Towards competency-oriented mathematics education

An investigation of task demands and teachers' knowledge of task demands from a competency perspective

Andreas Pettersen

Dissertation for the degree of Philosophiae Doctor

Department of Teacher Education and School Research Faculty of Educational Sciences

University of Oslo

2018

Acknowledgments

Although writing a doctoral thesis mainly is a lonesome journey, this dissertation would not have been possible without the many people that have contributed and supported me through what has been five enjoyable, memorable and enriching years.

First of all, I would like to thank my dear supervisor, Guri A. Nortvedt, who inspired and persuaded me to do a PhD in the first place and who has provided me with invaluable support and guidance every step of the way. I have certainly enjoyed our many working sessions and trips to conferences around the world. I would also like to thank my other dear supervisor, Johan Braeken, for his many clever answers, insights and advices throughout these years and for always providing support when I needed it.

Furthermore, I thank the Department of Teacher Education and School Research for giving me the opportunity to do a PhD and my research group, LEA, for academic and financial support. Sincere thanks also go to Rolf Vegar Olsen for your feedback in regards to mid-term evaluation, to Nils Buchholtz for your extensive comments on the final reading, and to Jelena Radišić for reviewing the extended abstract. Your constructive feedback was highly valuable for my work. In the same manner, I am truly grateful for being surrounded by so many inspiring, smart and fun colleagues and friends who make it a pleasure to come to work each day and provide a supportive and stimulating environment that I truly appreciate.

I would like to thank Ross Turner and Mogens Niss for providing me with supportive material and inspiration to adopt the MEG item analysis scheme. In addition, I want to express my gratitude to Ross Turner for providing me with the opportunity to visit ACER and for making my three months in Melbourne such a pleasant experience. Also, I am grateful for the funding provided by the Faculty of Education that made the visit possible.

Additionally, I would like to thank the teachers who participated in my studies, the Norwegian PISA group who provided me with support and data material, and The Norwegian Directorate for Education and Training who provided me with students' exam responses. Without your contribution, this thesis may never have been conducted.

Finally, I want to thank the two most important persons in my life. Thank you, Magali, for your endless encouragement and support, and for being the wonderful girlfriend and mother that you are. Thank you, Tim, for all the joy you bring me. Being a part of your life and coming home to your bright and smiling face means everything to me.

Summary

Notions of mathematical competence promoting an enriched view of mathematical mastery have influenced curriculum reforms around the world. However, there are concerns that the implementation of these notions has stalled in curriculum documents at the system level and that classroom practices still follow a traditional format in which mastering mathematics means possessing factual knowledge and procedural skills. This thesis has investigated the mathematical competency demands of tasks used in Norwegian secondary mathematics and teachers' knowledge about these task demands. In this investigation, six mathematical competencies play a leading role in providing an insight on the topic. These include Communication, Devising strategies, Mathematising, Symbols and formalism, Representation and Reasoning and argument.

A first study aimed to investigate teachers' ability to recognise competency demands of mathematical tasks through the use of an item analysis scheme involving the six aforementioned competencies with four levels of demand for each competency. The results showed a high consistency among the teachers in their ratings of demands, but also that they mainly used the lower levels of the rating scale (Article 1). This indicated that the teachers were able to recognise the mathematical competencies involved in the task solution, but struggled with identifying higher levels of demands. For further scrutinising the teachers' ratings of competency demands, an explanatory item response modelling approach was applied in which the rated demands were combined with students' responses to the tasks (Article 2). The results showed that the teachers' ratings of competency demands could explain around half of the variance in task difficulty, thus providing some empirical evidence supporting the validity of the teachers' ratings. When distinguishing the demands for individual competencies, the results showed that the ratings of some of the competencies (e.g. Symbols and formalism and Reasoning and argument) were related to the difficulty of the items, whereas those of others (e.g. Mathematising and communication) were not. This suggested that the teachers were more successful in recognising the demands for some of the competencies than for others. The results also indicate that for the Norwegian exam, the demands for only two of the competencies were identified and related to task difficulty. This questions the extent to which the exam captures the various cognitive skills and abilities that are represented in mathematical competence. The main methodological contribution of this study is the application of the explanatory item response modelling approach that was able to

empirically identify and separate the demands of individual competencies in mathematical tasks.

Partly inspired by the results from the first study, the second study aimed to investigate teachers' considerations of the demands of mathematical tasks they had used in their teaching practices to challenge high-achieving students. The results of content analysis of the teachers' considerations showed that the teachers mainly emphasised two competencies—Symbols and formalism and Devising strategies—and that these considerations mostly aligned with the identified competency demands of the tasks. Some differences were found when comparing individual teachers in terms of both their considerations of task demands and the competency demands of the tasks they had submitted. These differences were seen to reflect disparities in the teachers' mathematical-task knowledge with regards to their awareness of the mathematical competencies involved in mathematical tasks.

The findings of the two studies suggest that the demands of tasks typically used in Norwegian secondary mathematics classrooms are dominated by a few competencies, among which the Symbols and formalism competency plays the leading role. This indicates that the traditional focus on factual knowledge and procedural skills still pervades classroom activities. Furthermore, the participating teachers seem to mainly recognise and consider the demands for some of the competencies, especially Symbols and formalism, whereas they seem to focus less on the demands for other competencies, such as Mathematising. Even though the number of teachers involved in the two studies is too low to generalise the results, it is believed that these teachers were rather confident in their knowledge of mathematical tasks; thus, it can be considered that the challenges with recognising competency demands are not unique to the participating teachers. Thus, overall, these findings indicate that Norwegian secondary mathematics education lack some of the components needed to ensure that students develop a general mathematical competence as outlined in the national curriculum.

Table of contents

Part I: Extended abstract

Part II: Articles

Article 1: Pettersen, A., & Nortvedt, G. A. (2018). Identifying competency demands in mathematical tasks: Recognising what matters. *International Journal of Science and Mathematics Education, 16*(5), pp. 949–965.

Article 2: Pettersen, A., & Braeken, J. (2017). Mathematical competency demands of assessment items: A search for empirical evidence. *International Journal of Science and Mathematics Education*. doi:10.1007/s10763-017-9870-y

Article 3: Pettersen, A., & Nortvedt, A. G. (under review). Teachers' considerations of mathematical tasks used to challenge high-achieving students. *Scandinavian Journal of Educational Research*

Part I

Extended Abstract

1 Introduction

1.1 Background and rationale

For long, the focus of school mathematics was to develop students' knowledge of mathematical facts and procedural skills. However, in the second half of the 1900s, mathematicians and mathematics educators advocated an increased focus on the process-oriented aspects of mathematics (Apple, 1992; Niss, Bruder, Planas, Turner, & Villa-Ochoa, 2016; Schoenfeld, 1992). This led to an enriched view of mathematical mastery that emphasised the enactment of mathematics, with a focus on problem solving and the ability to apply mathematical knowledge and skills to solve extra-mathematical problems (Clarke, Goos, & Morony, 2007; Niss et al., 2016; Schoenfeld, 2016). In the 1990s, notions and frameworks of mathematical competence, mathematical literacy, and mathematical proficiency emerged and influenced curriculum reforms around the world by portraying a further enriched view of what it means to master mathematics (Kilpatrick, 2014a; Niss & Jablonka, 2014). One example of this influence is found in the Norwegian curriculum reform Kunnskapsløftet ("The Knowledge Promotion Reform") from 2006. The mathematics curriculum that followed the reform states that the subject of mathematics in compulsory education is intended to contribute to the development of the mathematical competence needed by both society and the individual through the development of competencies such as problem solving, modelling, reasoning, communicating, and the ability to use aids and technologies (Norwegian Directorate for Education and Training [Utdanningsdirektoratet], n.d.). The Norwegian curriculum and other mathematics curricula around the world have been influenced by the Danish KOM report (Kilpatrick, 2014a; Niss et al., 2016; Valenta, Nosrati, & Wæge, 2015) that identifies eight mathematical competencies that encapsulate the essence of what it means to master mathematics (Niss & Højgaard, 2011).

Nonetheless, it is not curriculum reforms that develop students' mathematical competencies, but rather the teaching practices and learning situations offered in mathematics classrooms. Tietze (1994) argues that the effectiveness of a curriculum is determined by classroom practices and the decisions, behaviours and attitudes of the teacher and not by the intentions and content of the curriculum. Studies (e.g. Boesen et al., 2014; Charalambous & Philippou, 2010) have shown that teachers seem to assimilate their understanding of curriculum reform messages to fit with current classroom practices rather than change practices, which means that the implementation of curriculum reforms does not necessarily lead to the intended changes in classroom practices. According to Niss and colleagues (Niss et al., 2016), the implementation of mathematical competencies in mathematics education so far mainly concerns the curriculum development and teacher education programmes and there is a lack of research and knowledge about quality teaching to foster and develop mathematical competencies.

The implementation of mathematical competencies beyond curriculum documents would mean adjusting classroom practices with the new view on what it means to master mathematics. Thus, as most teaching and learning in mathematics classrooms is situated around tasks (Bergem, 2016; Boesen et al., 2014; Doyle, 1988), the mathematical tasks in which students engage would need to provide opportunities for the development of a wide range of mathematical competencies (Niss & Højgaard, 2011; Turner, Blum, & Niss, 2015). Several authors have recognised the importance of selecting appropriate tasks as a key to successful mathematics teaching (Anthony & Walshaw, 2009; Chapman, 2013; Hiebert & Wearne, 1993; Tatto et al., 2012), and knowledge about the mathematical thinking and understanding stimulated by tasks is seen as a crucial part of mathematics teachers' knowledge (Ball, Thames, & Phelps, 2008; Baumert & Kunter, 2013; Chapman, 2013; Krauss et al., 2008; Tatto et al., 2012). Studies have shown that the types of tasks that dominate traditional and current classroom practices are rather uniform with a strong focus on procedural skills (Boesen et al., 2014; Hiebert et al., 2003; Kaur, 2010; Lithner, 2004; Palm, Boesen, & Lithner, 2011). According to Niss et al. (2016), to move mathematics education beyond its traditional confines, where mathematical knowledge is reduced to a combination of factual knowledge and procedural skills, it is crucial to support mathematics teachers in understanding and embracing notions of competency and in developing appropriate teaching practices. To accomplish this, more research is needed on the extent to which mathematical tasks demand the use of mathematical competencies as well as teachers' knowledge of such task demands.

1.2 Main objective and research questions

The main objective of this thesis is to contribute to knowledge about the mathematical competency demands of tasks used in Norwegian secondary mathematics and to examine teachers' knowledge about these task demands. Thus, this objective can be seen as twofold: to investigate both tasks and teachers' knowledge about tasks from a competency perspective. As the concept of mathematical-task knowledge is a complex and multi-dimensional construct (Chapman, 2013), this thesis focuses on two aspects of teachers' task knowledge: the ability to recognise demands for specific mathematical competencies and the consideration of competency demands when selecting tasks for teaching practices.

As illustrated in Figure 1, two complementary studies have been conducted which focus on different aspects of the main objective. Study 1 involved an analysis of mathematical competency demands of tasks from two different assessments and an investigation of teachers' ability to recognise competency demands of tasks based on the use of an item analysis scheme. This study is presented in Articles 1 and 2. Study 2, presented in Article 3, involved an analysis of teacher-selected tasks used in secondary mathematics and an investigation of the teachers' considerations of the demands of these tasks.

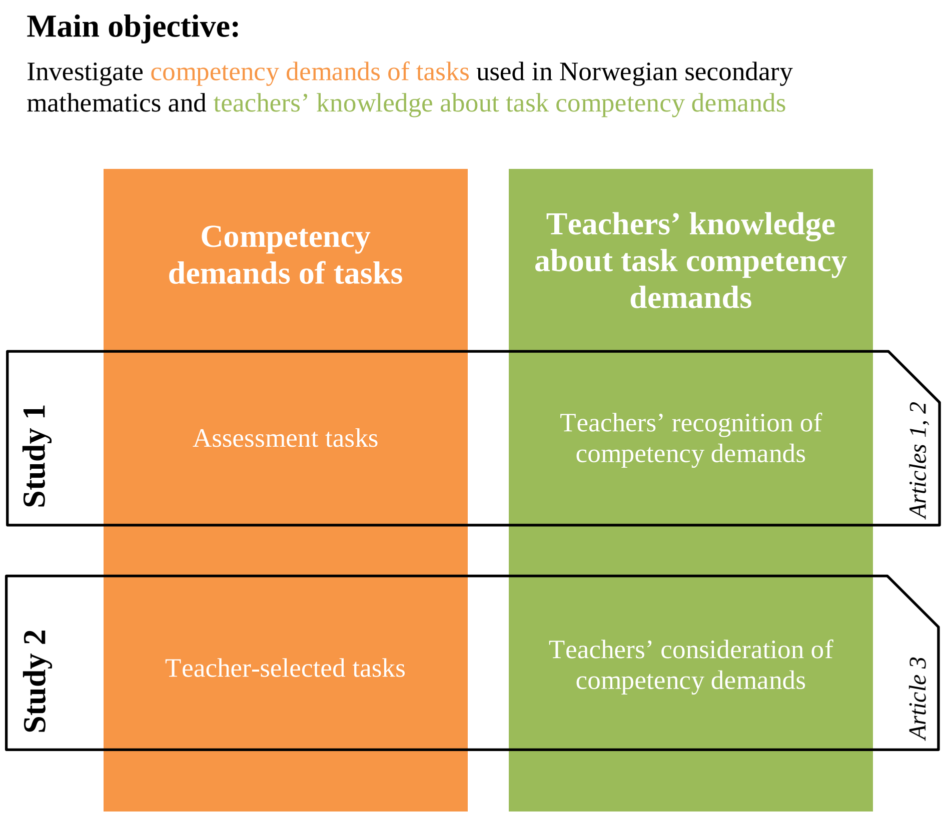

Figure 1. Overview of research project, its two studies' and their focus on different aspects of the main objective of this thesis.

Article 1

Pettersen, A., & Nortvedt, G. A. (2018). Identifying competency demands in mathematical tasks: Recognising what matters. *International Journal of Science and Mathematics Education*, *16*(5), pp. 949–965.

Article 1 presents an investigation of teachers' recognition of mathematical competency demands of tasks. A group of teachers applied an item analysis scheme to individually rate the demands for six mathematical competencies in 141 assessment tasks (from the PISA 2012 survey and a Norwegian national exam). The teachers' ratings' of competency demands provided the quantitative data which was used to answer the research question concerning the degree to which the group of teachers and prospective teachers consistently analyse the competency demands of tasks originally developed to assess students' mathematical competence.

Article 2

Pettersen, A., & Braeken, J. (2017). Mathematical competency demands of assessment items: A search for empirical evidence. *International Journal of Science and Mathematics Education*. doi:10.1007/s10763-017-9870-y

Expanding on the first article, Article 2 presents a psychometric approach for further scrutinising the teachers' rated competency demands of the tasks from the two assessments. By combining the rated demands from Article 1 with students' responses to the same tasks, an explanatory item response modelling approach was applied to address the following research question: To what extent do differences in teacher-rated competency demands in mathematics assessment items align with the differences in empirical item difficulty?

Article 3

Pettersen, A., & Nortvedt, A. G. (under review). Teachers' considerations of mathematical tasks used to challenge high-achieving students. *Scandinavian Journal of Educational Research*

Article 3 presents an investigation of teachers' considerations of the demands of mathematical tasks they have previously used to challenge their high-achieving students. Thus, while the first study focused on teachers' ability to recognise the competency demands of tasks based on the use of a theoretical framework, this study concerned teachers' considerations of the demands of tasks they have selected and used in their teaching practices.

Two research questions were addressed: (1) What characterises teachers' considerations of task demands? (2) How do these considerations align with the competency demands of the tasks according to a competency framework? Seven mathematics teachers from lower- and upper-secondary school submitted a total of 78 tasks they had used to challenge their students along with information about the use and their considerations of the task demands. To answer the two research questions, the teachers' considerations of task demands were analysed both deductively and inductively and compared with an analysis of the competency demands of the submitted tasks.

Mathematics teaching and learning are culturally embedded activities (D'Ambrosio, 1994). My research involves Norwegian mathematics teachers and tasks situated in Norwegian secondary schools based on a Norwegian curriculum. Thus, the results and findings should be seen in a Norwegian context. Still, competency-oriented mathematics curricula have been implemented in many countries (Niss et al., 2016), and therefore, the competency perspective adopted in my research should also be relevant to other educational contexts. Furthermore, mathematics curricula worldwide increasingly seem to be aligned (Cai & Howson, 2013) and some of the mathematical tasks involved in this thesis are adopted from an international assessment study (i.e. Programme for International Student Assessment - PISA, the 2012 Survey). Thus, this research could be relevant to countries with similar school systems and mathematics education frameworks.

1.3 A note on terminology

A large variety of terminologies exist in educational research literature, partly owing to different research traditions and fields and the lack of unanimity among these. This diversity can lead to ambiguity and confusion and create obstacles in the communication and progress of research. Therefore, the following sections aim to define and clarify two key terms in the current thesis.

In this thesis, the term mathematical competence refers to a capability to understand, do, and apply mathematics in a variety of contexts. This draws on the definition provided by Niss and Højgaard (2011), who defined mathematical competence as 'having knowledge of, understanding, doing, using and having an opinion about mathematics and mathematical activity in a variety of contexts where mathematics plays or can play a role' (p. 49). The constructs and notions of mathematical competence are sometimes referred to as competency frameworks (Kilpatrick, 2014a; Niss & Højgaard, 2011), in which mathematical competence is described through a set of sub-constructs or strands (Niss et al., 2016). These sub-constructs or strands are referred to as a mathematical competency (or competencies). Examples of such competencies are communication, mathematical reasoning, modelling, and problem solving (for a more exhaustive list of competencies, see e.g. Kilpatrick, 2014a; Niss et al., 2016).

