

Principles for designing mathematical tasks that enhance imitative and creative reasoning

Johan Lithner¹ 

Accepted: 13 May 2017 / Published online: 18 May 2017
© The Author(s) 2017. This article is an open access publication

Abstract The design research programme learning by imitative and creative reasoning (LICR) studies whether, how and why tasks and teaching that enhance creative reasoning lead to a more productive struggle and more efficient learning than the common but inefficient task designs based on imitating given solution procedures. The purpose of this paper is to synthesise the research outcomes determined to date by providing the following: a conceptual framework for key concepts and relationships among teaching, tasks, student activities and learning; a theoretical basis for analyses of causal effects between task/teaching design and learning outcomes; a design research methodology for transforming initial design ideas, through cycles of evaluation and revision, into firmer design principles; and an application of this theory and methodology to empirical studies carried out to date, in order to propose task-design principles related to imitative and creative reasoning.

Keywords Task design · Rote learning · Problem solving · Creative reasoning

1 Introduction

One of the most persistent challenges in mathematics education is replacing the dominant task and teaching designs, which are based on imitation of given solution methods. According to a review by Hiebert (2003), there are “massive amounts of converging data” showing that such teaching models fail to promote students’ development of central

mathematical competencies effectively and instead lead mathematics students to try to follow rote learning (i.e., by mechanical or habitual repetition) task-solution methods “like robots with poor memories” (p. 12). Superficial rote learning strategies can be a major obstacle to learning and using mathematics (Lithner 2000, 2003, 2008; Boesen et al. 2010). Hiebert has concluded that students have more opportunities to learn facts and simple procedures than to engage in more complex processes, and achievement data indicate that students are indeed learning simple facts and calculation procedures but are not learning how to find solution methods by themselves or how to engage in other mathematical processes. Similar opportunities to learn have been found in a Swedish study including observations of 200 mathematics classrooms (Boesen et al. 2014). Teaching, textbooks and assessments may promote rote learning, in the sense that algorithmic task-solution templates are provided by teachers and textbooks, and many practice and test tasks can be solved by imitating such templates (Lithner 2004; Stacey and Vincent 2009; Thompson et al. 2012; Bergqvist and Lithner 2012; Shield and Dole 2013; Boesen et al. 2014). In a study of common textbooks from Australia, Canada, Finland, India, Ireland, Nepal, Scotland, Singapore, South Africa, Sweden, Tanzania and the USA, Jäder et al. (2015) found that 79% of the tasks could be solved by imitating given procedures, 13% could be completed by mainly applying given procedures but making some minor modifications, and only 9% of the tasks required the construction of solution methods.

It is hardly reasonable to expect that students attain an in-depth understanding of all aspects of mathematics. Rote learning can reduce the demands on working memory and free up cognitive resources to be used for more advanced problem solving. In addition, rote learning and memorisation may have different roles and meanings in different

✉ Johan Lithner
johan.lithner@umu.se

¹ NMD, Umea University, 90187 Umea, Sweden

cultures. Leung (2014) noted that in East Asian cultures (including those of nations ranking high in TIMSS and PISA), there is a stress on, among other things, practice and memorisation. Stigler and Hiebert (1999) found that students in Japanese classrooms spend as much time solving challenging problems and discussing concepts as they do practicing skills. A possible conclusion is that a balance between rote learning and more creative mathematical activities may promote students' development of central mathematical competencies (Schoenfeld 1985; Hiebert 2003).

A review by Niss (2007) suggested that students need to engage in activities in which they must 'struggle' (in a productive sense) with important mathematics, but a delicate balance must be struck to prevent these struggles from becoming obstacles to rather than promoters of learning. However, in regard to proposals for more effective teaching, Hiebert and Grouws (2007) concluded in a review that, at the time of their writing, the state of education was far from providing a coherent and systematic knowledge base that documented robust links between teaching and learning outcomes. Little was known about how to translate this abstract idea of 'struggle' into the design of specific artefacts (for example, tasks) and activities useful in teaching and about the mechanisms that link such teaching to learning outcomes (Niss 2007). The productive struggle is rooted in the fact that developing central mathematical competencies (e.g., reasoning ability and conceptual understanding) requires active engagement in corresponding challenging learning processes (e.g., non-routine problem solving). There is little or no transfer to such competencies from easier learning processes, such as imitation of given solution templates (Schoenfeld 1985; Brousseau 1997; Niss 2007). Although there are important insights concerning how to provide good learning opportunities (NCTM 2000; Boaler 2002; Cobb et al. 2003; Niss 2003; Hiebert and Grouws 2007; Schoenfeld 2007, 2015; Stein et al. 2008), it is methodologically difficult to verify that the desirable learning outcomes result from teaching rather than from other variables (Niss 2007). The research programme learning by imitative and creative reasoning (LICR) seeks to add to the growing knowledge of how to actually translate this abstract idea of 'struggle' into the design of specific artefacts (e.g., tasks) and activities useful in teaching and of the mechanisms that link such teaching to learning outcomes. The focus is on the particular type of struggle

