

Blum (2015) conceptualized the modelling cycle through the inclusion of two different worlds, the mathematical and the real. These two worlds differ in terms of socio-cultural norms and conventions, and this can be acknowledged in theorizing competences. For instance, the mathematical world makes demands of mathematical rigor and precision, whereas the real world is accepting of rules of thumb and quick estimates (Williams & Wake, 2007). Being allowed to use extra-mathematical knowledge or own creative inventions in mathematical modelling tasks in the (pure) mathematical discourse might, therefore, confuse students. We have noticed that Lu and Kaiser’s operationalization of creativity focuses primarily on phases in the mathematical world (mathematizing and working mathematically). Yet, creativity also plays a role in understanding the problem situation. Students may, for example, employ a variety of resources, and perhaps even interview the “client” about the problem. Creativity may also play a role in explaining and visualizing



**Fig. 1** Analytic dimensions for mathematical modelling competencies

results. These two omitted aspects of creativity pertain to the “real world” side of the modelling cycle, which could be further elaborated.

Thus, regarding the theorizing on mathematical modelling competencies based on the modelling cycle, we suggest (1) enriching the cognitive dimension with other analytic, interacting dimensions, such as creativity, tool use, metacognition, and potentially further dimensions; and (2) including socio-cultural differences between the real world and the mathematical world (see Fig. 1).

Our suggestion of theorizing may influence some research interpretations and explain why some blockages in modelling persist. For instance, student difficulties regarding “assumption making” are interpreted as *cognitive blockages* when analyzed in the cognitive dimension (e.g., Galbraith & Stillman, 2006). However, when considering socio-cultural aspects in the two worlds, these may be re-analyzed as *cultural blockages*, since the didactical contract of the “mathematical classroom world” limits student agency and discourages creativity (Brousseau, 2002). Cultural blockages in the learning environment may be more difficult to overcome than cognitive blockages in an individual student.

### 3 Fostering

The term fostering is frequently used in educational research, as in the fostering of student discussion, of inquiry skills, of love for the subject, of resilience, and so forth. In mathematics education research, we see fostering of number sense at pre-school level (Baroody et al., 2012), and the fostering of collaboration (Chih, 2021) and creativity (Munakata & Vaidya, 2013). Fostering is rarely defined, so the term may convey meaningful information with an implicitly accepted meaning agreed by all. In encyclopedias, fostering is defined as “to promote the growth or development of [something]” and is used in sentences such as “[s]uch conditions foster the spread of the disease” (Merriam-Webster, n.d.). It appears to be associated with encouraging, nurturing, and facilitating. It takes time, is goal-directed,

and requires interactivity between a subject doing the fostering and an object being fostered.

Fostering clearly differs from direct instruction. We contend that fostering is more student-centered, more informal, more experiential, and more inspiring. By contrast, direct instruction is more teacher-centered, more content-directed, and more explanatory (Stephan, 2020). In this commentary, we lack the space to further discuss the many perspectives on teaching (passing on knowledge, facilitating learning, enculturating, inducing, etc.). We note, however, that mathematical modelling competencies cannot be taught through direct instruction since modelling is not a *spectator sport* (Blum, 2015). Instead, these competencies can be fostered through varied modelling experiences. This point is illustrated in a number of papers in this SI.

Brady and Jung explore the nature of emerging classroom modelling cultures. They enable the fostering of the social modelling competencies *collaboration* and *consultation* in modelling projects by adding a client perspective—an innovative characteristic of modelling tasks. The students participating in this study worked, among others, for the Alaska Department for Fish and Game and for volleyball tournament organizers. The modelling tasks were open-ended and had no “correct” solution and the developed models were discussed from various perspectives, thus fostering students’ *agency*. In addition, students had ample time for their modelling work and opportunities to adapt their presentations. Student presentations were analyzed qualitatively and sub-categorized into the sub-processes of *working mathematically*, *interpreting*, *validating* and *patching*. The latter term is defined as a “[p]resentation discourse that focuses on unruly features of the problem and/or explains adaptations or exceptions made when applying their model” (p. 10). Brady and Jung’s innovative construct of patching puts a focus on student reasoning regarding potential or performed adaptations to their models. Patching thus fosters student awareness of the providing and receiving of feedback in order to negotiate and justify their mathematical models. The major finding of this study is that the classroom culture is a matter of negotiations between the participants involved in the activity, which means that the fostering of mathematical modelling competencies does not depend solely on well-designed tasks and competent teachers. Researchers also need to take into consideration social interactions.