The concept of mathematical tasks is complex and multifaceted (see Chapter 2.2). An oft-cited definition of a mathematical task is presented by Stein, Grover, and Henningsen (1996) as 'a classroom activity, the purpose of which is to focus students' attention on a particular mathematical concept, idea or skill' (p. 460). This can be seen as a broad definition where a task could involve one or several mathematical problems to be solved, exercises to be performed, questions to be answered, or other mathematical activities (such as playing a mathematical game or engaging in a classroom discussion). In this thesis, the term 'mathematical task' is restricted to activities where students are expected to provide a solution or an answer. This involves both what can be regarded as mathematical *exercises* (i.e. routine-based tasks in which the answer is obtained through known strategies and algorithms (Borasi, 1986)) and mathematical *problems*[1] (i.e. non-routine tasks that are intellectually challenging and for which no methods, procedures, or algorithms for solving the problem are readily accessible to the problem solver (Blum & Niss, 1991)). Tasks used in mathematics teaching and learning are drawn from multiple sources (e.g. textbooks, the teacher, and tests) and can be used for different purposes (e.g. instruction and assessment). In the field of psychometrics, the term 'item' is used to refer to the questions, problems, or tasks involved in tests or assessments intended to measure certain abilities or attributes. Thus, the same task can be seen as an exercise, problem, instructional task, assessment task, or item depending on its use. The current thesis does not distinguish between tasks from different sources and used for different purposes; rather, it treats these tasks the same and considers all to be relevant with regards to investigating competency demands. Furthermore, the terms 'task' and 'item' are used interchangeably to refer to mathematical problems and exercises from assessments.

Curriculum is another concept that is understood and defined in numerous ways. In this thesis, the term 'curriculum' is used to refer to the formal written documents that involve the goals and expectations for the learning of mathematics at a system level. This is typically referred to as the intended curriculum (Cai & Howson, 2013), as contrasted with the

[1] The term 'mathematical task' is preferred over 'mathematical problem' as the latter has multiple and contradictory meanings in both research literature and curriculum documents (Schoenfeld, 1992). Furthermore, my thesis includes both typical routine-based tasks which are often not perceived as mathematical problems.

implemented curriculum (i.e. classroom practices) and the attained curriculum (i.e. what students learn). When referring to the curriculum at another level than the intended, this is explicitly stated in the text.

1.4 Structure and content of thesis

This thesis consists of two main parts. Part I contains the extended abstract, and Part II encompasses the three articles. As the extended abstract elaborates on the relationship between the articles and places their rationales and findings in a broader educational context, I would recommend reading the extended abstract before reading the articles. Still, each article stands on its own as a unique contribution.

Part I comprises four chapters. The current introductory chapter (Chapter 1) has stated the rationale and main objective of my thesis, and clarified some key terminology. Chapter 2 presents the theoretical foundation on which the thesis is based. It provides a brief historical outline that illustrates the complexity of the research field and the background for the growth of what can be considered competency-oriented mathematics education. Chapter 2 also provides an overview of research on features and demands of mathematical tasks, how such task characteristics relate to the potential learning opportunities provided by tasks, and teachers' use and knowledge of mathematical tasks. These overviews are used to position my research in the context of existing literature as well as to identify knowledge gaps. Finally, this chapter discusses the importance and challenges of developing assessments and assessment tasks that provide valid measures of mathematical competence. Chapter 3 presents the design of the research project and its two studies and the philosophical assumptions on which it is based. The methodological considerations, issues, and limitations are discussed and the research validity is addressed. This chapter also describes ethical concerns involved in my research. Chapter 4 presents a summary of the three articles and discusses the main contributions of the thesis in light of the main objective. This chapter aims to discuss the results and findings across the two studies and within the Norwegian educational context.

2 Theoretical background and framing

Mathematics education as a scientific field is situated at the nexus of education and mathematics and is grounded in a variety of fields such as psychology and philosophy (Ernest, 1991; Kilpatrick, 2014b; Schoenfeld, 2008; Sriraman & English, 2010). During the 1900s, research in mathematics education was shaped and steered by several shifting and coexisting epistemological and philosophical views, theories of learning, and research traditions (Blum, Artigue, Mariotti, Sträßer, & Van den Heuvel-Panhuizen, 2017; Ernest, 1991, 2010; Kilpatrick, 2014b; Schoenfeld, 2016; Sriraman & English, 2010) that built on a range of 'isms' including connectionism, behaviourism, and cognitivism (for a more thorough historical briefing, see e.g. Blum et al., 2017; Kilpatrick, 2014b; Schoenfeld, 2008, 2016). This pluralism led to different research traditions and views of mathematics teaching and learning and, according to Sriraman and English (2010), to a 'utilitarian mix-and-match culture' among mathematics education researchers. This is evident from the research and studies that are presented throughout this chapter which draw on different research traditions and methodologies.

2.1 Nature of mathematical knowledge

Different philosophies of mathematics lead to different educational practices, as mathematics curriculum and teaching practices are shaped by the philosophy and views on mathematical knowledge on which they are based (Ernest, 1991). Mathematics was long seen as the unique realm of certain knowledge and infallible objective truth established by logical deduction from axioms, where the axioms are considered as basic self-evident truths which do not need further justification (Ernest, 1991). In line with such an absolutist view is the belief that acquisition of mathematical knowledge means to remember and correctly apply mathematical rules and facts provided by authorities (such as teachers and textbooks), which again shapes teaching and learning practices in mathematics classrooms (Ernest, 1989)[2].

In the 1960s and 1970s, a growing number of mathematicians and philosophers questioned and criticised the absolutist view of mathematical knowledge (Ernest, 1991). A

[2] There is no one-to-one relation between epistemological beliefs about mathematical knowledge and views on instruction and school mathematics (for instance, it is possible to see mathematical knowledge as objective and certain and at the same time believe that students' should learn mathematics through discussion and engagement in personal real-life problems and activities). Still, several studies indicate that there is a strong relationship between epistemological beliefs and views on school mathematics and instruction (see Philipp, 2007).

highly influential critic, Imre Lakatos, argued that rather than being developed as a steady accumulation of infallible truths, mathematical knowledge is discovered in a process of human activity involving dialogues, exposure to criticism, reformulations, and, possibly, refutations (Lakatos, 1976). This criticism led to the evolution of a new tradition of philosophy of mathematics that sees mathematics as fallible and corrigible and as being developed through a *process* of human activity, thus dismissing the absolutist view of mathematical knowledge as a *product* of certain unchallengeable truths (Ernest, 1994; Sriraman & English, 2010). A similar shift can be seen in mathematics education. In the 1940s, mathematicians and mathematics were already stressing that mastering mathematics went beyond the traditional focus on knowledge of mathematical facts and rehearsing of procedures and involved aspects such as the enactment of and doing mathematics (Niss et al., 2016; Schoenfeld, 1987, 1992). One prominent influence of the process-oriented view on mathematics is the renewed attention to teaching and learning of problem solving that evolved within the field of mathematics education (Lesh & Zawojewski, 2007; Schoenfeld, 1992). While problem solving traditionally concerned whether students were able to solve problems, the process-oriented perspective shifted the focus to the cognitive activities involved in the problem solving processes, such as the strategies and metacognitive behaviour that were conducted (Schoenfeld, 1992).

2.1.1 Growth of competency frameworks in mathematics education

In the 1990s, notions of mathematical competence and similar concepts such as mathematical literacy, numeracy, and mathematical proficiency gained an increased foothold in mathematics education (Kilpatrick, 2014a; Niss et al., 2016). The term 'mathematical competence' in itself has long been used as a generic term for a persons' ability to handle or apply mathematics. For instance, Hiebert and Lefevre (1986) argued that '[b]eing competent in mathematics involves knowing concepts, knowing symbols and procedures, and knowing how they are related' (p. 16). However, the notions that emerged in the 1990s provided an enriched view of mathematical mastery beyond that of conceptual and procedural knowledge and promoted a more nuanced image of school mathematics that involved a variety of mathematical competencies (Kilpatrick, 2014a). There is no common definition or understanding of 'competence' in general (Blömeke, Gustafsson, & Shavelson, 2015; Pikkarainen, 2014; Westera, 2001) or of mathematical competence in particular (Niss et al., 2016), and several competency frameworks of mathematical competence have been developed and coexist in the field of mathematics education (Kilpatrick, 2014a). According to Boesen et al. (2014), constructs of mathematical competence can be seen to be inspired by ontological and epistemological development from the process-oriented focus that evolved in the 1960s and 70s as well as from ideas from social constructivism related to teaching and learning mathematics. With regard to the epistemological foundation, Niss et al. (2016) claim that theoretical constructs of mathematical competence have grown out of the experiences and minds of the proponents through observations, reflections and discussions and through systematic empirical and experimental work. Thus, notions of mathematical competence seem to have a pragmatic nature (Cherryholmes, 1992; Hildebrand, 2013) in that they are based on a range of research traditions and philosophies in which a main concern seems to relate to the consequences and usefulness of the proposed notions, that is, whether the proposed competence frameworks promote the 'right' kind of mathematics learning and teaching (Niss et al., 2016).

This diverse and pragmatic nature is also reflected in mathematics curricula in countries around the world where a wide range of different notions and conceptualisations of mathematical competence have been adopted (Niss et al., 2016). One conceptualisation of mathematical competence that has been highly influential on curriculum reforms in several European countries such as Norway (Valenta et al., 2015), Sweden (Boesen et al., 2014), and Denmark and Germany (Niss et al., 2016) originated from the Danish KOM project (Niss & Højgaard, 2011; Niss & Jensen, 2002)[3]. From this project, a framework was derived (hereafter referred to as the KOM framework) that identified and characterised eight distinct but overlapping mathematical competencies, namely, thinking mathematically, problem tackling competency, modelling competency, reasoning competency, representing competency, symbol and formalism competency, communicating competency, and aids and tools competency. The idea behind the proposed conceptualisation was that these competencies should function as a means for identifying and characterising mathematical mastery and that mathematics teaching and learning should focus on the development of these eight competencies across mathematical content domains and educational levels (Kilpatrick, 2014a).

Still, while notions of mathematical competence have gained a foothold in mathematics curricula documents worldwide, the implementation of competencies in

[3] KOM is the abbreviation for 'Competencies and Mathematical Learning' in Danish. This thesis mainly refers to the English version of the KOM report (see Niss & Højgaard, 2011) which was originally presented in Niss and Jensen (2002).

classroom practices is less pronounced (Niss et al., 2016). For instance, Boesen et al. (2014) found that even 15 years after the implementation of competency-oriented mathematics reform in Sweden, teaching practices were not in line with the reform message and were still dominated by practicing procedures. One reason for this may be that teachers seem to find it challenging to come to grips with and implement notions of mathematical competencies, which is a crucial factor for adopting teaching practices in line with a competency-oriented view of the learning of mathematics (Niss et al., 2016). The current thesis focuses on two key components of teaching and learning of mathematics that are vital to ensure that mathematical competencies are not implemented solely in the intended curriculum but also in classroom practices: the mathematical tasks in which students engage and teachers' knowledge of these tasks.

2.2 Mathematical tasks

Mathematical tasks are used to engage students in particular mathematical concepts or ideas and are often formulated or shaped as a problem, exercise, or question for the students to solve or answer (see Chapter 1.3). Mathematical tasks play a key role in mathematics teaching and learning as the tasks in which students engage are seen to determine their opportunities to develop mathematical understanding and skills and to engage in mathematical thinking (Ainley, Bills, & Wilson, 2005; Hiebert & Wearne, 1993; Richardson, Carter, & Berenson, 2010; Stein & Smith, 1998; Sullivan, Clarke, & Clarke, 2013). In their synthesis of what research can tell about quality mathematics teaching, Anthony and Walshaw (2007, p. 94) argue that 'it is through tasks, more than any other way, that opportunities to learn are made available to students'. Sierpinska (2004) argues that mathematical tasks are the fabric of both mathematics teaching and learning and of research in mathematics education. The crucial role of tasks in mathematics education is reflected in the research literature where tasks are found to be used as both (1) a means of research to investigate teaching practices or students' proficiency and as (2) an object of research to investigate the features, characteristics, or demands of tasks.

Research involving mathematical tasks is not grounded in a single theoretical perspective or philosophical stance. Rather, it draws on the different research traditions, methods, and theories of learning that coexist within the field of mathematics education (as discussed in previous sections). From a sociocultural perspective, a mathematical task is seen as tightly related to the social context in which the task is embedded and, as such, research based on this perspective is often concerned with the nature of and interaction between teachers and students directly and indirectly involved in the task activity (Shimizu, Kaur, Huang, & Clarke, 2010). An example of such a study was conducted by Henningsen and Stein (1997) who investigated how classroom-based factors such as classroom norms and teachers' and students' dispositions shape student engagement with tasks. Another research practice is to focus on the properties and characteristics of mathematical tasks as stated in curriculum and instructional material (e.g. textbooks) more or less independently of the classroom situation in which they are to be implemented. An example of such an approach is found in one of the publications from the COACTIV study (Professional Competence of Teachers, Cognitively Activating Instruction, and Development of Students' Mathematical Literacy) that analysed different characteristics of tasks as they were presented in instructional documents and used these characteristics as indicators of the potential for cognitive activation offered by the tasks (M. Neubrand, Jordan, Krauss, Blum, & Löwen, 2013). The different perspectives and approaches provide different possibilities and should be seen as complementary rather than competitive. As illustrated by the two aforementioned studies, the former perspective is often associated with studies investigating how mathematical tasks are implemented in teaching and learning situations whereas the latter perspective is concerned with the potential of tasks in terms of the learning opportunities they are seen to promote. According to J. Neubrand (2006), in analyses of mathematical tasks, it is vital to distinguish between problems as they are posed independent of their implementation and problems as they take place and are enacted in practice. This distinction formed a basis for the investigation of teaching practices conducted in the TIMSS 1999 Video Study, in which the analysis distinguished between potential demands of tasks as they were stated and how mathematical tasks were actually worked on in the classroom (Hiebert et al., 2003).

Numerous types of tasks with different characteristics and properties are described in the literature, such as routine- and non-routine problems, word-problems, exercises, procedural tasks, modelling tasks, representational tasks, contextual tasks, open-ended tasks, rich tasks, real-world tasks, and investigative tasks (see e.g. Borasi, 1986; de Lange, 1995; Haapasalo & Kadijevich, 2000; Haladyna & Rodriguez, 2013; Mayer & Hegarty, 1996; Sullivan et al., 2013; Yeo, 2017). Tasks can be used for instructional purposes (intended to promote learning of mathematics) and assessment purposes (intended to generate information about student learning) (Shimizu et al., 2010) and can be implemented in school or as homework. Tasks can be targeted for different grade levels, may involve different mathematical content (e.g. geometry or algebra), activities (classroom discourse, games, or individual seat work) and mental processes, and require different types of responses (e.g. answering a question or solving a problem with or without explanations and justifications, orally or in writing). This complexity makes it challenging to provide an exhaustive overview of research related to mathematical tasks. Instead, the following sections concentrate on studies concerned with the analysis of task characteristics and how these relate to the potential learning opportunities provided by the tasks.

2.2.1 Task analysis

One distinction that can be made in the analysis of mathematical tasks is between task features and task demands. Task features are related to the presentation and formulation of the task, the statements it involves, and its mathematical and contextual features. This involves the mathematical content and topics related to the task, possible visual features (e.g. illustrations and figures), textual quantity and quality, and the context or situation in which the problem is embedded. Some features (such as mathematical content, embedded context, and visual features) are regarded as surface characteristics (Arbaugh & Brown, 2005; Turner, Dossey, Blum, & Niss, 2013) and are more or less directly observable when examining the task, whereas other features are more latent, such as the openness of tasks in terms of goal and method(s) of solution (Yeo, 2017). Task demands refer to requirements assumed to be needed to be able to solve or complete the task. This typically entails the mathematical knowledge, operations, and other cognitive aspects (such as mathematical and mental processes) involved in the solution process.

Several frameworks have been developed to analyse mathematical tasks in terms of both their features and demands. In an early work, Goldin and McClintock (1979) presented a task variables framework based on a review of research on mathematical problems and problem solving. This framework consisted of four main categories: problem syntax, mathematical content and non-mathematical context, problem structure, and heuristic processes. These categories were seen to identify significant aspects of tasks in terms of task complexity and difficulty, and the framework was seen as a useful tool both for measuring and stimulating the learning of mathematical problem solving (Goldin & McClintock, 1979). Li (2000) argued that it is vital to analyse different features and requirements of textbook problems, and not only textbook content, which has been the traditional focus of interest, as textbook problems play an important role in students' learning of mathematics. Li (2000) developed a framework involving three dimensions—mathematical features, contextual features, and performance requirements—and applied this framework to analyse and compare textbook problems. Yeo (2017) developed a framework to analyse the openness of mathematical tasks based on a set of task variables including the goal of the task, method for solution, and task complexity. According to Yeo (2017), this framework could be useful for teachers to design and select tasks that engage their students in appropriate mathematical thinking processes and for researchers when investigating the relationship between task openness and learning in mathematics. The Task Analysis Guide developed by Stein, Smith, Henningsen, and Silver (2000) describes the characteristics of tasks associated with four levels of cognitive demand: doing mathematics, procedures with connection, procedures without connections, and memorization. Stein et al. (2000) argue that the framework can support teachers in differentiating between different levels of demand and raise their awareness of the demand of tasks and how they relate to goals for student learning. The Task Analysis Guide has been adopted and used in several studies to investigate the cognitive demands of tasks both as they are presented in educational material and set up and implemented in the classroom (e.g. Boston & Smith, 2009, 2011; Brändström, 2005; Charalambous, 2008).