when students construct task solutions instead of imitating them.

The purpose of this paper is to synthesise the research outcomes obtained to date in the form of task-design principles by providing the following:

- a conceptual framework for key concepts and relations among teaching, tasks, student activities and learning;
- a theoretical basis for analysis of causal effects between task/teaching design and learning outcomes (cf. the Theory of Didactical Situations, Brousseau 1997);
- a structure for transforming initial design ideas, through cycles of evaluation and revision, into firmer design principles, thus providing a design research methodology (McKenney and Reeves 2012);
- an application of this theory and methodology to the empirical studies carried out to date, in order to propose task design principles related to imitative and creative reasoning.

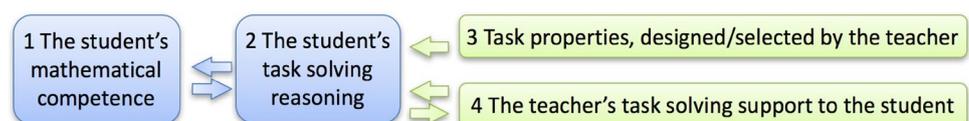
2 Conceptual framework, theory and methodology

2.1 Relating task properties, reasoning, interaction and learning

The model (Fig. 1) is inspired by, but not identical to, Stein et al. (1996) framework for relationships between mathematical tasks and learning. Its aim is to clarify the focus of LICR research, not to include all aspects of learning and teaching. Interventions and manipulations are carried out in components 3 and 4, and outcomes are measured in components 1 and 2.

1. A main aim during the past three decades of educational reform has been to help students acquire richer *mathematical competence*, i.e., the ability to understand, judge, do, and use mathematics. Basic competencies include problem solving ability (in which a problem is a challenging task in which the solver does not know a solution method in advance), reasoning ability (to justify choices and conclusions) and understanding. The internationally influential reform-oriented frameworks defining mathematical competence (NCTM 2000; Kilpatrick et al. 2001; Niss 2003) have

Fig. 1 Student components (1 and 2) and task/teacher design components (3 and 4)



also influenced Swedish official policy documents since 1994 (Boesen et al. 2014).

2. *Students' task-solving reasoning* affects the competence developed, or what is learnt from trying to solve the task. In contrast, students' existing competence affects what type of reasoning they can carry out.
3. Students' reasoning is affected by *task properties*, which are designed/selected by the teacher.
4. The teacher may interact with students to support task-solving reasoning.

2.2 The theory of didactical situations

Brousseau's (1997) theory of didactical situations in mathematics (TDS) is used as a theoretical clarification of the characteristics and consequences of rote learning and as a starting point for the design of a more constructive alternative. First, it is used to indicate why it may be attractive (and thus prevalent) in teaching to provide algorithmic solution templates: In TDS, students' temporary incomplete or faulty conceptions are not considered failures but are often inevitable and constitutive of knowledge formation processes. However, the teacher may try to overcome students' obstacles by providing task-solution templates. This relieves students of the need to take responsibility for their intellectual work, and then the struggle necessary for deeper learning will not take place.

Secondly, the theory explains why learning by imitating algorithms is ineffective. An algorithm is broadly defined to include all pre-specified task-solving methods, such as rules and template examples. An algorithm is a sequence of executable instructions for solving a class of tasks, and it can be determined in advance. The n th transition does not depend on any circumstance that was unforeseen in the $(n-1)$ th transition—it does not depend on new information, new decisions, interpretations, or thus on any meaning that could be attributed to the transitions. Therefore, the execution of an algorithm has high reliability and speed, which is a strength when the purpose is only to solve a task. However, if the purpose is to learn, an algorithm executed without considering its meaning may lead to rote learning. It is the domination of algorithmic solution templates in mathematics teaching and learning, not the algorithms themselves, that is problematic. However, algorithms are a fundamental and crucial part of mathematics. Fan and Bokhove (2014) have concluded in their literature survey that “learning of algorithms has suffered from an alleged dichotomy between procedures and understanding” (p. 481) but also that “the majority of more recent research seems to indicate that products and processes, procedures and understanding, go hand in hand” (p. 484).