Durandt et al.’s study evaluates two different teaching approaches in engineering education by measuring students’ competencies and attitudes in modelling. One approach was independence-oriented teaching and included groupwork; thus, it was closer to “fostering.” The other consisted of traditional teaching with direct instruction and individual work. The intervention lasted five lessons, and in both approaches, the same tasks were completed. Some tasks were more open and made students “run” a full modelling cycle, whereas others focused on just one aspect of the modelling cycle, such as interpreting a graph. Besides a test, the researchers also developed an innovative instrument to measure attitudes, the Survey of Attitudes Towards Mathematical Modelling (SATMM) with 6 scales (affect, perceived competence, value, difficulty, interest, effort). Their results show non-significant, but “descriptively more positive attitudes” (p. 17.) experienced by the students who experienced the independence-oriented approach. The study demonstrates that a balance between independent work and teachers’ guidance is beneficial for fostering modelling competencies.

Geiger et al. have developed the framework “Design and Implementation Framework for Mathematical Modelling Tasks (DIFMT).” Their study concerns the fostering of teacher and researcher *collaboration* to develop teaching practices suitable for fostering such competencies among students. Their research process of design-implement-reflect, which includes regular meetings with researchers and teachers, demonstrates that fostering is an iterative process. The DIFMT framework includes principles regarding both the design

of modelling tasks and classroom implementation of these. One aspect of the framework is termed *pedagogical architecture*, a truly innovative construct to identify socio-cultural aspects in a classroom. Other principles identify important teacher competencies for fostering modelling, such as knowledge about the nature of a problem, understanding of the modelling processes, and how findings should be presented. We want to highlight that Geiger et al.'s DIFMT framework includes *implemented anticipation*, a term, originally from Niss (2010), that describes a modeller's foresight in earlier phases of the modelling regarding that which is likely to be mathematically useful in later phases. Geiger et al. claim that teacher anticipation "is key to both the design and implementation of modelling tasks, as teachers must anticipate how students will respond, what scaffolds should be prepared, and where challenges are likely to emerge" (p. 6). Thus, anticipation is a competence to be fostered among teachers, and its fostering takes time.

A final innovative study, which also addresses the fostering of teacher anticipation, is that of Alwast and Vorhölter. Video clips of staged classroom situations were used to foster and measure the noticing competencies of pre-service teachers. By this is meant a teacher's abilities to attend to the experiences of students, as they happen in the classroom, without automaticity or habits (Mason, 2002). This relates to anticipation and to being prepared to deal with unexpected challenges in a flexible and creative manner. Video clips showing student groups working on modelling problems can be useful in teacher education to foster teacher flexibility and creativity in classroom practice regarding social norms and expectations in mathematical modelling activities. The video clips in this study showed students' difficulties regarding sub-competencies in understanding and simplifying real-world problems, and working mathematically, but the sub-processes of evaluation, validation, and presentation of the solution to a "client" were not included in the clips.

The four articles above provide empirical evidence of how mathematical modelling competencies, such as collaboration, negotiation, presentation, and agency, can be fostered both for students and for pre-service teachers. Moreover, it is also possible to foster the fostering competencies of teachers. However, we end with a theoretical question: how is fostering related to teaching? And we pose a methodological question, which takes us also into the next paragraph: if we want to foster group modelling competency, group creativity, group attitudes, and so forth, how can these be measured without disrupting the fostering process?