Despite the large number of studies that have analysed the features and demands of mathematical tasks, few empirical studies have been conducted involving an entire system of mathematical competencies (Niss et al., 2016). According to Boesen, Lithner, and Palm (2018), one reason for this might be that mathematical competency frameworks have been mainly developed for curriculum development and not for analysing empirical data. One exception is the research conducted by Lithner et al. (2010), who developed a research framework for mathematical competencies to serve as a basis for the analysis of tasks as well as other empirical data. Based on their analysis of Swedish national tests, Boesen et al. (2018) argue that this framework is useful for examining the extent to which tasks are evenly distributed across competencies and whether tests actually 'capture the whole spectrum of what it means to be mathematically competent' (p. 121). Another exception is the research conducted by the PISA Mathematics Expert Group (hereafter referred to as the MEG) which involved the development of an item analysis scheme to be used to analyse mathematical problems given to 15-year-old students with regard to the extent to which solving these problems required the activation of mathematical competencies (Turner et al., 2015). This scheme was used to analyse 48 mathematical problems used in both the PISA 2003 and PISA 2006 surveys by Turner et al. (2013). Turner et al. (2015) subsequently concluded that this scheme could be used effectively by experts to identify the competency demands of these problems.

2.2.2 MEG item analysis scheme

The MEG item analysis scheme (see Appendix 1) comprises six mathematical competencies: Communication, Devising strategies, Mathematising, Representation, Symbols and formalism, and Reasoning and argument. These were originally derived from the eight competencies included in the first PISA Mathematics Frameworks (Turner et al., 2015), which again is based on (and evolved in parallel and intertwined with) the KOM framework (Niss, 2015). The reduction from eight to six competencies followed from merging mathematical reasoning and mathematical thinking into the Reasoning and argument competency and omitting the mathematical aids and tools competency (Turner et al., 2015). As the notion of mathematical competence in the Norwegian curriculum is also based on the KOM framework (Valenta et al., 2015), the MEG scheme and its six competencies are regarded as highly relevant to mathematics education in Norway and a suitable tool for analysing mathematical tasks in Norwegian secondary education. However, in the Norwegian curriculum the ability to use aids and technologies is described as an important aspect of mathematics. The lack of an aids and tool competency in the MEG framework means that the analysis of tasks based on the MEG scheme does not capture this aspect of the Norwegian mathematics curriculum.

The MEG competencies are in line the definition proposed by Klieme and Leutner (as cited in Klieme, Hartig, & Rauch, 2008) in which competencies are defined as 'context-specific *cognitive* dispositions that are acquired by learning and needed to successfully cope with certain situations or tasks in specific domains' (p. 9). Several frameworks breakdown mathematical competency into separate dimensions such as knowledge (e.g. factual and procedural knowledge), cognitive processes (e.g. remember, understand, and apply), and mathematical content (e.g. algebra, geometry, and measurement) (Kilpatrick, 2014a). In the MEG framework, a mix of skills, knowledge, and mental processes are used to describe each of the individual competencies. Compared to other competency frameworks, such as the US National Research Council's five strands of mathematical proficiency (Kilpatrick, 2014a), the MEG framework has no clear emphasis on the importance of a deep understanding of mathematical concepts. This can be seen as a limitation, given the strong position of conceptual understanding in mathematics education. Still, this type of understanding is implicitly included in the handling of mathematical concepts involved in, for instance, the Representation competency and the Symbols and formalism competency. Another issue is the two-sided character of the mathematical competencies described in the KOM report, involving both an investigate side (i.e. the ability to understand, reflect, and analyse) and a productive side (i.e. the ability to carry out processes). Although the two sides are not explicitly stated they are implicitly involved in several of the competencies. For instance, the Communication competency involves both reading and interpreting statements and making sense of information as well as presenting and explaining mathematical work and reasoning. Similarly, Mathematising involves both interpreting outcomes and validating given mathematical models as well as constructing models based on extra-mathematical situations.

The operational definitions in the MEG scheme involve four levels of cognitive demands (0–3) for each of the six competencies, where level 0 implies no or very minimal demand for the activation of this competency and level 3 implies a demand at an advanced or complex level (Turner et al., 2015). To support reliable and consistent ratings of competency demands, the operational definitions have been developed to make the distinctions between the competencies as distinct as possible (Turner et al., 2015). However, mathematical competencies are seen to have an overlapping and intertwined nature (Niss et al., 2016), which is also the case for the KOM competencies (Niss & Højgaard, 2011). The study conducted by Turner et al. (2013) showed rather high correlations between the rated demands for some of the competencies, which could indicate challenges with obtaining such clear distinctions and thus with separating the demands for the different competencies.

2.2.3 Task analysis: Empirical research

In many studies, the analysis of task features and/or demands based on the abovementioned or similar frameworks has been used as a means to investigate teaching and learning practices of mathematics. This research is based on the assumption that certain task characteristics could be influential or beneficial with regard to the quality of teaching more or less independently of the classroom context in which the task is implemented and of the students engaging in the task. Results from these studies have contributed to an increased understanding of the learning opportunities students are provided in the mathematics classroom and of the types and use of mathematical tasks that are associated with quality mathematics teaching.

As a part of the COACTIV study, Baumert et al. (2010) analysed assessment tasks developed and used by Grade 10 mathematics teachers in Germany in terms of three dimensions of cognitive demands: type of task, level of mathematical argumentation required, and translation processes within mathematics. These assessment tasks were found to reflect the task structure found in teacher instruction, and the results showed that the use of more cognitively demanding tasks had a substantial positive effect on students learning gains (Baumert et al., 2010). Furthermore, the results also showed that tasks provided by German Grade 10 teachers had a low overall level of cognitive challenge (Baumert et al., 2010). Stein and Lane (1996) analysed instructional tasks as they were set up and implemented in classrooms and found that the use of tasks that involved high levels of cognitive demand (i.e. doing mathematics or procedures with connection) led to greater student gains on a performance assessment involving high levels of mathematical thinking and reasoning. Similarly, in the TIMSS 1999 video study, the cognitive demand of mathematical problems was analysed two times: first to characterise the problem, and second, to describe its implementation (Stigler & Hiebert, 2004). The results revealed that in the highest-achieving countries, mathematical problems characterized as high-demanding were, to a large extent, implemented as high-demanding. This was in contrast to what happened in US classrooms, where all high-demanding tasks changed into routine exercises or other cognitively low-demanding activities when they were implemented (Stigler & Hiebert, 2004).

Several studies have analysed the demands of mathematical tasks in textbooks to examine students' potential learning opportunities. For instance, in her doctoral thesis, Brändström (2005) analysed the cognitive demand of differentiated tasks (i.e. tasks located in separate strands according to ability level) in Swedish textbooks and found that the tasks at the low strands mainly involved a low cognitive demand in terms of memorisation and remembering. According to Brändström (2005), this breaks with the Swedish curriculum which states that all students should engage in higher-order thinking such as reasoning and reflecting with the help of mathematics. Similarly, Jones and Tarr (2007) analysed US middle-grade mathematics textbooks and found that for six of the eight textbook series, a vast majority of the tasks (>83%) required low cognitive demand (i.e. procedures without connections and memorization). Furthermore, they argued that the two other textbook series that involve a higher proportion of cognitively demanding tasks have a higher potential for developing a deeper understanding of mathematical content (Jones & Tarr, 2007). Several other studies have also analysed the features and demands of textbook tasks to investigate and compare students' potential for learning mathematics both within and between countries (e.g. Baker et al., 2010; Charalambous, Delaney, Hsu, & Mesa, 2010; Li, 2000; J. Neubrand, 2006).

The research framework for mathematical competencies developed by Lithner et al. (2010) was used to analyse the mathematical activities in Swedish mathematics classrooms (Boesen et al., 2014) and Swedish national mathematics tests (Boesen et al., 2018). Based on the observation of 197 lessons, they concluded that despite the implementation of a mathematical competence reform, carrying out procedures still dominated classroom practices (Boesen et al., 2014). The analysis of the national tests showed that these seem to capture, to a fairly high extent, all of the mathematical competencies but that the complex nature of the competencies is not fully captured as aspects such as the ability to evaluate and reflect on mathematics and to draw conclusions are not involved in the tests (Boesen et al., 2018).

The results from the aforementioned studies yield a somewhat coherent picture illustrating the importance of cognitively demanding tasks in teaching and learning mathematics and simultaneously identifying that the tasks provided in mathematics classrooms mainly engage students in activities that involve low cognitive demand.

2.3 Assessing mathematical competencies

Mathematical tasks also play an important role when it comes to assessing students' learning in mathematics. Suurtamm et al. (2016) distinguish between large-scale assessments (such as PISA, TIMSS, and national assessments) and classroom assessments (typically, teacher-made or teacher-selected tests) where the two types of assessments are grounded in different traditions with different epistemological perspectives and theories of learning. Classroom assessments are based on cognitive, constructivist, and/or sociocultural views of learning in which assessment is seen as a social practice mainly intended to support students' learning (Suurtamm et al., 2016; Wiliam, 2007). Large-scale assessments stem from a psychometric tradition (Glaser & Silver, 1994) associated with a postpositivist worldview (Creswell & Plano Clark, 2011) in which a main intention is to gain reliable measures of students' learning outcomes. In some cases, large-scale assessments also function as a means to implement reform messages as teaching practices are assumed to be adapted to fit with what is measured in the tests (Boesen et al., 2018; de Lange, 2007).

For both large-scale and classroom assessments, it is important for the assessment to reflect the type of mathematics that is valued and described in the curriculum (de Lange, 2007; Suurtamm et al., 2016; Wiliam, 2007). Thus, the shift to a competency-oriented mathematics education also requires a shift in assessment practices as the range of knowledge, skills, and cognitive processes involved in the notions of mathematical competence requires different types of tasks than what is involved in traditional knowledge tests. However, because of its complex nature, the assessment of students' mathematical competencies is regarded as highly challenging (Koeppen, Hartig, Klieme, & Leutner, 2008; Niss et al., 2016), and there are concerns over whether current assessment practices are able to measure the complex abilities and higher-order thinking involved in such competencies (e.g. Koeppen et al., 2008; Lane, 2004; Niss, 2007). Developing and selecting tasks that demand the use of these competencies is challenging for both teachers and test developers, and it is seen as one of the crucial aspects for the implementation of mathematical competencies in mathematics teaching and learning (Niss et al., 2016).

2.4 Mathematical-task knowledge for teaching

The selection of appropriate tasks is regarded as a key to successful mathematics teaching (Anthony & Walshaw, 2007, 2009; Ball et al., 2008; Doyle, 1988; Hiebert & Wearne, 1993), whether used for instructional or assessment purposes. Appropriate tasks are seen as tasks that challenge students at an appropriate level, extend current understanding and knowledge, and provide opportunities for students to struggle with important mathematical ideas and engage in high-level thinking (Anthony & Walshaw, 2009; Shimizu et al., 2010). For teachers to select and develop appropriate tasks when planning lessons, they need knowledge of mathematical tasks and the potential learning opportunities they provide. According to Ball (2000, p. xii), '[a]cquiring the ability to think with precision about mathematical tasks and their use in class can equip teachers with more developed skills in the ways they select, modify, and enact mathematical tasks with their students'. The importance of knowledge of the learning potential of mathematical tasks is also emphasised in many frameworks for mathematics teachers' professional knowledge (see, e.g., Ball et al., 2008; Baumert & Kunter, 2013; Krauss et al., 2008). Chapman (2013) uses the term 'mathematical-task knowledge for teaching' to refer to the knowledge teachers need to select and develop appropriate tasks and to optimize the learning potential of tasks. This task knowledge has many facets, including knowledge of cognitive demands of tasks and the learning and understanding they can promote as well as the ability to identify and create tasks that provide opportunities to develop meaningful and deep understanding in accordance with the learning needs of the students. According to Chapman (2013), mathematical-task knowledge is a key factor in teachers' treatment of tasks. This was, to some extent, confirmed in a study by Baumert et al. (2010), who found that the extent to which teachers provided cognitively challenging tasks was largely determined by their pedagogical content knowledge and their knowledge of mathematical tasks. The enriched view of mathematical mastery promoted by notions of mathematical competence can be seen to further add to the mathematical-task knowledge for teaching. To stimulate the development of mathematical competence, teachers need to select and develop tasks that provide opportunities for developing a range of competencies (Turner et al., 2015) or, in the words of Niss and Højgaard (2011, p. 31), orchestrate activities 'with the explicit aim of developing the mathematical competencies of the individual'. An essential factor for this is that teachers must grasp the notions of mathematical competencies and be empowered to develop teaching approaches that implement these competencies (Niss et al., 2016).

Several studies have shown that teaching practices are dominated by the selection and use of tasks that involve lower levels of cognitive demand (Baumert et al., 2010; Boston & Smith, 2009; Silver, Mesa, Morris, Star, & Benken, 2009). Furthermore, studies of teachers' (e.g. Boston, 2013) and pre-service teachers' (e.g. Osana, Lacroix, Tucker, & Desrosiers, 2006) evaluation of mathematics problems have shown that teachers might struggle with recognising and understanding the cognitive demands of problems involving high levels of complexity. Some studies also indicate that when analysing tasks, teachers tend to focus on surface characteristics (e.g. Arbaugh & Brown, 2005; Osana et al., 2006; Smith & Stein, 1998). Furthermore, Sproesser, Vogel, Dörfler, and Eichler (2018) found a rather large discrepancy between teachers' estimations of task solution rates and empirical solution rates of tasks for students of age 12 and 16 years and argued that the ability to accurately judge the difficulty of tasks is important to support students' learning. These findings suggest that mathematics teachers struggle with recognising or tend to not consider the cognitive challenges involved in the tasks they use in their teaching practices; this might indicate a lack of sufficient mathematical-task knowledge. Furthermore, this could result in students not being provided with appropriate tasks, thus jeopardising their opportunities to extend their mathematical understanding and knowledge and to become mathematically competent.

It is important to note that the recognition and selection of appropriate tasks in itself does not ensure successful mathematics teaching. The characteristics and demands of tasks can change when moving from instructional material to classroom implementation (Stein et al., 1996), and studies have shown that maintaining the cognitive demand of tasks can be challenging for teachers (Brodie, Jina, & Modau, 2009; Stein et al., 2000; Stigler & Hiebert, 2004). The importance of task implementation is also reflected in Chapman's (2013) concept of mathematical-task knowledge, which involves knowledge of how to orchestrate and organise students work and support their process of thinking without reducing or eliminating the cognitive challenge. Consequently, the appropriateness of a task cannot be determined solely by analysing its features and demands, as this also depends on the students who are to engage in the task (such as their abilities, interests, and motivation) and the social context in which the task is implemented. Thus, successful mathematics teaching requires knowledge of tasks, students, and the student-task interaction as well as the ability to implement and adjust tasks in accordance with the sociocultural context of the classroom. Still, all studies are bound by particular restrictions and limitations. This research project follows in the footsteps of several of the aforementioned studies in focusing on the potential of mathematical tasks and their features and demands as they are presented in instructional material, as well as on teachers' task knowledge in terms of their recognition and considerations of these.

3 Methodology and research design

Quantitative and qualitative research differ in that the former is characterised by a focus on deduction, hypothesis testing, statistical analysis, and a search for objective knowledge whereas the latter is associated with induction, exploration, and theory generation where subjective experiences and interpretations form a basis for data and analyses (Johnson & Onwuegbuzie, 2004). In practice, this dichotomy is not very fruitful as the problems addressed in social sciences are often complex and require the use of different methodological approaches (Brewer & Hunter, 2006; Creswell & Plano Clark, 2011). This pluralism can also be seen in research involving mathematical tasks, where different research traditions and epistemological positions coexist (see Chapter 2.2). Sierpinska (2004) argues that mathematical tasks can be seen to function as research tools on the same level as other methodological tools. This is reflected in the methodological design of my research project, in which mathematical tasks and task analysis pervade the research process, functioning both as a means and mode of data collection, influencing the methods used, and forming the theoretical underpinnings. Furthermore, the complex nature of mathematical tasks and the analysis of competency demands encouraged the inclusion of both quantitative and qualitative components.