Thirdly, the aim of TDS is the design of situations that allow for the construction of knowledge by the learner (as an alternative to imitation). One central aspect is the *devolution of problems*: Students must take responsibility for a part of the problem-solving process. The teacher's task is to arrange a suitable didactic situation in the form of a problem in such a way that if students solve it, then the students will obtain the desired target knowledge. From the point when the students accept the problem as their own to the moment when they produce an answer, the teacher refrains from interfering and suggesting how to solve the task. The teacher must therefore arrange the devolution of a good problem rather than describe what the students are supposed to learn. This does not imply that the teacher is more passive or has a less important role than that in the ‘solution-template providing’ approach. Designing a good problem for devolution is usually much more difficult than designing imitative tasks (Sect. 3.2), and it places higher demands on teacher interaction (Sect. 5.3).

2.3 Imitative and creative reasoning

A series of studies resulting in a research framework (Lithner 2008) have suggested that a key factor affecting learning outcomes is whether students engage in imitative or creative reasoning. Reasoning is the line of thought adopted to produce assertions and reach conclusions in task solving. It is not necessarily based on formal logic; thus, it is not restricted to proof and may even be simple, incorrect and/or superficial, as long as some sensible reasons (as perceived by the reasoner) support it.

Algorithmic reasoning (AR) consists of an attempt to solve a task by applying a given or recalled algorithm (Lithner 2008). Examples include following a memorised procedure of finding the line through two points or imitating an example given by the teacher of how to multiply two three-digit numbers and applying it to two other numbers. Another version of imitative reasoning, namely, to recall and repeat memorised non-algorithmic knowledge (such as a mathematical proof), is uncommon in school and is not treated in this paper (see Lithner 2008 for an account).

Opportunities for students to create knowledge in line with TDS have been found to be rare in teaching, textbooks and tests. When it is applied, students are able to make better progress with *Creative Mathematically founded Reasoning* (CMR, Lithner 2008). Empirical studies of the distinctions between AR and students' own constructions of solutions have defined this reasoning type as fulfilling three criteria: (1) Creativity: the learner creates a reasoning sequence not experienced previously, or re-creates a forgotten one (Silver 1997). (2) Plausibility: there are predictive arguments supporting the strategy choice and arguments for verification, explaining why the strategy implementation

and conclusions are true or plausible (Pólya 1954; Lithner 2008). (3) Anchoring: the arguments are anchored in the intrinsic mathematical properties of the components of the reasoning (Lithner 2008). A literature review has revealed two main uses of mathematical ‘creativity’ (Sriraman et al. 2013): the extraordinary creativity of geniuses and the everyday creativity that “can be fostered broadly in the general school population” (Silver 1997, p. 75). The latter meaning is used here, i.e., the creation of task solutions (or the re-creation of forgotten ones) that are original to the individual who creates them.

Solving a task using CMR is largely similar to what many others have written in terms of (non-routine) problem solving (NCTM 2000). The reason for introducing the notion is that in a large part of the literature a “problem” is equated with “any mathematical task”, including also routine exercises. In addition, when a “problem” is defined as a task for which students have no access to a solution method from the start, there are various additional requirements, for example, that a problem be challenging (Schoenfeld 1985) or require exploration (Niss 2003). A task requiring CMR does not have to meet similar criteria such as being challenging, requiring exploration or invoking modelling. Thus, the LICR programme can be more focussed on the distinction between imitation and creation instead of other types of struggle.