## 4 Measuring

Research methods and analytic theories are intertwined, both connected by the *research paradigm*, which embraces beliefs and norms regarding that which is regarded as "good" research (Burton, 2002). Cognitivism connects well to measurement research, following Weinert's writings on measuring the cognitive competencies of students in large-scale surveys (1999 as cited in Röhr-Sendlmeier & Käser, 2017). The instruments (questionnaires, tests) are administered at the individual level. However for analysis and the reporting of results, data are aggregated, making participants invisible as individuals. The associated research paradigm is known as *positivism* (Bryman, 2016), and strives to gain knowledge in a similar way as in the natural sciences by establishing as objective a knowledge as possible through measurement, with special emphasis on the reliability of instruments and data analysis. One problem which arises from interpreting the results is that an observed correlation does not necessarily imply causality. Secondly, measurements are disturbed

by social and cultural circumstances, creating “noise” and uncertainty over what is actually being measured. This translates into tensions between the *reliability* and the *validity* of instruments and analysis, which are captured by the well-known *attenuation paradox*, which states that increased reliability is achieved at the expense of validity (Nicewander et al., 1977). In educational measurement research, therefore, the development of valid measurement remains a challenge, as also described by Lu and Kaiser. Thirdly, we follow Clarke’s (1996) critique by noting that measurement can be an invasive form of assessment when it disrupts learning processes. Instead, researchers can turn to portraying competencies of individual students, of their group work, or of the classroom culture. In this paragraph, we discuss four papers in this SI with an eye on these three challenges.

Krawitz et al. study the effect of reading prompts on creating a *real model*. This was measured with 9th grade students and compared between two conditions (with and without prompts) and between two educational environments (Taiwan and Germany). This study is truly positivist, creating a measurement that is as reliable as possible, to enable statistical analyses. For instance, the students were randomly distributed between conditions, and were given a supplementary test with pure mathematics tasks, not to foster mathematical modelling, but to statistically *control* samples. Furthermore, the German students were selected from the Gymnasium track—by homogenizing the German sample, variation was reduced, thereby increasing Cronbach alpha. The students were given relatively difficult tasks; only 8% of the Taiwanese and 17% of the German Gymnasium students were successful, which makes the nature of measurement rather invasive (Clarke, 1996). In addition, the tasks appear artificial. For example, the exemplary parachute task asks for a diagonal distance travelled under certain crosswinds. This distance is necessary for testing student knowledge of Pythagoras’ theorem, but unnecessary in real-life parachuting, where it is the horizontal distance which is needed to avoid landing in the sea, against rocks, or in urban areas. Moreover, the illustration shows a steerable parachute, which can go upwind, whereas the task is about non-steerable parachutes. Thus, a student with knowledge about parachuting may be confused by incongruency between task, figure, and real-life scenarios. We therefore advise further research into cross-cultural aspects in tasks, for instance, whether a task which includes references to a military or expensive leisure context is valid in cross-national comparison. In addition, the topic of reading prompts focuses on mathematical modelling as a paper-and-pencil test activity where the reading of task situations requires support. While reading prompts may ease an invasive testing regime, they are unnecessary when fostering student modelling competencies through collaborative work, since students are then able to investigate situations and look up information on the internet.

The study by Cai et al. deals with the modelling competencies of mathematics teachers, as well as their *noticing* competencies regarding student answers. The study aimed to compare two groups of mathematics teachers in the USA: 21 pre-service teachers (novices) versus 21 recipients of a prestigious teacher award (experts). This study is not purely positivist: it is not large-scale, and the instrument involved two open modelling tasks. Having completed these, the teachers were given a number of solutions by individual students, on which they were asked to comment in a free, narrative format on the appropriateness of the students’ modelling approaches. Through an inductive, grounded theory approach, the researchers developed codes that portray teachers’ modelling competencies and noticing competencies. Since these codes were developed transparently in alignment with the data, they have a high level of validity. We consider it interesting that the resultant codes differed to quite an extent between the two tasks, both regarding teacher solutions and their reflections on student answers. Thus, the codes are task-dependent, but likely also dependent on the answers given by students, on the participating teachers, and on the instructional



culture. There also emerged common codes between the tasks, such as offering praise as part of feedback, which may be more consistent across tasks/populations/etc., and this cross-task consistency could be an object of further research. We recommend the replication of this study with other scenarios, in particular with group work, and with modelling tasks that address realistic problems from a *client*, see Brady and Jung. We disadvise the reuse of the artificial problems given. For instance, in the "seashell task," seashells are placed on a scale and the task asks for an estimation of the weight of some additional seashells. In real life, one would likely observe the readings given on the scale! With more fostering of mathematical modelling, we also hope to see teachers noticing *task flaws*, which will result in codes such as "need for task redesign."