3.1 General overview

3.1.1 Philosophical position

The philosophical worldview or paradigm that underlies research practice involves a range of assumptions about how researchers gain knowledge of the world, the research process, the nature of reality, and the researcher’s role in the research process (Creswell & Plano Clark, 2011; Lincoln, Lynham, & Guba, 2005). As this worldview influences how research is conducted, it is important that researchers be aware of and identify the assumptions and ideas underlying their research (Creswell & Plano Clark, 2011). The current research project can be seen as based on a postpositivist philosophy (Creswell & Plano Clark, 2011; Phillips & Burbules, 2000). Unlike positivism, in which sensory experience is regarded as the only source of knowledge and the possibility of gaining knowledge of unobserved entities is rejected, postpositivism sees knowledge as conjectural and that it can be supported by all available evidence including empirical evidence, arguments, controlled experiments, and interviews (Phillips & Burbules, 2000). This means that although it is usually associated with quantitative methods and statistical analyses (Creswell & Plano Clark, 2011; Lincoln et al., 2005), postpositivism is not restricted to certain methodologies or methods and does not reject qualitative approaches or interpretative methodologies in the search for knowledge (Phillips & Burbules, 2000). Thus, it is in line with the use of a quantitatively driven research approach applying a variety of methods. Postpositivism aims to discover or approximate a singular reality (although this might not be achievable) through objectivity and falsification of hypotheses (Lincoln et al., 2005; Phillips & Burbules, 2000). My project reflects the idea of singular reality and objectivity through the use of an item analysis scheme for task analysis, where the scheme is seen to ensure that information about task competency demands is collected objectively without being influenced by the subject conducting the analysis. This is also reflected in the use of statistical analyses to investigate teachers' recognition and considerations of task competency demands. For instance, the estimation of the consistency of teachers' ratings of demands to evaluate the accuracy of their ratings (Article 1) is based on the assumption that the tasks have a certain set of competency demands and that if teachers have independently arrived at the same set of ratings, these are likely to reflect an objective, singular reality. The focus on the potential demands of mathematical tasks as presented in educational and instructional materials, while ignoring several important aspects of the use of tasks in mathematics teaching, draws on ideas of reductionism, another characteristic of postpositivism (Creswell & Plano Clark, 2011). The idea of reductionism also underlies the use of ratings to represent the task demands in that these ratings are assumed to provide valuable knowledge about the potential of tasks in terms of the mathematical competencies they involve.

3.1.2 Research design

The design of my research project was guided by the two aspects of its main objective: to investigate competency demands of tasks and teachers' knowledge about these task demands. In this study, data collection and analyses are dominated by a quantitative orientation but are informed and supported by qualitative data and methods. The quantitative and qualitative approaches are sequential rather than mixed and combined, and thus, the research design resembles what is called a quantitatively driven approach to multimethod research (Brewer & Hunter, 2006; Hesse-Biber & Johnson, 2015; Mark, 2015).

Two studies were conducted intended to provide complementary knowledge with regards to the competency demands of mathematical tasks and teachers' knowledge of these demands. Study 1 involved one set of mathematical tasks used in Norwegian secondary mathematics from two different assessments and focused on participating teachers' ability to recognise the competency demands of these tasks. Study 2 involved mathematical tasks used and selected by teachers and focused on teachers' considerations of the demands of these tasks. Study 1 preceded Study 2 in chronological order, and the results of Study 1 influenced and inspired Study 2.

3.1.3 The two studies

Study 1 – Article 1 and Article 2: This study aimed to investigate the extent to which a group of teachers could recognise the competency demands of mathematical tasks, which constitute an important aspect of task knowledge for teaching (see Chapter 2.4). Because of the ambiguous nature of the concept of mathematical competence, the MEG item analysis scheme was used to provide a common theoretical framework for the teachers that involved a set of competencies relevant to Norwegian secondary mathematics. The use of this scheme also provided a standardised instrument for rating competency demands, thereby allowing a comparison and further examination of their ratings through statistical approaches. The participating teachers first attended a training session in how to use the scheme and then used the scheme individually to analyse and rate the competency demands of 151 assessment tasks. The outcome of the teachers' ratings was analysed based on consistency estimates and descriptive statistics to evaluate the extent to which the teachers seemed to be able to recognise the demands for the six competencies. Furthermore, by combining the teachers' ratings with students' scored responses to the same tasks, a psychometric approach was conducted to further evaluate the ratings of competency ratings and to determine whether they reflected actual demands as experienced by students. This approach can be considered two-sided, either as a validation of the teachers' rated task competency demands or as a validation of the two assessments in terms of the mathematical competencies they can measure. These two sides and their relations are further discussed in Article 2. The assessment tasks stemmed from the PISA 2012 assessment and the Norwegian 2014 grade 10 exam, which are two externally mandated assessments implemented at the end of compulsory education in Norway to measure students' mathematical competence.

Study 2 – Article 3: The second study aimed to investigate in-service secondary school teachers' considerations of the demands of mathematical tasks they have previously used in their teaching practices, which related to another aspect of teachers' mathematical-task knowledge. As the results from Study 1 indicated that the participating teachers might struggle with recognising the demands of more complex and high-demanding tasks, the teachers were asked to submit tasks they had used to challenge their high-achieving students. Along with each task, the teachers submitted a task questionnaire designed to provide information about the teachers' use of the submitted tasks and their considerations of task demands through the following question: what do you think makes this a demanding task for high-achieving students? The intention of this open-ended question was to have teachers' use their own terms and vocabulary to describe the aspects they consider most important, rather than providing a theoretical framework of mathematical competence to guide their reflections about task demands (as in Study 1). Thus, Study 2 adopts a more exploratory approach than Study 1. To examine teachers' considerations of task demands, content analysis was conducted deductively, wherein the six MEG competencies formed the coding categories, and then inductively. Furthermore, the results from this analysis were compared with an analysis of the competency demands of submitted tasks.

The teacher-submitted tasks were seen to further add to the pool of mathematical tasks. When put together, the two studies involve a rather large number of tasks used in Norwegian mathematics classrooms. Thus, the analysis of these is intended to provide a rough picture of the competency demands of tasks provided to Norwegian secondary students.

Table 1 provides an overview of the main elements of the research process involved in the three articles: main aims, research questions, and data material and analysis.

Table 1. Overview of main elements of research process involved in the three articles.

	Study 1		*Study 2*
	Article 1	**Article 2**	**Article 3**
Main aim	Investigate teachers' recognition of the competency demands of tasks through the use of an analysis scheme	Further scrutinise the teachers' rated competency demands of tasks	Investigate teachers' considerations of task demands
Research question(s)	To what degree do a group of teachers and prospective teachers consistently analyse the competency demands of [mathematical assessment] tasks?	To what extent do differences in teacher-rated demands of the six MEG competencies in mathematics assessment items align with the differences in empirical item difficulty?	(1) What characterises teachers' considerations of task demands? (2) How do these considerations align with the competency demands of the tasks according to a competency framework?
Data material	Teachers' (n = 5) ratings of competency demands of tasks (n = 141)	Teachers' (n = 5) ratings of competency demands of tasks (n = 141) Students' task responses (n = 4686 + 1312)	Tasks (n = 78) submitted by teachers (n = 7) Teacher considerations of task demands (n = 35)
Data analysis	Analysis of task demands Rater consistency estimates Descriptive statistics	Analysis of task demands Explanatory item response modelling	Analysis of task demands Content analyses (deductive and inductive) Descriptive statistics

3.2 Participants

The study participants were two groups of Norwegian mathematics teachers, including in-service, pre-service, and former teachers with teaching experience from lower and/or upper secondary school mathematics (see Articles 1 and 3 for a more thorough description of the teachers).

Neither Study 1 nor Study 2 aimed at providing a representative sample of teachers in Norway. The teachers in Study 1 (n = 5) were purposefully sampled based on their teaching experience, educational background, and availability for training and analysing tasks. These teachers were all linked to the University of Oslo as students or former students at the teacher education programme or as employees. In Study 2, the rather low number of teachers (n = 7) was due to recruitment difficulties. More than 100 schools were approached by email. In addition, teachers were approached at lectures held at the University campus and through personal emails. Many of these teachers were positive to participating and received information and questionnaires. Nonetheless, a vast majority finally chose not to participate. The participating teachers in Study 2 worked at schools located in Oslo and Akershus.

3.3 Data collection and analysis

My research revolved around two main concepts: competency demands of mathematical tasks and teachers' knowledge about these task demands. The following sections describe the data collection and analysis conducted to investigate these concepts.

3.3.1 Mathematical tasks

In this thesis, the mathematical tasks involved can be seen to function both as an object of research (to examine the competency demands of tasks used in Norwegian secondary mathematics) and as a means of research (to provide an insight into teachers' recognition and considerations of the competency demands of tasks), both of which are involved in each of the two studies.

Study 1: The mathematical tasks involved in the first study were taken from two different assessments: 85/84 mathematical tasks[4] from the PISA 2012 survey and 56 tasks from the Norwegian 2014 grade 10 national mathematics exam. Both the PISA survey and the national exam have been developed to assess 15-year-old students' mathematical competence at the end of their compulsory education (i.e. grade 10 in Norway), and their underlying frameworks have been influenced by the notion of mathematical competence provided by the KOM report (Niss, 2015; Valenta et al., 2015). Thus, these tasks are seen to be rooted in a competency-oriented view on mathematics education and are therefore relevant to be studied from a competency perspective.

[4] While the teachers' analysis in Article 1 involved 85 PISA tasks, Article 2 involved only 84 of these tasks because one task was deleted from the data at a national level owing to problems with the functioning of the task (OECD, 2014). Thus, because this task lacked data for student responses, it could not be used in the psychometric modelling approach.

The selection of assessment tasks from a Norwegian exam and PISA 2012 was partly motivated by the validation approach used in Article 2. This approach required a rather large number of student responses, i.e. data that could be obtained for research purposes (see Chapter 3.3.4). Ideally, to investigate whether the ratings of competency demands are reflected in task responses, the psychometric modelling approach should be based on assessments developed to measure the competencies of interest (i.e. MEG competencies), where the assessment items vary systematically with regard to these competencies (Wilson, De Boeck, & Carstensen, 2008). Although the PISA and exam tasks are not developed to vary systematically across the six MEG competencies and four levels of demand, the two assessments can be seen to involve these competencies at a more conceptual level in the item development process through the influence of the KOM framework.

Study 2: The mathematical tasks involved in the second study were submitted by teachers based on the following two criteria: the tasks had to be (1) previously used at a secondary school (i.e. grades 9–11[5]) with the intention to (2) challenge high-achieving students. The intention of the first criteria was to provide an insight into the teachers' considerations of the demands of tasks they had used in teaching practices. The second criteria was inspired by the results from Study 1 that indicated that it was more challenging for the teachers to recognise the competency demands of higher-demanding tasks. Thus, the intention was to further explore teachers' analysis of high-demanding tasks when they were not provided with an analysis scheme. The 78 submitted tasks involved instructional and assessment tasks and consisted of both teacher-made and teacher-selected tasks, where the sources included textbooks, tests, and websites. It would be fair to assume that the submitted tasks do not represent typical tasks used by Norwegian teachers or even the participating teachers but rather the tasks they consider 'worthy' and confident of submitting to a research project. In this case, the tasks can be seen to represent tasks they associate with successful mathematics teaching.

3.3.2 Analysis of mathematical tasks

A key component of this research project was the analysis of mathematical tasks in terms of their demands for six mathematical competencies. Both the tasks from the two assessments (Study 1) and the teacher-submitted tasks (Study 2) are in a written format involving numeric

[5] In Norway, lower secondary school includes grades 8–10 and upper secondary school includes grades 11–13

and mathematical symbols, text, graphs, tables, illustrations, formulas, and/or other mathematical representations. The mathematical tasks pose a problem or a question for the students to solve or answer. The task analysis procedure applied in this research started with what can be regarded as a mathematical analysis where the mathematical knowledge, skills, and cognitive processes involved in the solution of the task was inspected. Next, the MEG scheme was implemented, and the perceived demands were matched with the operational definitions of four rating levels (0–3) for each of six mathematical competencies: Communication, Devising strategies, Mathematising, Representation, Symbols and formalism, and Reasoning and argument (for a more exhaustive description of the competency rating process, see Article 1). Thus, the competency demands of a task were represented by a set of six ratings on a scale of 0–3. The MEG item analysis scheme was used in this research based on the fact that the notion of mathematical competence in the MEG scheme was built on the same framework (i.e. the KOM framework) that had influenced the Norwegian mathematics curriculum. Thus, the identified competency demands were relevant to the Norwegian context of this thesis.

This type of analysis of textual data through the use of categories and coding resembles content analysis, a widely used and well-established method in both qualitative and quantitative research (Silverman, 2011). While content analysis was originally regarded as a quantitative method ‘for the objective, systematic, and quantitative description of the manifest content of communication’ (Berelson, 1952, as cited in Flick, 2014, p. xx), recently, it has been used for a range of analytical approaches involving both impressionistic, intuitive, and interpretive analyses as well as more systematic and strict textual analyses (Hsieh & Shannon, 2005). The task analysis conducted in the current research can be viewed as a systematic approach to ‘discover’ the objective descriptions of the competency demands of mathematical tasks in line with a postpositivistic perspective (see Chapter 3.1.1). However, the argument here is that the task analysis process did not resemble an automatic procedure involving counting distinct elements as in quantitative content analysis (Mayring, 2014). Identifying cognitive processes involved in mathematical tasks is recognised as highly challenging (Haladyna & Rodriguez, 2013; Lane, 2004), and coding of competency demands required a profound analysis which sometimes was closer to an act of interpretation rather than automated coding.

The rated demands of mathematical tasks played a somewhat different role in the two studies involved in the research. In Study 1, the analysis of competency demands functioned both as data collection (to investigate teachers' ability to recognise these demands) and data analysis (to investigate the demands of the assessment task). In Study 2, the task analysis was only a part of the analytical process of teachers' submitted tasks.

3.3.3 Teacher considerations and use of tasks

In Study 2, a questionnaire was used to collect data about the teachers' reflections about and their use of the mathematical tasks they submitted. This task questionnaire (see Appendix 2) asked the teachers about how they had used the submitted tasks in their teaching practices, for instance, whether a task had been used for individual work or group work or for instructional or assessment purposes. In addition, the questionnaire included an open-ended question: "What do you think makes this a demanding task for high-achieving students?" This was intended to gain insight into what the teachers considered to be key aspects of the task with regard to task demands. The use of interviews involving follow-up questions or prompts instead of questionnaires might have yielded richer data in terms of teacher considerations and mathematical-task knowledge. However, teachers' personal reflections unaffected by the presence of the researcher and his or her questions can be seen to better reflect the considerations teachers make when planning and selecting tasks for their teaching practices. The questionnaire was piloted to ensure that the questions were unambiguous and understood as intended. Even though the intention of the second study was to investigate teachers' considerations of the competencies required to solve mathematical tasks, it was decided not to include the term 'competency' in the open-ended question to avoid steering the teachers' thoughts in a certain direction and to have them consider and put forth all types of task features and characteristics as a source of task demand.

The teachers' considerations of task demands were analysed using a content analysis approach, as further described in Article 3. The content analysis consisted of a deductive and an inductive phase. The deductive coding adopted a competency perspective to examine the extent to which references to the six MEG competencies could be observed in the teachers' considerations. As it was expected that the considerations involved aspects not related to the six competencies, inductive coding was conducted to explore the further explanations provided by the teachers.

3.3.4 Secondary analysis of PISA and exam data

The term secondary analysis refers to the analysis of already existing data conducted by researchers that have not been involved in the data collection process, where the purpose of the analysis was probably not envisaged by those responsible for the data collection (Bryman, 2016). This is the case for the psychometric modelling described in Article 2 which involved student responses to PISA and exam tasks gathered through the PISA 2012 survey and the Norwegian Ministry of Education and Research, respectively. Bryman (2016) lists several advantages of secondary analysis. One advantage that is crucial to my research is the possibility of using numerous data sets that would not be possible to collect yourself. An example of this is the 1300–1400 student responses to each of the 140 tasks involved in Article 2. The validation of the teachers' ratings based on the explanatory item response modelling approach required such a large number of responses to provide measures that were sufficiently accurate to identify and disentangle the demands of the six competencies.

An important limitation with secondary analysis is that because the data has been collected for other purposes than yours, some theoretically important variables and aspects might be missing (Bryman, 2016). Even though the two assessments were not developed explicitly to measure the six competencies involved in the MEG framework, the constructs of mathematical competence involved in both the PISA assessment and the Norwegian exam are heavily influenced by the notion of mathematical competence presented in the KOM framework on which the development of the MEG scheme is based (Niss, 2015; Turner et al., 2015; Valenta et al., 2015). Accordingly, the linking of students' scored responses with the rated demands of these tasks was assumed to be appropriate at a theoretical level, although the missing Aids and tools competency resulted in some misalignment with the Norwegian curriculum (see Chapter 2.2.2).

Explanatory item response modelling was used to provide empirical evidence about the extent to which the teachers' ratings of competency demands actually reflected the demands students encounter when solving the tasks. Explanatory item response models offer the opportunity to examine the effects of specific item features, that is, whether properties of items, such as competency demands, can explain the responses generated by them (Koeppen et al., 2008; Wilson et al., 2008). This made it possible to link the rated demands for the individual competencies to student responses (for a more exhaustive description of this approach, see Article 2). Being able to model and explain item responses, usually represented by item difficulty (proportion of correct/incorrect responses), based on the identified features of test items is important to understand what is measured in tests (De Boeck, Cho, & Wilson, 2016; Graf, Peterson, Steffen, & Lawless, 2005) and has been used as empirical evidence validity for many tests (see e.g. Enright, Morley, & Sheehan, 2002; Freedle & Kostin, 1993; Gorin & Embretson, 2006). In my thesis, this approach was also used to further evaluate the extent to which the teachers involved in Study 1 were able to recognise the demands for the different competencies in the assessment tasks. A covariation between the rated demand for a certain competency and item difficulty shows a relationship between higher ratings and more demanding tasks, which indicates that the ratings reflect an actual demand experienced by the students when engaging the task. Conversely, if no covariation is found, this shows that higher rated demands do not necessarily reflect a higher experienced demand. This missing covariation could be due to teachers' lack of ability to recognise competency demands but could also result from other factors. For instance, the variation in the assessment tasks in terms of demands for different competencies could be too low to obtain statistically significant relationships between rated demands and item difficulty. Another possible source is the fact that the MEG item analysis scheme and its operational definitions involve strong overlaps between competencies or do not adequately represent the demands students actually experience. Article 2 further discusses the different plausible explanations of the results from the explanatory item response modelling.