2.4 Methodology: educational design research

Because mathematics learning is immensely complex (Niss 2007) and the difficulties in designing and analysing interventions are underestimated (Schoenfeld 2007), a structure for this type of research is helpful. Some of the fundamental questions in design research concern how to base the design itself in relevant experiential and scientific knowledge, how to evaluate and revise the design, and how to reach conclusions in a format that is both generalisable for building scientific theoretical knowledge and concretely applicable for educational development. Methodological approaches combining these aspects are therefore advocated (Brown 1992; Cobb et al. 2003; Schoenfeld 2007). In contrast to most methodologies, the theoretical products of design experiments have potential for rapid pay-off for practice, because they are empirically evaluated principles for the development of tasks and teaching (Cobb et al. 2003). Principles guide rather than strictly determine a design, and their use in practical design requires creative input, imaginative extensions and development through feedback from trials (Swan 2008), because there is often not enough research to support detailed prescriptions (Smith and Stein 2011). Hence, in design research, there is often an “emphasis more on sensitizing the designer to crucial issues than on specifying particular courses of action”

(Ruthven et al. 2009, p. 341). By focussing closely on the imitative-creative dimension, the LICR programme aims at more specific and directive design principles.

Design propositions provide initial guidance on how to achieve the goal. During cyclic processes of design and formative evaluation, propositions are revised and transformed into research results in the form of *design principles* that are theoretical insights that recommend how to address a specific class of issues (McKenney and Reeves 2012).

3 Initiation phase

A design research project is initiated by formulating requirements and propositions (McKenney and Reeves 2012).

3.1 Design requirements in three research contexts

Design requirements specify criteria that the intervention should meet and essentially describe what the intervention will address in a particular context (ibid.).

It is clear from the start that the LICR programme will not produce final and complete principles for designing tasks and teaching optimally enhancing learning. In the best-case scenario, some substantial progress towards such a utopian goal can be made. Therefore, the aim of the research design is to be useful not only in classroom design but also for further design research. Three main types of contexts are studied, with partially different requirements, as described in the following sections.

3.1.1 Experimental research

The purpose of the experimental studies is to reduce the number of variables manipulated and measured, in order to facilitate statistical analyses of relations, particularly between task design and learning processes/outcomes. Therefore, one requirement is that the task design commonly found in teaching, textbooks and tests be able to be modelled in a reduced experimental setting. Another requirement is that CMR tasks be designed in such a way that they differ from the AR tasks only in terms of the absence of given solution methods (Figs. 2, 3 below). A third requirement is that to avoid statistical floor/ceiling effects, the tasks should not be too difficult to solve during practice and not too easy to solve during testing.

3.1.2 Explanatory research

These studies concern a basic understanding of the phenomena related to learning processes and outcomes. Here,

When squares are put in a row it looks like the figure to the right.



If x is the number of squares then the number of matches y can be calculated by the function $y=3x+1$

Example: If 4 squares are put in a row then $y=3x+1=3\cdot 4+1=13$ matches are needed.

How many matches are needed to get 50 squares in a row?

Fig. 2 Task with solution method

When squares are put in a row it looks like the figure to the right.



If x is the number of squares then the number of matches y can be calculated

Example: If 4 squares are put in a row then 13 matches are needed.

How many matches are needed to get 50 squares in a row?

Fig. 3 Task without solution method

the main requirement is to design tasks and research contexts so that these phenomena can be identified and understood.

3.1.3 Clinical research

The main requirement in clinical classroom interventions is that both the design and the research evaluation should work in the increased complexity of a real classroom. In addition, the interventions should align at least relatively well with the official curricula documents and, for ethical reasons, be expected not to have negative consequences for the students' learning.

3.2 Characteristics of tasks that enhance imitation and creation of solutions

Designing tasks that enhance algorithmic reasoning in a school context is relatively easy: first, mathematics is full

of powerful standard methods, developed over centuries, for solving many types of tasks. For example, there are arithmetic calculation algorithms, rules for determining the properties of geometrical objects and methods for solving various types of equations. It is straightforward to construct a task that requires the application of such methods. Second, textbooks are full of tasks that are solvable by copying worked examples and other types of templates. Third, because the student does not have to understand the meaning or underlying concepts of a task accompanied with an algorithmic solution template, there are relatively low requirements not only for students' creative ability but also for their conceptual understanding (Lithner 2008). Finally, if solution templates are available, students are likely to use them (Hiebert 2003; Lithner 2003, 2008; Boesen et al. 2014).