A similar study on measuring teacher modelling competencies and their didactical competencies was carried out by Yang et al., focusing on the comparison of pre-service teachers in three countries (Mainland China, Germany, and Hong Kong) and collecting large-scale data. The same type of instrument as in Cai et al. was used, with small differences: four, rather than just one, didactical questions were posed regarding the presented student answers. Furthermore, codes were not inductively developed from the data. Instead, teacher answers were evaluated on an ordinal scale (from high to low). This results in codes which are more unified than in Cai et al. and more suitable for statistical analysis. This increased reliability comes at the expense of validity, since details are lost and percentages of high, average, or low levels conceal the phase of modelling at which the blockages occur, as well as how praise was offered to students by teachers as part of their feedback. By obscuring details of individuals and their reasoning, this study has more validity issues, in particular cross-culturally.

The study by Greefrath et al. is a pre-test–post-test experiment focusing on pre-service teachers regarding the effect of two different types of modelling seminars on their pedagogical content knowledge for the teaching of mathematical modelling (PCKMM). A control group of pre-service teachers who had not received any professionalization on modelling was also tested. The modelling seminars focused either on (1) modelling tasks, the analysis and development of these, and without a focus on adaptive teacher interventions; or on (2) adaptive teacher interventions, preserving students' independence, and without a focus on developing tasks. The paper is positivist, drawing on sophisticated statistical notions and dehumanizing the participants by writing about "test subjects" (p.14). The results showed that both experimental groups improved on their PCKMM, the "task group" to a greater extent than the "intervention group," whereas the control group did not. The seminars were thus effective, and different emphases were shown to yield different effects. However, with such positivist research, differences, while observed, cannot easily be interpreted, especially with an opaque test instrument. Fortunately, this instrument was published elsewhere (Wess et al., 2021); closer scrutiny of it gives reason for concern, again, regarding validity. In order to measure PCKMM, the instrument had 71 items on four scales. Two of these—"knowledge about interventions" and "knowledge about tasks"—aim to differentiate between the "intervention group" and the "task group." Knowledge about interventions was measured by 24 complex, verbose items that asked participants to evaluate a teacher intervention as "suitable," "unsuitable," or "I don't know." Such complex items concerning teacher-student interactions require *analytic skills*, which are fostered through many experiences, and were thus feasible for students in both conditions. By contrast, the 17 items on knowledge about tasks were all yes/no questions on characteristics of modelling tasks (see Fig. 2). These items ask for *jargon knowledge*, which students in the "task group" likely had learnt in their seminar. Differences between item types can thus explain differences in scores between groups. In other words, the instrument may have led to the (lack of) differences that it intended to measure. At a more general level, the use of dichotomous

5.1 Characteristics of modelling tasks			
5.1.1.	Modelling tasks...	True	False
	... can be underdetermined.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	... can be overdetermined.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	... are as closed as possible.	<input type="checkbox"/>	<input checked="" type="checkbox"/>
5.1.2.	Modelling tasks...	True	False
	... can also be Fermi tasks.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	... are always also “dressed-up” word problems.	<input type="checkbox"/>	<input checked="" type="checkbox"/>
	... can also be context-related word problems.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
5.1.3.	Modelling tasks...	True	False
	... are as close to real life as possible.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	... are as authentic as possible.	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	... are as relevant as possible to the students.	<input checked="" type="checkbox"/>	<input type="checkbox"/>

**Fig. 2** Excerpt of instrument to measure teachers’ PCKMM (Wess et al. 2021, p. 101)

items should be avoided (Bryman, 2016), since they disallow nuance, discussion, or conditional answers such as “it depends...” or “yes, unless...,” which is typical of the dialogic and reflective nature of mathematical modelling. We advise future adaptations to include more open tasks, focusing on the collaborative analysis of classroom situations, and to ask (teams of) teachers to develop their own modelling tasks, as in Geiger et al.