3.4 Research validity

The term validity refers to the rigor of research, and several concepts of validity can be found in literatures related to both quantitative and qualitative research (Creswell & Plano Clark, 2011; Kleven, 2008; Newton & Shaw, 2014). Shadish, Cook, and Campbell (2002) define validity as the approximate truth of an inference, whereas Creswell and Plano Clark (2011) describe validity as the accuracy and trustworthiness of the interpretations and conclusions drawn by the researcher A key idea is that validity is not a property of the data, design, or methods involved in the research but of whether the inferences and conclusions drawn by the researcher based on the results of these data, design, and methods are trustworthy (Creswell & Plano Clark, 2011; Shadish et al., 2002). The following discussion draws on the validity system presented by Shadish et al. (2002) involving four types of validity: construct validity, statistical conclusion validity, internal validity, and external validity. In this validity system, evidence to support inferences is based on both empirical evidence and other sources of knowledge, such as findings and theories obtained from previous research (Shadish et al., 2002). As such, the system can be seen to be rooted in postpositivism and is relevant to both qualitative and quantitative research (Kleven, 2008).

3.4.1 Construct validity

Construct validity refers to the correspondence between a theoretical construct and the operationalisations and measures of this constructs (Shadish et al., 2002). In my research, the main construct of interest is overall mathematical competence and individual competencies, and the following discussion mainly concerns the ratings of competency demands as valid operationalisations of the mathematical competencies needed to be able to solve mathematical tasks. The analysis and rating (i.e. measurement) of the competency demands of mathematical tasks was based on an item analysis scheme. Two key issues concern the soundness of these competency ratings and the extent to which the conclusions I draw based on these are valid: the appropriateness of the MEG item analysis scheme and the process of rating competency demands of tasks.

The MEG scheme has previously been shown as an efficient instrument to rate the competency demands of mathematical problems, where these ratings were found to relate to the difficulty of the problems (Turner et al., 2013). This supports the MEG scheme as an appropriate operationalisation of competency demands in mathematical tasks. Still, a main issue with operationalising and measuring a set of individual mathematical competencies is that these competencies are often overlapping by definition (Niss et al., 2016). With regard to the rating process, this could mean that raters fail to distinguish between demands for different competencies and that tasks perceived to have high demand receive high, inflated ratings across competencies that do not reflect the actual competency demands. This phenomenon, wherein a global evaluation is found to potentially influence the evaluation of several individual attributes, is recognised in literature as the so-called halo effect and is considered a threat to the validity of ratings (Feeley, 2002; Nisbett & Wilson, 1977). Another issue related to this is the clarity of the definitions in the MEG scheme, as unclear and fuzzy categories are recognised as another source of halo errors (Feeley, 2002). According to Turner et al. (2015), in the development of the MEG scheme, there was a focus on maximising the distinction between the competencies, clarifying definitions and descriptions and reducing the use of relative terms with subjective meaning. To further strengthen the validity of the ratings as adequate measures of competency demands, the teachers that analysed the assessment tasks in Study 1 participated in a training session to enhance their understanding of the six competencies and provided experience with applying the scheme. The use of rater training is found to potentially reduce rater bias (Feeley, 2002). Nonetheless, the analysis of competency demands can be seen to involve interpretive elements (see Chapter 3.3.2) and the scheme includes subjective words with relative meanings (e.g. 'simple' and 'complex'). Thus a different group of raters may have ended up with slightly different sets of ratings. Furthermore, different approaches were implemented to increase the quality of the ratings and to provide evidence to support their appropriateness as measures of competency demands. In both studies, multiple raters were used to ensure the reliability of the ratings. For Study 1, the high consistency reported in Article 1 indicated that agreement among the teachers strengthens the validity of the ratings. Similarly, check-coding (Lewis, 2009) was performed as a validation approach in Study 2, where two researchers first individually rated the demands of all tasks before comparing ratings and discussing differences to reach mutual agreement about competency ratings.

Several authors (e.g. Embretson & Gorin, 2001; Messick, 1995) argue that being able to understand and identify the processes that underlie item responses is a key aspect of construct validity, where the modelling of task responses and item difficulty is regarded as a possible source of empirical evidence. The explanatory item response modelling approach presented in Article 2 links the teachers' rated competency demands with student responses and thus functions as a source of empirical evidence for construct validity. The results of this approach showed that considerable variation in task difficulty could be explained by the rated competency demands, thus strengthening the validity of the ratings as adequate measures of task demands. Furthermore, the results revealed considerable differences between the different competencies in terms of their relation to task difficulty, thus confirming the adequateness of some measures of competencies and invalidating others. The empirical evidence from the explanatory item response modelling shape the discussion of the results and findings presented in Chapter 4.

Another question regarding the use of the scheme is whether different task competency demands would have been identified through the use of another theoretical framework. The answer to this is yes, as different constructs of mathematical competence coexist. According to Niss et al. (2016), the discussion of interest is not whether a construct of mathematical competence is right or wrong or whether it is correct or incorrect but whether it contains relevant features and whether it serves its intended purpose. Nonetheless, although different competency frameworks involve different strands and competencies, they are similar in that they intend to capture what it means to master mathematics at large including a variety of knowledge and processes (Kilpatrick, 2014a; Niss & Jablonka, 2014). This means that results related to certain individual competencies might depend on the use of the MEG scheme but that the main findings related to the presence of a general mathematical competence focus in Norwegian secondary mathematics classrooms should not. Furthermore, the MEG scheme and its six competencies are seen as highly relevant for the Norwegian context of this thesis (see Chapter 2.2.2). However, the omission of an Aids and tools competency is a source of construct underrepresentation. Construct underrepresentation refers to the possibility that the indicators fail to cover the entire construct is seen as a threat to construct validity (Messick, 1995). Thus, for the analysis of task demands, the inclusion of an aids and tools competency would have further strengthened the validity and relevance of my findings. This lack of an aids and tools competency in the ratings of competency demands of tasks is considered in the inferences and conclusions drawn in Chapter 4 to strengthen their validity. In the analysis of teacher considerations of task demands in Study 2, an inductive content analysis was conducted to identify aspects that were not captured by the six MEG competencies, such as an aids and tools competency. This inductive analysis phase was conducted to counter the validity threat of construct underrepresentation.

3.4.2 Statistical conclusion validity

Statistical conclusion validity deals with statistical inferences and the covariation between variables (Shadish et al., 2002). Kleven (2008) argues that this is not solely about statistical methods such as significance tests and estimates of effect size but also, more generally, about whether a tendency is substantial enough to be worthy of interpretation. Several covariations between variables related to the ratings of competency demands are involved in the different articles, such as consistency estimates (Article 1), correlations in ratings between competencies (Article 2), and regression coefficients (Article 3). The consistency estimates and correlations relate to the reliability of the ratings and the overlap between the competencies, whereas the regression coefficients relate to the linking of rated demands and task responses through the explanatory item response modelling approach. Consequently, these covariations are important with regard to the validity of the inferences and conclusions that are drawn in this thesis. The significance and meaning of the covariations are discussed in the different articles.

3.4.3 Internal validity

While Shadish et al. (2002) refer to internal validity as inferences about causal relationships based on observed covariations between variables, Kleven (2008) applies a broader perspective arguing that internal validity ‘is important whenever we infer that something has an influence on something’ (p. 228). My research is not designed or intended to investigate causal relationships, and the results and findings do not search for or draw such causal inferences. Nonetheless, in the discussions of the results, some relationships are implied. For instance, one inferred implication is that the use of the MEG item analysis scheme could support teachers in their ability to recognise competency demands. This inference is built on rational arguments rather than statistical control, partly based on the results of the two studies involved in this research project and partly based on findings from previous studies. It is not my intention to confirm this or other possible causal relationships.

3.4.4 External validity

External validity addresses the issue of generalization and transferability, namely, whether the inferences drawn from the context of the study are also valid for other contexts outside the study (Shadish et al., 2002). While statistical generalizations require probability samples of the population of interest, which is seldom obtained in educational research, the external validity of research often relates to non-statistical judgement-based generalizations based on rational arguments (Kleven, 2008).

The intention of this research project was not to make claims about the mathematical-task knowledge of Norwegian mathematics teachers in general but rather to examine teachers’ knowledge about task competency demands at a smaller scale and to discuss the relevance and transferability of these to a broader population of Norwegian mathematics teachers. Relatively few teachers have participated in the two studies and in terms of mathematical-task knowledge, it is probable that they not reflect typical teachers. The teachers in the second study were a few of the very many who had been approached for participation. One reason for their participation could be that they are more confident in their knowledge about mathematical tasks and their teaching practices. In addition, a high proportion of the submitted tasks were developed by the teachers themselves, which could further indicate that the teachers were confident about their mathematical-task knowledge. Teachers with such confidence might not represent typical teachers. Thus, when discussing the results and findings, I am cautious not to generalise to the whole population of teachers, for instance, by referring to *the* (participating) teachers rather than teachers in general.

The other aspect of the main objective relates to mathematical tasks used in Norwegian secondary education. The mathematical tasks (n = 219) involved in the two studies are different types of tasks, such as assessment tasks and instructional tasks, that have been used in Norwegian secondary mathematics. As teaching practices are assumed to be adapted to fit with what is measured in high stakes national assessments (Boesen et al., 2014; de Lange, 2007), such as the Norwegian grade 10 exam, the exam tasks are likely to represent tasks typically used in Norwegian classrooms. As discussed in Chapter 3.3.1, the tasks involved in the second study are likely to be tasks the participating teachers consider as 'worthy' tasks used to challenge high-achieving students, but where around half of the tasks have been used in whole-class settings. Thus, rather than providing an extensive mapping of tasks in use in mathematics education in Norwegian secondary schools, the two studies investigate the competency demands of a somewhat broad but limited selection of tasks used in Norwegian secondary education. Nonetheless, I argue that this provides important knowledge that adds to the existing body of knowledge on mathematics teaching and learning as well as classroom practices.

3.5 Ethical considerations

The ethical issues in my research project mainly concern the anonymity and privacy of the participating teachers, and ensuring the same was a main focus in data collection and data handling. All participating teachers were approached through email. In line with the guidelines from Norwegian Social Science Data Services (NSD) and Norwegian National Research Ethics Committees (NESH, 2016), the participants were informed of the purpose of the study they were involved in, that their involvement was voluntary and that they could withdraw from the study any time. They were also informed about the anonymity and confidentiality of the data collected and submitted free and informed consent before participation (Appendix 3). No personal data, i.e. data that could directly or indirectly be related to a person (NESH, 2016), was gathered from the participants. The teachers were prompted to not provide their name or any other identifiable information in the submitted material, and written consent including participants' names were separated from the data material. Furthermore, to ensure anonymity, both the teachers' analysis of tasks (Articles 1 and 2) and the teachers' submitted tasks and questionnaires (Article 3) were submitted in a paper-based format to prevent data from being linked back to individual teachers (for instance through email and IP addresses).

My research also used existing data from the PISA survey and the Norwegian 2014 Grade 10 examination. The (re)use of data (e.g. PISA and exam data) for secondary analysis can be considered a gentle approach with regard to students' anonymity, as this is ensured in the original data collection and handling process. Students' scored responses to the Norwegian exam were provided by the Norwegian Directorate for Education and Training. The students, schools, and 'oppmenn' (i.e. trained teachers scoring the responses) involved in the data set were anonymised by the directorate before the data were provided to me. The use of the data relied on several conditions, for instance that my research project must be found worthy of using these data, that the data is only used for the research described in the application, and that a declaration of confidentiality had to be signed. The Norwegian Directorate for Education and Training also granted access to the set of tasks involved in the Norwegian 2014 Grade 10 examination. The students' scored responses to the PISA tasks were obtained through the Norwegian PISA group (but are also available through the official PISA webpage). These data were collected after obtaining participants' free and informed consent (in Norway, participation in large international studies, such as PISA and TIMSS, was made mandatory from 2014). The Norwegian PISA group also granted permission to use tasks from the PISA survey with certain reservations about the confidentiality of non-published tasks. Confidentiality schemes were submitted by both me and the teachers involved in Study 1, where it was requested that the no one else would be informed of the tasks and that no copies of the tasks could be made. The PISA tasks shown in Articles 1 and 2 are all released tasks that are publicly available on the official PISA website.

4 Towards competency-oriented mathematics education

Notions of mathematical competence that have influenced mathematics curricula around the world have introduced an enriched view of what it means to master mathematics that involves a range of mathematical competencies, thus extending the traditional focus on factual knowledge and procedural skills. The main objective of this thesis is to investigate the competency demands of mathematical tasks used in Norwegian secondary mathematics classrooms and teachers' recognition and considerations of these demands, both of which are important components of competency-oriented mathematics education. Two studies, resulting in three separate articles, have been conducted to investigate different aspects of this main objective. In this final chapter, I summarise the three articles with a focus on their results and findings before discussing the findings in light of the main objective and the broader educational context. Following this discussion, I discuss the contributions, implications, and limitations of my research and provide some concluding remarks. Finally, I suggest some future directions for realizing competency-oriented mathematics education.

4.1 Summary of articles

4.1.1 Identifying competency demands in mathematical tasks: Recognising what matters (Article 1)

Pettersen, A., & Nortvedt, G. A. (2018). Identifying competency demands in mathematical tasks: Recognising what matters. *International Journal of Science and Mathematics Education, 16*(5), pp. 949–965. doi:10.1007/s10763-017-9807-5

The aim of Article 1 was to investigate teachers' ability to recognise the demands for six mathematical competencies involved in assessment tasks through the use of the MEG item analysis scheme. The research question regarded the degree to which the teachers consistently analysed the competency demands of tasks. In this study, five mathematics teachers and prospective teachers (hereafter referred to simply as teachers) attended a training session intended to enhance their understanding of these competencies and how to apply the item analysis scheme to rate the competency demands of tasks. Following the training session, the teachers applied the scheme to individually rate a set of assessment tasks from the PISA 2012 survey (n = 56) and a Norwegian grade 10 national exam (n = 85). Both assessments have been developed to measure students' mathematical competence, and the tasks were seen to represent different types of tasks in which Norwegian mathematics teachers should be familiar (see Chapter 3.3.1).

The teachers' ratings of the competency demands of the assessment tasks were analysed using different statistical approaches. In response to the research question, estimations of rater reliability and agreement revealed fairly high consistency among the teachers' in their ratings for all six competencies across the two assessments, ranging from agreement measures of .88 for Devising strategies to .80 for Mathematising. Further inspection of the estimates revealed a somewhat lower consistency for more complex tasks, namely, tasks that demanded the activation of multiple competencies, especially for the Mathematising competency. Descriptive statistics of the distribution of the teachers' ratings showed that a vast majority of the ratings were at the lower levels of cognitive demand, indicating that the teachers rarely judged a task to demand a high level of competency. This latter result was somewhat surprising considering the fact that students' success rates in solving the tasks show that highly challenging tasks are included in both assessments. Based on these results, we argued that the teachers were rather successful in identifying the involvement of the six competencies in tasks but that demands at higher levels and in more complex tasks were more challenging to recognise.

The small number of teachers and prospective teachers participating in this study means that no general claims should be made about teachers. However, the rather large number of tasks and the high consistency in the teachers' ratings provided a foundation for further statistical analysis and validation of the teacher-rated competency demands of the assessment tasks through the use of an explanatory item response modelling approach. This approach is described in Article 2.

4.1.2 Mathematical competency demands of assessment items: A search for empirical evidence (Article 2)

Pettersen, A., & Braeken, J. (2017). Mathematical competency demands of assessment items: A search for empirical evidence. *International Journal of Science and Mathematics Education*. doi:10.1007/s10763-017-9870-y

The main aim of Article 2 was to further scrutinise the teachers' ratings of the competency demands of the assessment tasks involved in Article 1 through what can be considered an empirical validation of the ratings. The following research question was addressed: To what extent do differences in teacher-rated demands of the six MEG competencies in mathematics assessment items align with the differences in the empirical item difficulty?

The scored responses of around 1400 and 4700 students for the exam and PISA tasks, respectively, were linked to the teachers' ratings of task demands through an explanatory item response modelling approach. The average ratings on each of the six competencies given by the five teachers were used as tasks demands. This use of averaged ratings was supported by the high interrater consistency found in Article 1.

Addressing the research question, the results showed that when including the demand for all six competencies in the explanatory model, slightly more than and less than half of the variance in task difficulty could be explained for the PISA and exam data, respectively. This level of explanatory power for modelling item difficulty has previously been considered strong (e.g. Embretson & Daniel, 2008) and thus provides some empirical evidence supporting the validity of the teachers' ratings of competency demands. When examining individual competencies' relation to task difficulty, considerable differences were observed. Across the two assessments, the rated demands for Symbols and formalism and Reasoning and argument had a significant and strong association with task difficulty, whereas no such relationship was observed for Mathematising and Communication. For Devising strategies and Representation, a statistically significant relationship was found between the rated demands and task difficulty for the PISA data but not for the exam data. These findings suggest that the teachers were more successful in recognising and rating the demands for some competencies.

Some plausible explanations were provided for the unexplained variance in task difficulty. For instance, a substantial correlation in ratings was found between Mathematising and Devising strategies and between Mathematising and Reasoning and argument. This could indicate difficulties with separating the demands for the different competencies and could explain the low explanatory power of the rated demand for some competencies (e.g. Mathematising). According to Niss et al. (2016), the overlapping and intertwined nature of mathematical competencies is one of the challenges with measuring them. In literature, training and experience as well as clearly defined and concrete categories have been recommended for increasing raters' ability to discriminate between categories (Feeley, 2002). Thus, we argued that further clarifying definitions and descriptions in the MEG scheme as well as more rater training could have better supported the teachers in separating the demands for the different competencies and increased the accuracy of their ratings. Another explanation relates to features or aspects that are not captured by the rated demands for the MEG competencies, such as the lack of the Aids and tools competency in the analysis scheme (further discussed in Chapter 2.2.2).