It is easy to design a task that requires CMR to be solved; one must ensure only that the student does not know the solution method in advance. However, it is more difficult to design such a task in a way that is simultaneously not too difficult for the student to solve. Designing a task that requires CMR, that is not too difficult and whose solution also leads the student to construct a particular target knowledge (see TDS) is even more challenging. According to the CMR definition above, it must be possible for students to construct arguments anchored in mathematics that support the task-solution reasoning. If students do not have access to a solution method (recalled or given) to follow, only two possibilities remain for solving the task. One is to guess, but although guesswork can be a constructive part of problem solving, it is almost never possible to solve a task only by guessing. The other possibility is to construct (part of) the solution, and this construction requires some guidance, some type of (explicit or implicit) argument to support the choices and conclusions.

A task in which a complete solution method is available (given or recalled) to a particular student from the start, is denoted an *AR task*. For example, students can solve the Eq. $3x+4=19$ if they already know a general method for solving linear equations. In this paper, a *CMR task* is a task in which (1) no complete solution method is available from the start to a particular student, and (2) it is reasonable for students to justify the construction and implementation of a solution. These two categories include all tasks of interest for school task design, because the complements are tasks without solution templates that are not possible to solve by mathematical arguments, i.e., those that are intended to be solved by pure guesswork or not at all.

3.3 Design propositions

Design propositions serve the practical goals of design research by helping to sharpen the focus of an intervention

and providing grounds upon which design choices can be made. They serve the theoretical goals by providing starting points for the framework and are validated, refuted, or refined when interventions are tested during evaluation and reflection (McKenney and Reeves 2012).

The purpose when formulating design propositions here is not to carry out a broad literature overview in a search for the aggregated best design. Such overarching frameworks (NCTM 2000) tend to be far too complex for the limited quest for relations between CMR/AR task designs and learning processes/outcomes. Instead, the propositions are largely based on two fundamental works of mathematics education: Schoenfeld's (1985) work on problem solving and Brousseau's (1997) idea of devolution of problems.

The design propositions, as well as the design principles below, are intended to be prescriptive and are stated here in a somewhat simplified version based on the one suggested by van den Akker (2010): *If the goal is G, then this can be achieved by claim C, owing to the empirical and/or theoretical argument A*. The propositions contain claims about how task or teaching design may promote the two types of reasoning and how this may affect learning (Fig. 1). For simplicity and clarity, the statements are written in a categorical format aiming at capturing main strands but not all variations that exists in reality.

3.3.1 Task-design propositions

1. If the goal is for students to develop mathematical competence, learning by CMR tasks is more efficient than learning by AR tasks.
Argument: this general formulation is supported by TDS. However, the study of mathematical competence in general is outside the scope of this paper. Therefore, propositions related to more specific aspects are formulated.
2. If the goal is for students to develop problem solving ability, learning by CMR tasks is more efficient than learning by AR tasks.
Argument: students develop problem solving ability if and only if they practice problem solving (Schoenfeld 1985; Hiebert 2003). This statement is somewhat oversimplified, and it is important that not only task design but also teaching design be adapted to problem solving (Stein et al. 2008). However, this argument is based on research on problem solving, and CMR tasks are not exactly the same as problems (see the Sect. 6 above). Thus, it is not known whether students will develop problem-solving ability by practicing on non-challenging CMR tasks. However, the research referred to above has shown that when given AR tasks, students mainly apply the algorithms without trying to construct any parts of the solution, thus making it unlikely for

problem solving ability to develop. It seems that not even procedural ability is well developed by such tasks (Hiebert 2003).

3. If the goal is for students to develop mathematical understanding, learning by CMR tasks is more efficient than learning by AR tasks.

Argument: because it is necessary to consider mathematical properties in CMR but not in AR, it is likely that CMR will better develop students' understanding.

4 Results obtained within the design research process

This section presents examples of empirical studies evaluating the propositions above.

4.1 Experimental research

Experiments are carried out comparing (a) background variables, such as cognitive capacity, grades, gender, and motivation, (b) practice format, i.e., learning through AR or CMR task design, and (c) learning processes and learning outcomes, which are mainly measured by performance on post-test tasks but also by eye-tracking and brain imaging methods. The task formats are designed to resemble ordinary school tasks but are adapted for the data collection methods (Fig. 2). Eye-tracking and particularly brain-imaging environments impose strong restrictions on task-student interaction, thus also affecting the experimental task design. Jonsson et al. (2014) have allowed students work alone with tasks presented on a computer screen. Eighty-nine students were matched into two groups on the basis of mathematics grades, gender and cognitive ability tests. The AR group received training through 14 task sets with laboratory versions of a design that is common in schools (Fig. 2): a context, a given solution method, an example of how to apply the method and questions that could be solved by the method.