The four studies discussed in this paragraph provide reliable instruments for the measurement of mathematical modelling competencies to narrow the research gap. With the exception of Cai et al., these papers reflect a positivist research approach, favoring reliability over validity. However, the papers by Cai et al., and by Brady and Jung and Alwast and Vorhölter, discussed earlier, illustrate that measurement can have a higher validity when striving to portray competencies through a combination of measurement and fostering. We can therefore pose a more specific question than the one central in this SI: how can we develop measurement instruments and methods that have high validity; that can portray more complex competencies on the competence continuum beyond those assessed through paper-and-pencil tests (Blömeke et al., 2015), in particular group competencies; that do not invasively disrupt that which is being fostered; and that moreover have a reasonable reliability?

## 5 Discussion and conclusion

This SI focuses on mathematical modelling *competencies* related to the modelling cycle, a topic frequently addressed in research on mathematical modelling education, as observed by both Stillman (2019) and Geiger and Frejd (2015). This focus yields a productive research corpus, bringing mathematical modelling to the attention of the wider

mathematics education research community and providing a clear connection to competence-oriented curricula. It also yields methodologies and empirical evidence, on which teachers, researchers, and policy makers can further build. We observed that the SI fills the observed research gap mentioned in Section 1. This SI presents innovative instruments and methods for measuring pre-service teacher noticing skills in modelling contexts (Alwast & Vorhölter; Cai et al., Greefrath et al.), attitudes (Durandt et al.), and creativity (Lu & Kaiser). In some studies, the qualitative work was transformed into measurement by carefully and transparently developing adequate codes (Alwast & Vorhölter; Brady & Jung; Cai et al.; Geiger et al.). We argue that a cognitivist analysis, focusing on data at the individual level, challenges the collaborative aspect in modelling. A positivist research stance also provides high reliability at the expense of validity.

Some studies in this SI (Brady & Jung; Geiger et al.) demonstrate that fostering is about the development of a productive classroom culture. Mathematical modelling is teamwork (Blum, 2002) and fostering collaboration is *not* direct instruction (Chih, 2021). Geiger et al. provide a productive example on how fostering relates to teaching, where researchers and teachers collaboratively and iteratively adapt the didactical contract in mathematical modelling. To foster the multidimensional sense-making goals of modelling (Blum, 2015), new possibilities may be offered by team-based learning (TBL); "while the assumption that teaching is 'imparting' knowledge tends to be concerned only with knowledge acquisition, the studies we examined suggest that the benefits of TBL extend well beyond this singular learning goal" (Haidet et al., 2014, p. 308). However, a broader perspective on mathematical modelling, which considers the cultures, norms, and classroom expectations, requires further research on *global competencies* (Cevikbas et al.). Some competencies are easier than others to measure (Blömeke et al., 2015). If we therefore assume that mathematical modelling is a team activity to be fostered, it is also necessary to develop reliable and valid research tools to measure *group competencies of mathematical modelling, group attitudes*, etc. Brady and Jung's study provides a springboard, offering other researchers a method and codes to "measure" the modelling culture of a classroom that involves students working in groups on tasks from "clients." In terms of Clarke (1996), their measurement method is not *invasive* and rather *portrays* mathematical modelling competence in a social cultural context. We hope this SI inspires other researchers to continue this important work regarding both the validity and reliability of instruments and analyses.

Cevikbas et al. conclude that "there is a great need to investigate modelling competencies using a variety of theoretical frameworks and to extend existing frameworks by using innovative approaches" (p.19). In this commentary, we have introduced perspectives that can further serve theory development. Firstly, we drew on Blömeke et al.'s (2015) competence continuum, which we adapted to the conceptualization of modelling competencies. This provides theoretical principles for the classification of both observable competencies measured by clinical tests and the more complex holistic competencies that are difficult to measure but can be portrayed. This competence continuum puts classroom work in a perspective that includes professional practices of expert modellers. Secondly, in relation to the modelling cycle, we suggest overarching dimensions to the cognitive dimension, such as creativity, tool use, and metacognition (see Fig. 1). Geiger and Frejd (2015) claim that the local theories from mathematical modelling research (the modelling cycle and mathematical modelling competencies) reveal a set of *white spots*, which invite further theorizing. Several of these white spots connect to the *social turn* and address a research gap which adjusts that addressed in this SI, namely regarding the valid measurement, or portraying, of mathematical modelling *group* competencies, of competencies for dealing with

socio-cultural issues in modelling, and, for teachers, of essences of fostering. We propose further developing a social turn in mathematical modelling research.