Another aspect of the psychometric modelling of the cognitive complexity of assessment items is providing empirical evidence to ensure and support claims about the validity of tests and assessments (Embretson & Gorin, 2001; Lane, 2004; Messick, 1995). This aspect is highly relevant to the assessments of students' mathematical competence, which is an important aspect of the implementation of competency-oriented mathematics curricula (see Chapter 2.3). The results revealed that only two of the competencies are related to task difficulty for the national exam, compared to four for the PISA assessment. The weak link between the six competencies and the exam tasks was not unexpected, as many exam tasks can be regarded as non-contextualised tasks that mainly assess procedural skills. Based on these results, we questioned the extent to which the Norwegian grade 10 national exam captures the various cognitive skills and abilities that are represented in mathematical competence.

4.1.3 Teachers' considerations of mathematical tasks used to challenge high-achieving students (Article 3)

Pettersen, A., & Nortvedt, A. G. (under review). Teachers' considerations of mathematical tasks used to challenge high-achieving students. *Scandinavian Journal of Educational Research*

Article 3 aimed to investigate teachers' considerations of tasks that they had selected and used previously in their mathematics teaching to challenge high-achieving students. This focus was partly shaped by the results of the first study in which the teachers seemingly struggled with recognising competency demands in complex and high-demanding tasks (Article 1). Two research questions were raised: (1) What characterises teachers' considerations of task demands? (2) How do these considerations align with the competency demands of the tasks according to a competency framework? The data for this study involved seven secondary school teachers (grades 9–11) who submitted 78 mathematical tasks along with their considerations of task demands and information about the source and usage of the tasks. Compared to Study 1, this study followed a more exploratory approach as the teachers were not guided by a certain competency framework in their reflections on task demands and as the mathematical tasks were self-selected by the teachers. Surprisingly, about half of the teacher-submitted tasks were developed by the teachers themselves. Deductive and inductive content analysis was conducted to analyse teachers responses to the following open-ended question: “What do you think makes this a demanding task for high-achieving students?” Furthermore, the submitted tasks were analysed using the MEG scheme.

For the first research question, the results showed that in their considerations, the teachers mainly emphasised two competencies: Symbols and formalism and Devising strategies. These accounted for nearly 60% of teachers’ explanations in the deductive coding process. Other competencies were far less emphasised by the teachers. In particular, few references were made to the Mathematising competency; it appeared in only three considerations. For the second research question, strong similarities were found when comparing teachers’ considerations with the rated competency demands of the submitted tasks. However, when comparing individual teachers, some differences were observed both in considerations of tasks and in the competency demands of the submitted tasks. Some teachers provided considerations that were more aligned with the rated competency demands of the tasks, and some submitted tasks that involved a wider range of competencies at higher levels of competency demands. We argued that this indicated that there were disparities in teachers’ reflections and awareness of the mathematical competencies involved in mathematical tasks and that this could influence their ability to select appropriate tasks that engage students in a variety of competencies.

Based on these findings, we argued that the teachers provided insightful considerations of the tasks they submitted. However, from a competency perspective, the prominent role of some competencies illustrated a somewhat unbalanced focus when selecting and analysing tasks and task demands. The findings also suggested large differences between the teachers both with regard to the challenges they provide for their students as well as their knowledge and considerations of task demands. We argue that these aspects would influence teaching practices and the opportunities provided to students to engage in and develop a variety of mathematical competencies.

4.2 Main contributions

In line with its main objective, the main contribution of this thesis is to add to the empirical understanding of mathematical competencies in Norwegian secondary mathematics. The analysis of the demands for six mathematical competencies in assessment tasks (Study 1) and teacher-submitted tasks (Study 2) based on the MEG scheme provided insights into the range of competencies and levels of cognitive demands involved in Norwegian mathematics classrooms. Furthermore, insights into two aspects of teachers' mathematical-task knowledge were obtained by examining teachers' ability to recognise the competency demands of tasks through the use of the MEG scheme (Study 1) and teachers' considerations of task demands (Study 2). In addition, the thesis includes theoretical contributions regarding the nature of mathematical competencies and the use of the MEG scheme as well as a main methodological contribution with regard to the explanatory item response modelling approach. These contributions are discussed further below.

4.2.1 Empirical contributions

The main empirical contribution of this research is increased knowledge about the competency demands of mathematical tasks used in Norwegian mathematics education and teachers' knowledge of these task demands. Although notions of mathematical competence comprising a range of individual competencies have gained a foothold in the Norwegian mathematics curriculum (Valenta et al., 2015) and curricula worldwide (Niss et al., 2016), the findings of this thesis suggest that not all individual competencies have a similarly strong standing in teaching practices in Norwegian classrooms.

The prominent role of the Symbols and formalism competency was apparent throughout this thesis. The demands of the tasks in the two studies are dominated by the Symbols and formalism competency, and the teachers seem well aware of the demand for this competency in their task analysis and their consideration of task demands. The Symbols and formalism competency focuses on the application of mathematical procedures, rules, and conventions (Niss & Højgaard, 2011) and thus resembles what traditionally has been referred to as procedural knowledge (Haapasalo & Kadijevich, 2000; Hiebert, 1986). The central role of procedural knowledge in mathematics education has been confirmed in many studies. For instance, in her thesis, Pedersen (2014) found a strong emphasis on the application of procedures and methods in the Norwegian curriculum for upper secondary school. Other studies have found a strong emphasis on procedural skills when investigating mathematical tasks and classroom activities (Boesen et al., 2014; Boesen et al., 2018; Dole & Shield, 2008; Hiebert et al., 2003; Kolovou, van den Heuvel-Panhuizen, & Bakker, 2011).

Over the last couple of decades, there has been a concern about Norwegian students' apparent lack of algebraic and procedural skills (Nilsen, Angell, & Grønmo, 2013; Pedersen, 2014). Accordingly, and considering the strong position of procedural knowledge in mathematics and mathematics education, it is not surprising that the Symbols and formalism competency appears to play such a dominant role in Norwegian secondary mathematics. Nonetheless, selecting and implementing tasks that involve the Symbols and formalism competency does not ensure that students master this competency sufficiently. For instance, a task in which student engagement is reduced solely to mindless copying and rehearsing of procedures is not likely to provide any development of the Symbol and formalism competency. Rather, successful mathematics teaching is associated with engaging students in tasks that promote a deep understanding and involve higher cognitive demands, for instance, tasks in which the use of procedures is connected to concepts or understanding (Stein et al., 1996). Identifying and using such tasks is a part of the nature of mathematical task-knowledge for teaching (Chapman, 2013).

A second finding concerns the demand for the Mathematising competency. The findings in the two studies suggest that teachers might not be aware of and recognise the role of the Mathematising competency in mathematical tasks and that the tasks which Norwegian students encounter do not elicit its use. The Mathematising competency deals with the translation of the extra-mathematical context into mathematical structures (and vice versa) and with interpreting and modifying mathematical models (Niss & Højgaard, 2011). This is considered one of the competencies or processes involved in mathematical modelling, that is the solving of real-world problems using mathematics (Blum, 2015; Blum & Ferri, 2009). Mathematical modelling has received increased attention in mathematics education and curricula over the last few decades (Blum, 1993; Lesh & Zawojewski, 2007; Lingefjärd, 2006), in line with the increased attention to process-oriented aspects of mathematics (see Chapter 2.1.1). Blum (2015) argues that the teaching of mathematical modelling needs to engage students actively and cognitively (and meta-cognitively), with a focus on modelling as a whole and on the sub-competencies of modelling (such as mathematising).

In this thesis, one apparent issue is the strong overlap between the teachers' ratings of the demand for the Mathematising competency with the demand for the Devising strategies competency. Turner et al. (2013) found a similar overlap in their analysis. The Devising strategy competency refers to the ability to select, construct, and activate a solution strategy. Thus, while the Mathematising competency involves the ability to translate a real-world problem into mathematics, the Devising strategies competency involves the next stage in a problem solution process where a strategy for handling this mathematics must be worked out. The results in the second study showed that the participating teachers were aware of and concerned with the demand for the Devising strategies competency in mathematical tasks. However, the demand for the Mathematising competency was either not recognised or not emphasised by the teachers when explaining the challenging aspects of a task. The high involvement of the Devising strategies competency could have resulted from the request for tasks that had been used to challenge high-achieving students as the teachers might consider that it is especially beneficial for these students to engage in tasks where they are required to select or devise problem solving strategies. Nonetheless, a majority of the tasks had been used in whole-class settings, indicating that teachers feel that students across achievement levels should engage in this competency.

Another empirical contribution relates to the Norwegian Grade 10 exam and the PISA survey as valid measures of mathematical competence. For both assessments, the substantial proportion of the variance in item difficulty that could be explained by the rated competency demands showed the relevance of the competencies in describing the task. However, for the Norwegian exam, the demands for only two of the competencies were identified and related to task difficulty. National assessments are considered an important tool for implementing curriculum ideas as teachers are assumed to tailor classroom activities and practices so that they reflect the demands of assessments (Boesen et al., 2014; de Lange, 2007), which is often referred to as the washback effect (Buck, 1988). The strong position of the Symbols and formalism competency in the teacher-submitted tasks and the teachers' considerations of task demands might result from such an effect, considering the dominant role of this competency in the Norwegian exam. Thus, ensuring that the Norwegian grade 10 exam involves a range of mathematical competencies would be an important means for influencing the types of tasks and learning opportunities teachers provide for their students.

4.2.2 Theoretical contributions

The main theoretical contribution relates to a theoretical understanding of the nature of mathematical competence. A substantial correlation in ratings of demand was observed between some competencies, for instance, between Mathematising, Devising strategies, and Reasoning and argument (reported in Article 2). This confirms the overlapping and intertwined nature of mathematical competencies illustrated in theoretical constructs of mathematical competence, such as the KOM framework (Niss et al., 2016). Being aware of these overlaps is important both from a research perspective, for example, when investigating mathematical competencies empirically, and a teaching perspective, for example, when planning mathematics teaching and learning activities to stimulate the development of these competencies. Nonetheless, further development of the MEG scheme and other theoretical frameworks that further clarifies and identifies features that distinguish these competencies would be important for future research and for the implementation of mathematical competencies in educational practices. For instance, when presenting and discussing notions of mathematical competence in teacher education and professional development programmes, it would be important to be aware of these overlaps and the challenges with distinguishing these competencies in tasks and other classroom activities.

Another theoretical contribution concerns the applicability of the MEG item analysis scheme. While the scheme was originally used by the PISA Mathematics Expert Group to analyse mathematical problems in the PISA survey, the results from Study 1 suggest that the scheme can also be used by teachers to analyse tasks and recognise demands for several mathematical competencies. Thus, the use of this scheme could support teachers in grasping notions of mathematical competencies and in implementing these competencies when planning lessons and teaching approaches.

4.2.3 Methodological contributions

The mathematical competency perspective adopted in my research, which involves an entire system of individual competencies, proved to be fruitful in the analysis of both the demands of mathematical tasks and in teachers' considerations of task demands. Through this perspective, considerable differences were observed between the tasks involved in two assessments with regard to the number and types of mathematical competencies involved. This perspective also made it clear that teachers recognised and emphasised some of these competencies but not others when analysing task demands.

The results presented in this thesis, where the demand for four individual mathematical competencies were distinguished and identified, illustrate some of the strengths and possibilities of explanatory item response modelling for investigating the competency demands of mathematical tasks. Although explanatory item response modelling by itself is not new, there are relatively few applications of this technique in mathematics education and for investigating the demands and features of mathematical tasks. One exception is the study conducted by Embretson and Daniel (2008), who applied an explanatory item response model (LLTM) to examine the sources of the cognitive complexity of items on a test of quantitative ability. Twelve item features were involved in the modelling approach, for instance, the number of words in the item stem, whether the required equation was given in words, and the number of equations needed to be recalled, and the implications of the results are mainly of interest to item development and test design rather than teaching practices. Most of the 12 item features were found to be significant predictors of item difficulty and, in total, around half of the variance in difficulty could be explained. Based on these results, Embretson and Daniel (2008) argued that the results supported the validity of the postulated model of cognitive complexity for mathematical problem solving. Another exception is Hohensinn and Kubinger (2009) who applied explanatory item response modelling to investigate the effect of different response formats on item difficulty and found that although the format may influence the difficulty of items, it does not seem to change the proficiency measured by the item.

Compared to the two abovementioned studies which involved item features, many of which can be regarded as surface characteristics, the current thesis demonstrates the possibility of identifying the influence of task competency demands that are more closely related to understanding, problem solving, and student learning in mathematics. As there has been a call for more empirical evidence to ensure that complex abilities and mental processes are properly captured by assessment items (Koeppen et al., 2008; Lane, 2004; Messick, 1995), this thesis shows that the explanatory item response modelling approach could potentially be one such source of empirical evidence and could provide valuable information for developing tasks that involves mathematical competencies.

4.3 Implications and concluding remarks

Based on the current investigation of the demands of mathematical tasks and teachers' knowledge about task demands, there seems to be a need for some further steps on the road toward competency-oriented mathematics education. While notions of mathematical competence implemented in curriculum reforms worldwide stress the importance of developing a variety of competencies, this enriched view on mathematical mastery does not seem to be reflected in classroom practices where carrying out procedures has and still does play a dominant role (Boesen et al., 2014; Hiebert et al., 2003; Kaur, 2010; Niss et al., 2016; Palm et al., 2011). The main objective of this thesis has been to contribute knowledge about the demands for competencies in mathematical tasks used in Norwegian secondary mathematics as well as teachers' recognition and considerations of these demands. The findings in the two studies suggest that the types of tasks used in Norwegian secondary mathematics do not seem to stimulate the development of the variety of competencies portrayed in notions of mathematical competence. The Norwegian exam tasks and the teacher-submitted tasks mainly require the use of few mathematical competencies above a very basic level, with the Symbols and formalism competency playing a prominent role.

In Norway, the mathematics curriculum is currently being revised. An important change is the implementation of core elements, or big ideas, intended to guide progression and promote students' development of the understanding of content and relationships within the subject (Stortingsmelding (White Paper) nr. 28, 2015–2016). The core elements in the subject of mathematics are exploring and problem solving, modelling and applications, reasoning and argumentation, representation and communication, abstraction and generalisation, and mathematical knowledge domains. These, to a large extent, match the competencies involved in my thesis. The implementation of these core elements in the Norwegian mathematics curriculum means that mathematics teachers at all educational levels need to be aware of these elements and plan and implement classroom activities that stimulate their development. The findings in this thesis indicate the necessity of a shift in current practices.

How then could ambiguous and complex mathematical competencies move from being merely theoretical phenomena referenced in academic literature to being implemented in classroom activities? Niss and Højgaard (2011) argue that a first step toward promoting mathematics teaching that is in line with a competency perspective on mathematical mastery is to increase the awareness of mathematical competence and competence thinking when planning, arranging, and implementing teaching. Hopefully, this thesis could contribute to such awareness. Furthermore, Sierpinska (2004) argues that one of the strengths of the concept of mathematical tasks is that it forms a common ground for where teachers and researchers meet. Thus, tasks and task analysis can be seen as a gateway for supporting teachers to grasp competencies and implement them in their teaching practices. This thesis has shown that the use of an analysis scheme could support teachers in recognising competency demands in tasks. Similar promising results have been obtained by other studies through the use of similar tools for recognising key characteristics of cognitively demanding tasks (e.g. Arbaugh & Brown, 2005; Boston, 2013; Boston & Smith, 2011; Stein et al., 1996). Thus, providing in-service and preservice teachers with such tools to support and encourage reflections on task demands and the types of mathematical thinking and understanding they can promote could enhance teachers' mathematical-task knowledge. In a Norwegian context, such a tool should involve the competencies and new core elements included in the curriculum, and thus, it could draw on and further extend the MEG scheme by involving other competencies such as Aids and tools. Revisions and renewals of tools and frameworks for analysing tasks are also important for research purposes as, for instance, analyses of tasks as a means for investigating the learning potential provided by tasks or teachers' mathematical-task knowledge need to be based on the notion of mathematical competence that is relevant to the mathematics education in which the research is situated.

A possible further step toward a competency-oriented mathematics education relates to assessment practices and the grade 10 examination. White Paper no. 28 (Stortingsmelding (White Paper) nr. 28, 2015–2016) describes the need for an inspection of exam tasks and the possible need to start with pilot testing of exams. Both the findings in this thesis and the concerns of a range of authors regarding the validity of tests (e.g. Koeppen et al., 2008; Lane, 2004; Niss, 2007) strongly support such a practice, as it is regarded as highly challenging to develop tests and assessments that are able to capture the types of high-level thinking skills and complex abilities involved in the core elements that are to be implemented in the Norwegian curriculum. In such a development process, both the type of task analysis and the psychometric modelling approach involved in my thesis would serve as useful tools for test developers to provide additional information about the extent to which the competencies or other cognitive activities of interest are captured adequately by the assessments. In addition, as what is measured in national assessments is seen to shape teachers' classroom practices (Boesen et al., 2014; de Lange, 2007), ensuring that the assessment tasks involve a full spectrum of competencies could also affect the mathematical tasks teachers use in their practices and the competencies they emphasise when considering task demands.