The other group practiced with similar tasks, and the only difference was that no solution procedures were provided; thus, CMR was required to solve the tasks (Fig. 3). As expected, because the students had been given solution methods, the AR group outperformed the CMR group during practice (Fig. 4).

Each student practiced for approximately one-half hour on one occasion. One week later, students from both groups took identical tests with various mathematical questions related to the practice tasks. As shown in Fig. 4, the CMR group outperformed the AR group on the test. In addition, it was found that this performance difference was largest for the students with the lowest cognitive proficiency (measured by standard psychology

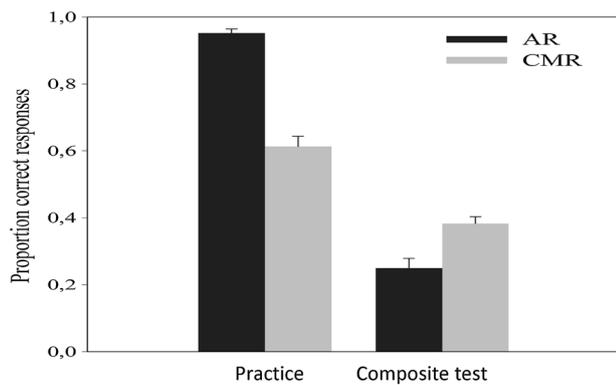


Fig. 4 Practice results and post test results

tests; Operation span and Ravens APM). In other words, it was the students with the lowest cognitive proficiency that had most to gain by CMR practice compared to AR practice. This finding contradicts the common belief that tasks requiring creative reasoning are more suitable for high-performing students. Norqvist (2016) has hypothesised that complementing the AR task design (exemplified in Fig. 2) with written explanations by experienced teachers of why the given solution method work would increase post-test performance. However, in Norqvist's study ($n = 104$), the hypothesis was rejected, and no post-test improvement relative to the task design of Fig. 2 was found.

A functional magnetic resonance brain imaging study ($n = 73$, Karlsson et al. 2015) has found similar performance results, as shown in Fig. 4. Concerning brain activity, one hypothesis was that the CMR group outperformed the AR group because the former had some type of higher activity during the post-test. In fact, the opposite hypothesis was confirmed by the study. Those who learnt by creative reasoning had lower brain activity during the post-test (Fig. 5) and somehow were able to use their mental resources more economically and still perform better. It is difficult to draw inferences from the task

design regarding brain activity, but it seems that practicing by CMR tasks leads to some type of better memory encoding.

4.2 Explanatory research

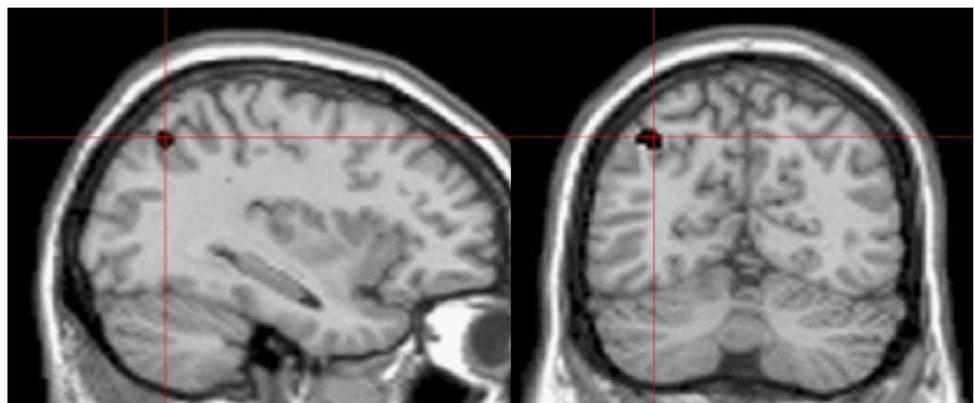
Explanatory research examines how experimental results can be understood. Adhering to the design requirements above, the main priority is to design tasks and settings that enable rich investigations of students' reasoning. Therefore, in contrast to experimental research, explanatory research is based on think-aloud protocols. It may also include small-group work and more complex tasks than those exemplified in Figs. 2 and 3.