**Availability of data and material** Not applicable

**Code availability** Not applicable

**Funding** Open access funding provided by Linköping University.

## Declarations

**Conflicts of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ärlebäck, J. B., & Frejd, P. (2013). Modelling from the perspective of commognition – An emerging framework. In G. Stillman, G. Kaiser, W. Blum, & J. Brown (Eds.), *Teaching mathematical modelling: Connecting to research and practice* (pp. 47–56). Springer. [https://doi.org/10.1007/978-94-007-6540-5\\_3](https://doi.org/10.1007/978-94-007-6540-5_3)
- Baroody, A. J., Eiland, M. D., Purpura, D. J., & Reid, E. E. (2012). Fostering at-risk kindergarten children's number sense. *Cognition and Instruction*, 30(4), 435–470. <https://doi.org/10.1080/07370008.2012.720152>
- Blömeke, S., Gustaffsson, J. E., & Shavelson, R. (2015). Beyond dichotomies: Viewing competence as a continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Blum, W. (2002). ICMI Study 14: Applications and modelling in mathematics education—Discussion document. *Educational Studies in Mathematics*, 51(1), 149–171. <https://doi.org/10.1023/A:1022435827400>
- Blum, W. (2015). Quality teaching of mathematical modelling: What do we know, what can we do? In S. J. Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education* (pp. 73–96). Springer.
- Brousseau, G. (2002). *Theory of Didactical Situations in Mathematics: Didactique des Mathématiques, 1970–1990*. Springer.
- Bryman, A. (2016). *Social research methods*. Oxford University Press.
- Burton, L. (2002). Methodology and methods in mathematics education research; where is the why? In S. Goodchild & L. D. English (Eds.), *Researching mathematics classrooms: A critical examination of methodology* (pp. 1–10). Greenwood Publishing Group.
- Chih, T. (2021). *Fostering collaboration in an asynchronous IBL modern geometry course* (pp. 1–25). Problems, Resources & Issues in Mathematics Undergraduate Studies. <https://doi.org/10.1080/10511970.2021.1955781>
- Clarke, D. J. (1996). Assessment. Chapter 9. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International Handbook of Mathematics Education* (pp. 327–370). Kluwer.
- Competence. (2021). In *Wikipedia*. Retrieved July 28, 2021, from [https://en.wikipedia.org/wiki/Competence\\_\(human\\_resources\)](https://en.wikipedia.org/wiki/Competence_(human_resources)).
- Frejd, P. (2010). Revisiting perspectives on mathematical models and modelling. In C. Bergsten, E. Jablonka, & T. Wedege (Eds.), *Mathematics and mathematics education: Cultural and social dimensions: Proceedings of Madif 7, Stockholm, 26-27, January, 2010* (pp. 80–90) SMDF.