As pointed out by Stein et al. (1996), tasks may change as they move through different phases. For instance, the demands and features of a task as it appears in instructional material might change when the task is implemented in the classroom. This thesis has been concerned with tasks as they are presented in instructional and educational material and, from this perspective, contributed knowledge about some aspects of the competency demands of tasks used in Norwegian mathematics education and teachers' knowledge of these task demands. Nonetheless, more research is needed on teachers' implementation and students' engagement in tasks to further clarify the extent to which the use of tasks in mathematics classrooms provides opportunities for developing a range of competencies and to determine whether teachers' practices successfully maintain or enhance the competency demands and optimise the learning potential of tasks. Such knowledge is vital to ensure that competency-oriented mathematics education not only exists in curriculum documents but also pervades mathematics teaching and learning activities in schools.

References

Ainley, J., Bills, L., & Wilson, K. (2005). Designing spreadsheet-based tasks for purposeful algebra. *International Journal of Computers for Mathematical Learning, 10*(3), 191–215.

Anthony, G., & Walshaw, M. (2007). *Effective pedagogy in mathematics/Pangarau: Best evidence synthesis*. Wellington: New Zealand Ministry of Education.

Anthony, G., & Walshaw, M. (2009). Characteristics of effective teaching of mathematics: A view from the West. *Journal of Mathematics Education, 2*(2), 147–164.

Apple, M. W. (1992). Do the standards go far enough? Power, policy, and practice in mathematics education. *Journal for Research in Mathematics Education, 23*(5), 412–431.

Arbaugh, F., & Brown, C. A. (2005). Analyzing mathematical tasks: A catalyst for change? *Journal of Mathematics Teacher Education, 8*(6), 499–536.

Baker, D., Knipe, H., Collins, J., Leon, J., Cummings, E., Blair, C., & Gamson, D. (2010). One hundred years of elementary school mathematics in the United States: A content analysis and cognitive assessment of textbooks from 1900 to 2000. *Journal for Research in Mathematics Education, 41*(4), 383–423.

Ball, D. (2000). Foreword. In M. K. Stein, M. S. Smith, M. Henningsen, & E. A. Silver (Eds.), *Implementing standards-based mathematics instruction: A casebook for professional development* (pp. ix–xiv). New York: Teachers College Press.

Ball, D., Thames, M., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 25–48). New York, NY: Springer.

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

Bergem, O. K. (2016). "Usually we are not where the teacher is". In K. Klette, O. K. Bergem, & A. Roe (Eds.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (pp. 47–61). Cham: Springer International Publishing.

Blum, W. (1993). Mathematical modelling in mathematics education and instruction. In T. Breiteig, I. Huntley, & G. Kaiser-Messmer (Eds.), *Teaching and learning mathematics in context* (pp. 3–14). New York, USA: Ellis Horwood.

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–Intellectual and attitudinal challenges* (pp. 73–96): Springer.

Blum, W., Artigue, M., Mariotti, M. A., Sträßer, R., & Van den Heuvel-Panhuizen, M. (2017). European didactic traditions in mathematics: Aspects and examples from four selected cases. In G. Kaiser (Ed.), *Proceedings of the 13th International Congress on Mathematical Education* (pp. 291–303): Springer.

Blum, W., & Ferri, R. B. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application, 1*(1), 45–58.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects — State, trends and issues in mathematics instruction. *Educational Studies in Mathematics, 22*(1), 37–68.

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Viewing competence as a continuum. *Zeitschrift für Psychologie, 223*(1), 3–13.
Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *Journal of Mathematical Behavior, 33*, 72–87.
Boesen, J., Lithner, J., & Palm, T. (2018). Assessing mathematical competencies: An analysis of Swedish national mathematics tests. *Scandinavian Journal of Educational Research, 62*(1), 109–124.
Borasi, R. (1986). On the nature of problems. *Educational Studies in Mathematics, 17*(2), 125–141.
Boston, M. D. (2013). Connecting changes in secondary mathematics teachers' knowledge to their experiences in a professional development workshop. *Journal of Mathematics Teacher Education, 16*(1), 7–31.
Boston, M. D., & Smith, M. S. (2009). Transforming secondary mathematics teaching: Increasing the cognitive demands of instructional tasks used in teachers' classrooms. *Journal for Research in Mathematics Education, 40*(2), 119–156.
Boston, M. D., & Smith, M. S. (2011). A 'task-centric approach' to professional development: Enhancing and sustaining mathematics teachers' ability to implement cognitively challenging mathematical tasks. *ZDM, 43*(6-7), 965–977.
Brewer, J., & Hunter, A. (2006). *Foundations of multimethod research*. Thousand Oaks, CA: SAGE Publications.
Brodie, K., Jina, Z., & Modau, S. (2009). Challenges in implementing the new mathematics curriculum in Grade 10: A case study. *African journal of research in mathematics, science and technology education, 13*(1), 19–32.
Bryman, A. (2016). *Social research methods* (5th ed.). Oxford: Oxford University Press.
Brändström, A. (2005). *Differentiated tasks in mathematics textbooks: An analysis of the levels of difficulty.* (Doctoral dissertation), Luleå tekniska universitet, Luleå. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-18110
Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal, 10*(1 & 2), 15–42.
Cai, J., & Howson, G. (2013). Toward an international mathematics curriculum. In M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 949–974). New York, NY: Springer New York.
Chapman, O. (2013). Mathematical-task knowledge for teaching. *Journal of Mathematics Teacher Education, 16*(1), 1–6.
Charalambous, C. Y. (2008). *Mathematical knowledge for teaching and the unfolding of tasks in mathematics lessons: Integrating two lines of research.* Paper presented at the Proceedings of the 32nd Conference of the International Group for the Psychology of Mathematics Education.
Charalambous, C. Y., Delaney, S., Hsu, H.-Y., & Mesa, V. (2010). A comparative analysis of the addition and subtraction of fractions in textbooks from three countries. *Mathematical Thinking and Learning, 12*(2), 117–151.
Charalambous, C. Y., & Philippou, G. N. (2010). Teachers' concerns and efficacy beliefs about implementing a mathematics curriculum reform: Integrating two lines of inquiry. *Educational Studies in Mathematics, 75*(1), 1–21.
Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. *Educational Researcher, 21*(6), 13–17.

Clarke, D., Goos, M., & Morony, W. (2007). Problem solving and working mathematically: An Australian perspective. *The International Journal on Mathematics Education, 39*(5), 475–490.
Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed. ed.). Los Angeles: Sage.
D'Ambrosio, U. (1994). Cultural framing of mathematics teaching and learning. In R. Biehler, R. W. Scholz, R. Sträßer, & B. Winkelmann (Eds.), *Didactics of mathematics as a scientific discipline* (pp. 443–455). Dordrecht: Kluwer Academic Publishers.
De Boeck, P., Cho, S. J., & Wilson, M. (2016). Explanatory item response models. In A. A. Rupp & J. P. Leighton (Eds.), *The handbook of cognition and assessment: Frameworks, methodologies, and applications* (pp. 249–268). Hoboken, NJ, USA: John Wiley & Sons, Inc.
de Lange, J. (1995). Assessment: No change without problems. In T. A. Romberg (Ed.), *Reform in school mathematics and authentic assessment* (pp. 87–172). Albany, NY: State University of New York Press.
de Lange, J. (2007). Large-scale assessment and mathematics education. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1111–1142). Charlotte, NC: Information Age.
Dole, S., & Shield, M. (2008). The capacity of two Australian eighth-grade textbooks for promoting proportional reasoning. *Research in Mathematics Education, 10*(1), 19–35.
Doyle, W. (1988). Work in mathematics classes: The context of students' thinking during instruction. *Educational Psychologist, 23*(2), 167–180.
Embretson, S. E., & Daniel, R. C. (2008). Understanding and quantifying cognitive complexity level in mathematical problem solving items. *Psychology Science, 50*(3), 328–344.
Embretson, S. E., & Gorin, J. S. (2001). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement, 38*(4), 343–368.
Enright, M. K., Morley, M., & Sheehan, K. M. (2002). Items by design: The impact of systematic feature variation on item statistical characteristics. *Applied Measurement in Education, 15*(1), 49–74.
Ernest, P. (1989). The knowledge, beliefs and attitudes of the mathematics teacher: A model. *Journal of Education for Teaching, 15*(1), 13–33.
Ernest, P. (1991). *The philosophy of mathematics education*. London: Falmer Press.
Ernest, P. (1994). The dialogical nature of mathematics. In P. Ernest (Ed.), *Mathematics, education and philosophy: An international perspective* (pp. 33–48). London: The Falmer Press.
Ernest, P. (2010). Reflections on theories of learning. In B. Sriraman & L. D. English (Eds.), *Theories of mathematics education: Seeking new frontiers*. Heidelberg: Springer.
Feeley, T. H. (2002). Comment on halo effects in rating and evaluation research. *Human Communication Research, 28*(4), 578–586. doi:10.1111/j.1468-2958.2002.tb00825.x
Flick, U. (2014). Mapping the field. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 3–18). London: SAGE Publications Ltd.
Freedle, R., & Kostin, I. (1993). The prediction of TOEFL reading item difficulty: Implications for construct validity. *Language Testing, 10*(2), 133–170.
Glaser, R., & Silver, E. A. (1994). Assessment, testing, and instruction: Retrospect and prospect. *Review of Research in Education, 20*, 393–419.
Goldin, G. A., & McClintock, C. E. (1979). *Task variables in mathematical problem solving*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.

Gorin, J. S., & Embretson, S. E. (2006). Item difficulty modeling of paragraph comprehension items. *Applied Psychological Measurement, 30*(5), 394–411.
Graf, E. A., Peterson, S., Steffen, M., & Lawless, R. (2005). Psychometric and cognitive analysis as a basis for the design and revision of quantitative item models. *ETS Research Report Series, 2005*(2).
Haapasalo, L., & Kadijevich, D. (2000). Two types of mathematical knowledge and their relation. *Journal für mathematik-didaktik, 21*(2), 139–157.
Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New York, NY: Routledge.
Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 524–549.
Hesse-Biber, S. N., & Johnson, R. B. (2015). *The Oxford handbook of multimethod and mixed methods research inquiry*: Oxford University Press.
Hiebert, J. (1986). *Conceptual and procedural knowledge : the case of mathematics*. Hillsdale, N.J: Erlbaum.
Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., . . . Stigler, J. W. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: National Center for Education Statistics.
Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge. In *Conceptual and procedural knowledge : The case of mathematics* (pp. 1–27). Hillsdale, N.J: Erlbaum.
Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal, 30*(2), 393–425.
Hildebrand, D. (2013). Dewey's pragmatism: instrumentalism and meliorism. In A. R. Malachowski (Ed.), *The Cambridge companion to pragmatism* (pp. 55–80): Cambridge University Press.
Hohensinn, C., & Kubinger, K. D. (2009). On varying item difficulty by changing the response format for a mathematical competence test. *Austrian Journal of Statistics, 38*(4), 231–239.
Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288.
Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.
Jones, D. L., & Tarr, J. E. (2007). An examination of the levels of cognitive demand required by probability tasks in middle grades mathematics textbooks. *Statistics Education Research Journal, 6*(2), 4–27.
Kaur, B. (2010). Mathematical tasks from Singapore classrooms. In Y. Shimizu, B. Kaur, R. Huang, & D. Clarke (Eds.), *Mathematical tasks in classrooms around the world* (pp. 15–33). Rotterdam, Netherlands: Sense Publishers.
Kilpatrick, J. (2014a). Competency frameworks in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 85–87). Dordrecht, Netherlands: Springer.
Kilpatrick, J. (2014b). History of research in mathematics education. In *Encyclopedia of mathematics education* (pp. 267–272): Springer.
Kleven, T. A. (2008). Validity and validation in qualitative and quantitative research. *Nordic Studies in Education, 28*(03), 219–232.
Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 3–22). Göttingen: Hogrefe Publishing.

Koeppen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current issues in competence modeling and assessment. *Zeitschrift für Psychologie, 216*(2), 61–73.
Kolovou, A., van den Heuvel-Panhuizen, M., & Bakker, A. (2011). Non-routine problem solving tasks in primary school mathematics textbooks–a needle in a haystack. *Mediterranean Journal for Research in Mathematics Education, 8*(2), 29–66.
Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, *100*(3), 716–725.
Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge, England: Cambridge University Press.
Lane, S. (2004). Validity of High-Stakes Assessment: Are Students Engaged in Complex Thinking? *Educational Measurement: Issues and Practice, 23*(3), 6–14.
Lesh, R. A., & Zawojewski, J. (2007). Problem solving and modeling. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 763–804). Charlotte, NC: Information Age.
Lewis, J. (2009). Redefining qualitative methods: Believability in the fifth moment. *International Journal of Qualitative Methods, 8*(2), 1–14.
Li, Y. (2000). A comparison of problems that follow selected content presentations in American and Chinese mathematics textbooks. *Journal for Research in Mathematics Education, 31*(2), 234–241.
Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2005). Paradigms and perspectives in contention. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (pp. 183–190).
Lingefjärd, T. (2006). Faces of mathematical modeling. *ZDM, 38*(2), 96–112.
Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *Journal of Mathematical Behavior, 23*(4), 405–427.
Lithner, J., Bergqvist, E., Bergqvist, T., Boesen, J., Palm, T., & Palmberg, B. (2010). *Mathematical competencies—A research framework*. Paper presented at the MADIF 7 Mathematics and mathematics education: Cultural and social dimensions, Stockholm, Sweden.
Mark, M. M. (2015). Mixed and multimethods in predominantly quantitative studies, especially experiments and quasi-experiments. In S. N. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford handbook of multimethod and mixed methods research inquiry* (pp. 21–41): Oxford University Press.
Mayer, R. E., & Hegarty, M. (1996). The process of understanding mathematical problems. *The Nature of Mathematical Thinking*, *12*, 24–59.
Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution.* Klagenfurt, Austria. Retrieved from http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173
Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741–749.
NESH. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi [Guidelines for research ethics in social-sciences, humaniora, law and theology]*. Retrieved from https://www.etikkom.no/globalassets/documents/publikasjoner-som-pdf/60125_fek_retningslinjer_nesh_digital.pdf
Neubrand, J. (2006). The TIMSS 1995 and 1999 video studies. In F. K. S. Leung, K. D. Graf, & F. J. Lopez-Real (Eds.), *Mathematics education in different cultural traditions: A comparative study of East Asia and the West* (pp. 291–318): Springer.

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, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 125–144). New York: Springer.

Newton, P., & Shaw, S. (2014). *Validity in educational and psychological assessment*. London: SAGE Publications.

Nilsen, T., Angell, C., & Grønmo, L. S. (2013). Mathematical competencies and the role of mathematics in physics education: A trend analysis of TIMSS Advanced 1995 and 2008. *Acta Didactica Norge, 7*(1), 1–21.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250–256.

Niss, M. (2007). Reflections on the state of and trends in research on mathematics teaching and learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1293–1312). Charlotte, NC: Information Age.

Niss, M. (2015). Mathematical competencies and PISA. In K. Stacey & R. Turner (Eds.), *Assessing mathematical literacy* (pp. 35–55): Springer.

Niss, M., Bruder, R., Planas, N., Turner, R., & Villa-Ochoa, J. A. (2016). Survey team on: conceptualisation of the role of competencies, knowing and knowledge in mathematics education research. *ZDM, 48*(5), 611–632.

Niss, M., & Højgaard, T. (2011). *Competencies and mathematical learning. Ideas and inspiration for the development of mathematics teaching and learning in Denmark (English edition).* Roskilde: IMFUMFA.

Niss, M., & Jablonka, E. (2014). Mathematical literacy. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 391–396). Dordrecht, Netherlands: Springer.

Niss, M., & Jensen, T. H. (2002). *Kompetencer og matematiklæring – Idéer og inspiration til udvikling af matematikundervisning i Danmark* (Vol. 18). Copenhagen, Denmark: The Ministry of Education.

Norwegian Directorate for Education and Training [Utdanningsdirektoratet]. (n.d.). Curriculum for the common core subject of mathematics. Retrieved from https://www.udir.no/kl06/MAT1-04?lplang=http://data.udir.no/kl06/eng

OECD. (2014). *PISA 2012 technical report*. Retrieved from Paris: https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf

Osana, H. P., Lacroix, G. L., Tucker, B. J., & Desrosiers, C. (2006). The role of content knowledge and problem features on preservice teachers' appraisal of elementary mathematics tasks. *Journal of Mathematics Teacher Education, 9*(4), 347–380.

Palm, T., Boesen, J., & Lithner, J. (2011). Mathematical reasoning requirements in Swedish upper secondary level assessments. *Mathematical Thinking and Learning, 13*(3), 221–246.

Pedersen, I. F. (2014). *Insights from TIMSS advanced on critical aspects of the advanced mathematics program in Norwegian upper secondary school: Content, competence, and motivation.* Department of Teacher Education, Faculty of Educational Sciences, University of Oslo, Oslo. (no. 213)

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Charlotte, NC: Information Age.

Phillips, D. C., & Burbules, N. C. (2000). *Postpositivism and educational research*. Lanham, Maryland: Rowman & Littlefield Publishers.

Pikkarainen, E. (2014). Competence as a key concept of educational theory: A semiotic point of view. *Journal of Philosophy of Education, 48*(4), 621–636.