- Frejd, P. (2020). Teachers' possibilities to generate science capital for modelling. In G. A. Stillman, G. Kaiser, & E. Lampen (Eds.), *Mathematical modelling education and sense-making* (pp. 39–49). Springer. [https://doi.org/10.1007/978-3-030-37673-4\\_4](https://doi.org/10.1007/978-3-030-37673-4_4)
- Frejd, P., & Bergsten, C. (2016). Mathematical modelling as a professional task. *Educational Studies in Mathematics*, 91(1), 11–35. <https://doi.org/10.1007/s10649-015-9654-7>
- Frejd, P., & Bergsten, C. (2018). Professional modellers' conceptions of the notion of mathematical modelling: Ideas for education. *ZDM-Mathematics Education*, 50(1-2), 117–127. <https://doi.org/10.1007/s11858-018-0928-2>
- Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *ZDM-Mathematics Education*, 38(2), 143–162. <https://doi.org/10.1007/BF02655886>
- Geiger, V., & Frejd, P. (2015). A reflection on mathematical modelling and applications as a field of research: Theoretical orientation and diversity. In G. A. Stillman, W. Blum, & M. S. Biembengut (Eds.), *Mathematical Modelling in Education Research and Practice: Cultural, Social and Cognitive Influences* (pp. 161–171). Springer. [https://doi.org/10.1007/978-3-319-18272-8\\_12](https://doi.org/10.1007/978-3-319-18272-8_12)
- Greefrath, G., Siller, H.-S., & Weitendorf, J. (2011). Modelling considering the influence of technology. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in Teaching and Learning of Mathematical Modelling* (pp. 315–329). Springer. [https://doi.org/10.1007/978-94-007-0910-2\\_32](https://doi.org/10.1007/978-94-007-0910-2_32)
- Haidet, P., Kubitz, K., & McCormack, W. T. (2014). Analysis of the team-based learning literature: TBL comes of age. *Journal on Excellence in College Teaching*, 25(3-4), 303–333 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4643940/>
- Hernandez-Martinez, P., & Vos, P. (2018). "Why do I have to learn this?" - A study from mathematical modelling education about the relevance of mathematics. *ZDM-Mathematics Education*, 50(1-2), 245–257. <https://doi.org/10.1007/s11858-017-0904-2>
- Houston, K., Galbraith, P., & Kaiser, G. (2009). *ICTMA: The first twenty-five years*. History of ICMI. Retrieved July 28 2021, from <https://www.icmihistory.unito.it/ictma.php>.
- Inglis, M., & Foster, C. (2018). Five decades of mathematics education research. *Journal for Research in Mathematics Education*, 49(4), 462–500. <https://doi.org/10.5951/jresmetheduc.49.4.0462>
- Kaiser, G. (2017). The teaching and learning of mathematical modeling. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 267–291). National Council of Teachers of Mathematics.
- Kaiser, G., & Brand, S. (2015). Modelling competencies: Past development and further perspectives. In G. A. Stillman, W. Blum, & M. Salett Biembengut (Eds.), *Mathematical Modelling in Education Research and Practice* (pp. 129–147). Springer. [https://doi.org/10.1007/978-3-319-18272-8\\_10](https://doi.org/10.1007/978-3-319-18272-8_10)
- Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19–44). Ablex.
- Maaß, K. (2006). What are modelling competencies? *ZDM-Mathematics Education*, 38(2), 113–142. <https://doi.org/10.1007/BF02655885>
- Mason, J. (2002). *Researching your own practice: The discipline of noticing*. RoutledgeFalmer. <https://doi.org/10.4324/9780203471876>
- McLeod, D. B. (1989). Beliefs, attitudes, and emotions: New views of affect in mathematics education. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving* (pp. 245–258). [https://doi.org/10.1007/978-1-4612-3614-6\\_17](https://doi.org/10.1007/978-1-4612-3614-6_17)
- Merriam-Webster. (n.d.). Foster. In Merriam-Webster.com dictionary. Retrieved July 28, 2021, from <https://www.merriam-webster.com/dictionary/foster>
- Morgan, C., & Sford, A. (2016). Investigating changes in high stakes mathematics examinations: A discursive approach. *Research in Mathematics Education*, 18(2), 92–119. <https://doi.org/10.1080/14794802.2016.1176596>
- Munakata, M., & Vaidya, A. (2013). Fostering creativity through personalized education. *Primus*, 23(9), 764–775. <https://doi.org/10.1080/10511970.2012.740770>
- Nicewander, W. A., Price, J. M., Mendoza, J. L., & Henderson, D. (1977). The attenuation paradox and the distribution of ability. *British Journal of Mathematical and Statistical Psychology*, 30(2), 204–209. <https://doi.org/10.1111/j.2044-8317.1977.tb00740.x>
- Niss, M. (2010). Modeling a crucial aspect of students' mathematical modeling. In R. Lesh, P. Galbraith, C. Haines, & A. Hurford (Eds.), *Modelling students' mathematical competencies* (pp. 43–59). Springer. [https://doi.org/10.1007/978-1-4419-0561-1\\_4](https://doi.org/10.1007/978-1-4419-0561-1_4)
- Niss, M., & Blum, W. (2020). *The learning and teaching of mathematical modelling* (1st ed.). Routledge. <https://doi.org/10.4324/9781315189314>
- Röhr-Sendmeier, U., & Käser, U. (2017). Kompetenz. In L. Kühnhardt & T. Mayer (Eds.), *Bonner Enzyklopädie der Globalität* (pp. 235–248). Springer. [https://doi.org/10.1007/978-3-658-13819-6\\_19](https://doi.org/10.1007/978-3-658-13819-6_19)