Richardson, K., Carter, T., & Berenson, S. (2010). Connected tasks: The building blocks of reasoning and proof. *Australian Primary Mathematics Classroom, 15*(4), 17–23.
Schoenfeld, A. H. (1987). *Cognitive science and mathematics education*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). New York, N: Macmillan.
Schoenfeld, A. H. (2008). Research methods in (mathematics) education. In L. D. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (pp. 481–533).
Schoenfeld, A. H. (2016). Research in mathematics education. *Review of Research in Education, 40*(1), 497–528.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
Shimizu, Y., Kaur, B., Huang, R., & Clarke, D. (2010). The role of mathematical tasks in different cultures. In Y. Shimizu, B. Kaur, R. Huang, & D. Clarke (Eds.), *Mathematical tasks in classrooms around the world* (pp. 1–14). Rotterdam, Netherlands: Sense Publishers.
Sierpinska, A. (2004). Research in mathematics education through a keyhole: Task problematization. *For the Learning of Mathematics*, *24*(2), 7–15.
Silver, E. A., Mesa, V. M., Morris, K. A., Star, J. R., & Benken, B. M. (2009). Teaching mathematics for understanding: An analysis of lessons submitted by teachers seeking NBPTS certification. *American Educational Research Journal, 46*(2), 501–531.
Silverman, D. (2011). *Interpreting qualitative data: A guide to the principles of qualitative research* (4th ed. ed.). Los Angeles, Calif: SAGE.
Smith, M. S., & Stein, M. K. (1998). Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3*(5), 344–350.
Sproesser, U., Vogel, M., Dörfler, T., & Eichler, A. (2018). *Teachers' judgement accuracy of task difficulty related to functions.* Paper presented at the Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education, Umeå, Sweden.
Sriraman, B., & English, L. D. (2010). Surveying theories and philosophies of mathematics education. In B. Sriraman & L. English (Eds.), *Theories of mathematics education: Seeking new frontiers*. Heidelberg: Springer.
Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455–488.
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.
Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3*(4), 268–275.
Stein, M. K., Smith, M. S., Henningsen, M., & Silver, E. A. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development*. New York: Teachers College Press.
Stigler, J. W., & Hiebert, J. (2004). Improving mathematics teaching. *Educational Leadership, 61*(5), 12–17.
Stortingsmelding (White Paper) nr. 28. (2015–2016). *Fag – Fordypning – Forståelse — En fornyelse av Kunnskapsløftet [Subjects – Depth – Understanding — A renewal of The Knowledge Promotion]*. Oslo: Ministry of Education and Research,.

Sullivan, P., Clarke, D., & Clarke, B. (2013). *Teaching with tasks for effective mathematics learning* (Vol. 9). New York: Springer Science & Business Media.

Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., . . . Vos, P. (2016). *Assessment in mathematics education: Large-Scale assessment and classroom assessment*. Switzerland: Springer Open.

Tatto, M. T., Schwille, J., Ingvarson, L., Rowley, G., Bankov, K., Peck, R., . . . Senk, S. L. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).

Tietze, U.-P. (1994). Curricula and goals. In R. Biehler, R. W. Scholz, R. Sträßer, & B. Winkelmann (Eds.), *Didactics of mathematics as a scientific discipline* (pp. 41–53). Dordrecht: Kluwer Academic Publishers.

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

Turner, R., Dossey, J., Blum, W., & Niss, M. (2013). Using mathematical competencies to predict item difficulty in PISA: A MEG study. In M. Prenzel, M. Kobarg, K. Schöps, & S. Rönnebeck (Eds.), *Research on PISA* (pp. 23–37). New York: Springer.

Valenta, A., Nosrati, M., & Wæge, K. (2015). *Skisse av den «ideelle læreplan i matematikk» [Draft of the «ideal curriculum in mathematics»]*. Trondheim: Nasjonalt senter for matematikk i opplæringen Retrieved from https://docplayer.me/1644531-Skisse-av-den-ideelle-laereplan-i-matematikk.html.

Westera, W. (2001). Competences in education: A confusion of tongues. *Journal of curriculum studies, 33*(1), 75–88.

Wiliam, D. (2007). Keeping learning on track. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1053–1098). Charlotte, NC: Information Age.

Wilson, M., De Boeck, P., & Carstensen, C. H. (2008). Explanatory item response models: A brief introduction. In E. Klieme & D. Leutner (Eds.), *Assessment of competencies in educational contexts: State of the art and future prospects* (pp. 91–120). Göttingen: Hogrefe & Huber.

Yeo, J. B. (2017). Development of a framework to characterise the openness of mathematical tasks. *International Journal of Science and Mathematics Education, 15*(1), 175–191.

Appendix 1 – MEG item analysis scheme[6]

Competency definitions and level descriptions

Communication

The communication competency has both ‘receptive’ and ‘constructive’ components. The receptive component includes understanding what is being stated and shown related to the mathematical objectives of the task, including the mathematical language used, what information is relevant, and what is the nature of the response requested. The constructive component consists of presenting the response that may include solution steps, description of the reasoning used and justification of the answer provided.

In written and computer-based items, receptive communication relates to understanding text and images, still and moving. Text includes verbally presented mathematical expressions and may also be found in mathematical representations (for example titles, labels and legends in graphs and diagrams).

Communication does not include knowing how to approach or solve the problem, how to make use of particular information provided, or how to reason about or justify the answer obtained, rather it is the understanding or presenting of relevant information. It also does not apply to extracting or processing mathematical information from representations. In computer-based items, the instructions about navigation and other issues related to the computer environment may add to the general task demand, but is not part of the communication competency.

Demand for the receptive aspect of this competency increases according to the complexity of material to be interpreted in understanding the task; the need to link multiple information sources or to move backwards and forwards (to cycle) between information elements. The constructive aspect increases with the need to provide a detailed written solution or explanation.

Definition: Reading and ***interpreting*** statements, questions, instructions, tasks, images and objects; ***imagining*** and ***understanding*** the situation presented and ***making sense*** of the information provided including the mathematical terms referred to; ***presenting*** and ***explaining*** one’s mathematical work or reasoning.

Level 0: Understand short sentences or phrases relating to concepts that give immediate access to the context, where all information is directly relevant to the task, and where the order of information matches the steps of thought required to understand what the task requests. Constructive communication involves only presentation of a single word or numeric result.

Level 1: Identify and link relevant elements of the information provided in the text and other related representation/s, where the material presented is more complex or extensive than short sentences and phrases or where some extraneous information may be present. Any constructive communication required is simple, for example it may involve writing a short statement or calculation, or expressing an interval or a range of values.

Level 2: Identify and select elements to be linked, where repeated cycling within the material presented is needed to understand the task; or understand multiple elements of the context or task or their links. Any constructive communication involves providing a brief description or explanation, or presenting a sequence of calculation steps.

[6] Reprinted from Turner et al. (2015) with permission of Springer International Publishing, copyright 2015.

Level 3: Identify, select and understand multiple context or task elements and links between them, involving logically complex relations (such as conditional or nested statements). Any constructive communication would involve presenting argumentation that links multiple elements of the problem or solution.

Devising strategies

The focus of this competency is on the strategic aspects of mathematical problem solving: selecting, constructing or activating a solution strategy and monitoring and controlling the implementation of the processes involved. 'Strategy' is used to mean a set of stages that together form the overall plan needed to solve the problem. Each stage comprises a sub-goal and related steps. For example a plan to gather data, to transform them and to represent them in a different way would normally constitute three separate stages.

The knowledge, technical procedures, mathematising and reasoning needed to actually carry out the solution process are taken to belong to those other competencies.

Demand for this competency increases with the degree of creativity and invention involved in identifying a suitable strategy, with increased complexity of the solution process (for example the number, range and complexity of the stages needed in a strategy), and with the consequential need for greater metacognitive control in the implementation of the strategy towards a solution.

Definition: ***Selecting*** or ***devising*** a mathematical strategy to solve a problem as well as ***monitoring*** and ***controlling*** implementation of the strategy.

Level 0: Take direct actions, where the solution process needed is explicitly stated or obvious.

Level 1: Find a straight-forward strategy (usually of a single stage) to combine or use the given information.

Level 2: Devise a straight-forward multi-stage strategy, for example involving a linear sequence of stages, or repeatedly use an identified strategy that requires targeted and controlled processing.

Level 3: Devise a complex multi-stage strategy, for example that involves bringing together multiple sub-goals or where using the strategy involves substantial monitoring and control of the solution process; or evaluate or compare strategies.

Mathematising

The focus of this competency is on those aspects of the modelling cycle that link an extra-mathematical context with some mathematical domain. Accordingly, the mathematising *competency has two components. A situation outside mathematics may require translation into a form amenable to mathematical treatment. This includes making simplifying assumptions, identifying variables present in the context and relationships between them, and expressing those variables in a mathematical form. This translation is sometimes referred to as mathematising. Conversely, a mathematical entity or outcome may need to be interpreted in relation to an extra-mathematical situation or context. This includes translating mathematical results in relation to specific elements of the context and validating the adequacy of the solution found with respect to the context. This process is sometimes referred to as de-mathematising.*

The intra-mathematical treatment of ensuing issues and problems within the mathematical domain is dealt with under other competencies. Hence, while the mathematising *competency deals with representing extra-mathematical contexts by means of mathematical entities, the representation of mathematical entities is dealt with under the* representation *competency.*

Demand for activation of this competency increases with the degree of creativity, insight and knowledge needed to translate between the context elements and the mathematical structures of the problem.

Definition: ***Translating*** an extra-mathematical situation into a mathematical model, ***interpreting*** outcomes from using a model in relation to the problem situation, or ***validating*** the adequacy of the model in relation to the problem situation.

Level 0: Either the situation is purely intra-mathematical, or the relationship between the extra-mathematical situation and the model is not relevant to solving the problem.

Level 1: Construct a model where the required assumptions, variables, relationships and constraints are given; or draw conclusions about the situation directly from a given model or from the mathematical results.

Level 2: Construct a model where the required assumptions, variables, relationships and constraints can be readily identified; or modify a given model to satisfy changed conditions; or interpret a model or mathematical results where consideration of the problem situation is essential.

Level 3: Construct a model in a situation where the assumptions, variables, relationships and constraints need to be defined; or validate or evaluate models in relation to the problem situation; or link or compare different models.

Representation

The focus of this competency is on decoding, devising, and manipulating representations of mathematical entities or linking different representations in order to pursue a solution. By 'representation of a mathematical entity' we understand a concrete expression (mapping) of a mathematical concept, object, relationship, process or action. It can be physical, verbal, symbolic, graphical, tabular, diagrammatic or figurative.

Mathematical tasks are often presented in text form, sometimes with graphic material that only helps set the context. Understanding verbal or text instructions and information, photographs and graphics does not generally belong to representation competency – that is part of the communication *competency. Similarly, working exclusively with symbolic representations lies within the* using symbols, operations and formal language *competency. On the other hand, translation between different representations is always part of the* representation *competency. For example, the act of transforming mathematical information derived from relevant text elements into a non-verbal representation is where* representation *commences to apply.*

While the representation *competency deals with representing* mathematical *entities by means of other entities (mathematical or extra-mathematical), the representation of extra-mathematical contexts by mathematical entities is dealt with under the* mathematising *competency.*

Demand for this competency increases with the amount of information to be extracted, with the need to integrate information from multiple representations, and with the need to devise representations rather than to use given representations. Demand also increases with added complexity of the representation or of its decoding, from simple and standard representations requiring minimal decoding (such as a bar chart or Cartesian graph), to complex and less standard representations comprising multiple components and requiring substantial decoding perhaps devised for specialised purposes (such as a population pyramid, or side elevations of a building).

Definition: ***Decoding***, ***translating*** between, and ***making use*** of given mathematical representations in pursuit of a solution; ***selecting*** or ***devising*** representations to capture the situation or to present one's work.

Level 0: Either no representation is involved, or read isolated values from a simple representation, for example from a coordinate system, table or bar chart; or plot such values; or read isolated numeric values directly from text.

Level 1: Use a given simple and standard representation to interpret relationships or trends, for example extract data from a table to compare values, or interpret changes over time shown in a graph; or read or plot isolated values within a complex representation; or construct a simple representation.

Level 2: Understand and use a complex representation, or construct such a representation where some of the required structure is provided; or translate between and use different simple representations of a mathematical entity, including modifying a representation.

Level 3: Understand, use, link or translate between multiple complex representations of mathematical entities; or compare or evaluate representations; or devise a representation that captures a complex mathematical entity.

Using symbols, operations and formal language

This competency reflects skill with activating and using mathematical content knowledge, such as mathematical definitions, results (facts), rules, algorithms and procedures, recalling and using symbolic expressions, understanding and manipulating formulae or functional relationships or other algebraic expressions and using the formal rules of operations (e.g. arithmetic calculations or solving equations). This competency also includes working with measurement units and derived quantities such as 'speed' and 'density'.

Developing symbolic formulations of extra-mathematical situations is part of mathematisation. *For example, setting up an equation to reflect the key elements of an extra-mathematical situation belongs to* mathematisation, *whereas solving it is part of the* using symbols, operations and formal language *competency. Manipulating symbolic expressions belongs to the* using symbols, operations and formal language *competency even though they are mathematical representations. However, translating between symbolic and other representations belongs to the* representation *competency.*

The term 'variable' is used here to refer to a symbol that stands for an unspecified number or a changing quantity, for example C and r in the formula $C = 2\pi r$.

Demand for this competency increases with the increased complexity and sophistication of the mathematical content and procedural knowledge required.

Definition: Understanding and ***implementing*** mathematical procedures and language (including symbolic expressions, arithmetic and algebraic operations), using the mathematical ***conventions*** and ***rules*** that govern them; ***activating*** and ***using knowledge*** of definitions, results, rules and ***formal systems***.

Level 0: State and use elementary mathematical facts and definitions; or carry out short arithmetic calculations involving only easily tractable numbers. For example, find the area of a rectangle given the side lengths, or write down the formula for the area of a rectangle.

Level 1: Make direct use of a simple mathematical relationship involving variables (for example, substitute into a linear relationship); use arithmetic calculations involving fractions and decimals; use repeated or sustained calculations from level 0; make use of a mathematical definition, fact, or convention, for example use knowledge of the angle sum of a triangle to find a missing angle.

Level 2: Use and manipulate expressions involving variables and having multiple components (for example, by algebraically rearranging a formula); employ multiple rules, definitions, results, conventions, procedures or formulae together; use repeated or sustained calculations from level 1.

Level 3: Apply multi-step formal mathematical procedures combining a variety of rules, facts, definitions and techniques; work flexibly with complex relationships involving variables, for example use insight to decide which form of algebraic expression would be better for a particular purpose.

Reasoning and argument

This competency relates to drawing valid inferences based on the internal mental processing of mathematical information needed to obtain well-founded results, and to assembling those inferences to justify or, more rigorously, prove a result.

Other forms of mental processing and reflection involved in undertaking tasks underpin each of the other competencies. For example the thinking needed to choose or devise an approach to solving a problem is dealt with under the devising strategies *competency, and the thinking involved in transforming contextual elements into a mathematical form is accounted for in the* mathematising *competency.*

The nature, number or complexity of elements that need to be brought to bear in making inferences, and the length and complexity of the chain of inferences needed would be important contributors to increased demand for this competency.

<u>Definition:</u> ***Drawing inferences*** by using logically rooted thought processes that explore and connect problem elements to ***form, scrutinise*** or ***justify arguments*** and conclusions.

<u>Level 0:</u> Draw direct inferences from the information and instructions given.

<u>Level 1:</u> Draw inferences from reasoning steps within one aspect of the problem that involves simple mathematical entities.

<u>Level 2:</u> Draw inferences by joining pieces of information from separate aspects of the problem or concerning complex entities within the problem; or make a chain of inferences to follow or create a multi-step argument.

<u>Level 3:</u> Use or create linked chains of inferences; or check or justify complex inferences; or synthesise and evaluate conclusions and inferences, drawing on and combining multiple elements of complex information, in a sustained and directed way.

Appendix 2 – Task questionnaire[7]

Name or number of task: ______________________________

1) At what grade level did you use this task? ________________

2) Which students worked on this task?

☐ High-achieving students only ☐ Students at different achievement level

3a) In which learning situation was the task used?

☐Instruction ☐Homework ☐Assessment/test

3b) If the task was used for instructional purposes, how did the students work with the task?

☐Individually ☐Pairs/Groups ☐Whole-class

4) What is the source of the task?

☐Textbook ☐Internet ☐Colleague ☐Self-made ☐Other

5) What do you think makes this a demanding task for high-achieving students?

[7] The questionnaire has been translated from Norwegian

Appendix 3 – Consent form

Studien gjennomføres i samsvar med retningslinjer utarbeidet av Personvernombudet for forskning, NSD. Innsamlingen, oppbevaringen og rapportering av data skjer i tråd med disse retningslinjene. Det innsamlede materialet vil bli fullstendig anonymisert og det vil ikke være mulig å knytte resultater og funn fra studien til enkeltlærere eller skoler.

- Jeg erklærer at samtykke er gitt frivillig og at jeg er informert om hva det skal brukes til
- Jeg er informert om at samtykke når som helst kan trekkes tilbake

☐ Jeg samtykker til å dele oppgavene jeg sender inn med de andre lærerne som deltar i studien.

__________________ ____________________________________

Dato Underskrift

____

Navn med blokkbokstaver

Part II

The Articles

Article 1

Pettersen, A., & Nortvedt, G. A. (2018). Identifying competency demands in mathematical tasks: Recognising what matters. *International Journal of Science and Mathematics Education*, *16*(5), pp. 949–965.

1

Article 2

Pettersen, A., & Braeken, J. (2017). Mathematical competency demands of assessment items: A search for empirical evidence. *International Journal of Science and Mathematics Education*. doi:10.1007/s10763-017-9870-y

2

Article 3

Pettersen, A., & Nortvedt, A. G. (under review, 2018). Teachers' considerations of mathematical tasks used to challenge high-achieving students. *Scandinavian Journal of Educational Research*

3