- Schukajlow, S., Kaiser, G., & Stillman, G. (2018). Empirical research on teaching and learning of mathematical modelling: A survey on the current state-of-the-art. *ZDM-Mathematics Education*, 50(1-2), 5–18. <https://doi.org/10.1007/s11858-018-0933-5>
- Stephan, M. (2020). Teacher-centered teaching in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 836–840). Springer. [https://doi.org/10.1007/978-3-030-15789-0\\_150](https://doi.org/10.1007/978-3-030-15789-0_150)
- Stillman, G. (1998). The emperor's new clothes? Teaching and assessment of mathematical applications at the senior secondary level. In P. Galbraith, W. Blum, G. Booker, & I. D. Huntley (Eds.), *Mathematical modelling: Teaching and assessment in a technology-rich world* (pp. 243–253) Horwood.
- Stillman, G. A. (2019). State of the art on modelling in mathematics education—Lines of inquiry. In G. A. Stillman & J. P. Brown (Eds.), *Lines of Inquiry in Mathematical Modelling Research in Education, ICME-13 Monographs* (pp. 1–20). Springer. [https://doi.org/10.1007/978-3-030-14931-4\\_1](https://doi.org/10.1007/978-3-030-14931-4_1)
- Vorhölter, K. (2018). Conceptualization and measuring of metacognitive modelling competencies: Empirical verification of theoretical assumptions. *ZDM-Mathematics Education*, 50(1), 343–354. <https://doi.org/10.1007/s11858-017-0909-x>
- Vos, P. (2020). On science museums, science capital and the public understanding of mathematical modelling. In G. A. Stillman et al. (Eds.), *Mathematical Modelling Education and Sense-making* (pp. 63–73). Springer. [https://doi.org/10.1007/978-3-030-37673-4\\_6](https://doi.org/10.1007/978-3-030-37673-4_6)
- Vos, P., & Frejd, P. (2020). The object-tool duality in mathematical modelling - A framework to analyze students' appropriation of Sankey diagrams to model dynamic processes. *Avances de Investigación en Educación Matemática*, 17, 52–66. <https://doi.org/10.35763/aiem.v0i17.305>
- Vos, P., Hernandez-Martinez, P., & Frejd, P. (2020). Connections of science capital and the teaching and learning of mathematical modelling. In G. A. Stillman, G. Kaiser, & E. Lampen (Eds.), *Mathematical Modelling Education and Sense-making* (pp. 33–38). Springer. [https://doi.org/10.1007/978-3-030-37673-4\\_3](https://doi.org/10.1007/978-3-030-37673-4_3)
- Watson, W. E., Michaelsen, L. K., & Sharp, W. (1991). Member competence, group interaction, and group decision making: A longitudinal study. *Journal of Applied Psychology*, 76(6), 803–809. <https://doi.org/10.1037/0021-9010.76.6.803>
- Wessels, H. (2014). Levels of mathematical creativity in model-eliciting activities. *Journal of Mathematical Modelling and Application*, 1(9), 22–40.
- Williams, J., & Wake, G. (2007). Metaphors and models in translation between college and workplace mathematics. *Educational Studies in Mathematics*, 64(3), 345–371. <https://doi.org/10.1007/s10649-006-9040-6>
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4), 458–477. <https://doi.org/10.2307/749877>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.