Sidenvall et al. (2015) found that students in ordinary classrooms mainly use AR and that apart from obtaining solution templates from books and teachers, the students' peer-peer interaction commonly involves copying one another's solutions. A study by Granberg and Olsson (2015) found that the dynamic software GeoGebra supports collaboration and CMR by providing students with a shared working space and feedback that enhances their creative reasoning. Van Steenbrugge and Norqvist (2016) identified relationships among task design, student characteristics and reasoning type.

4.3 Clinical research

In the experimental and explanatory out-of-school research contexts above, it is possible to design tasks in which the primary priority is to create research data. Thus, it is not necessary that students actually learn anything in line with their curricula goals. In clinical in-school research, the learning goals become the starting point, and the tasks must be designed to align with them. Brousseau (1997) has emphasised that tasks should be designed with the desired target knowledge in mind. In clinical studies, this is achieved with the method of *Hypothetical Learning Trajectory* (Simon

Fig. 5 The AR group activated more of the left angular gyrus brain region during the post-test



1995; Clements et al. 2004), which starts by establishing the student's prior competence in relation to the desired learning goal. Then, a developmental sequence is anticipated, i.e., the student's progression through knowledge levels from the initial state to the learning goal. Finally, a set of tasks that intended to take the student through the developmental sequence is designed. For example, in order for students to create (instead of being given) the standard set of rules for congruent triangles, a sequence of tasks with increasing complexity can be designed.

4.4 Learning goals: task solving understanding and fluency

Design principles include learning goals to aim for. Instead of trying to handle the complexity of broad learning goals (NCTM 2000) and aim for precision, the present versions of the design principles focus on two limited but central aspects of competence: students' understanding of why specific solution methods are suitable and their ability to use these methods.

Mathematics students spend most of their study time on tasks (Boesen et al. 2014), and solution procedures are important (Kilpatrick et al. 2001) but under-researched in mathematics education (Star 2005). Mathematical understanding is often defined in terms of networks, representations and connections. Relating to the NCTM (2000) standards for *representations* of abstract and real mathematical entities and *connections* between representations, the LICR programme developed a definition of mathematical understanding that aims to find not a universally agreed-upon definition but a restricted one that is functional for the purposes of this study:

Task-solving understanding is defined as the ability to justify mathematically the key representations and connections of the methods used in strategy choices and implementations.

This definition, based on the ability to justify, can be extended to aspects of understanding other than task-solution methods, and it is likely that learning by CMR also affects such aspects, but this possibility is outside the scope of this paper. Task-solving understanding is knowing *why* a solution method is suitable for a specific task. Modifying a definition of procedural fluency given by Kilpatrick et al. (2001), we obtain a characterisation of knowing *how* to solve a task.

Task-solving fluency is defined as the skill to choose and implement methods flexibly, accurately, efficiently and appropriately.

5 Design principles

The results from the studies exemplified above and from other studies form the basis for the revision of the design propositions (Sect. 3.3) into the present version of design principles presented in this section. Most of the LICR data are from the Swedish context, and most of the other studies referred to relate to Western culture. The principles may be culture specific, as are other mathematics education results (Leung 2014). The origins of the claims are indicated in the arguments following each principle. The statements are written in a categorical format aiming at capturing the main strands but not all variation that exists in reality. Relating to Fig. 1, the principles are presented in three groups: task design affecting reasoning, task design affecting learning and embryos for teaching design principles.

5.1 Task-design principles related to the reasoning used

5.1.1 The AR task-design principle

If the goal is to design a task such that a student will use AR to solve it, this can be achieved by either providing a task-solution method in connection with the task or judging that the student already knows a method.

Argument: In all LICR studies that have, in various ways, analysed students' work with such tasks, the conclusion is that if a task-solution method is given (by the book, the teacher or a peer) or known in advance, the students will apply the method and seldom explore it further (for example, reflect on why the method is suitable or construct alternative methods; Lithner 2003; Boesen et al. 2010; Sidenvall et al. 2015). In more general terms, this conclusion is also supported by a literature survey by Hiebert (2003). Sometimes, it has been argued that the mere presence of interactive dynamic software leads students into creative explorations, but Olsson (2017a) refuted this claim and indicated that it is the task design, not the presence of the dynamic software, that mainly determines whether the students will use AR or CMR. Boesen et al. (2010) found that when designing tasks, it is possible to approximate what standard solution procedures students know by analysing how present these procedures are in the students' previous textbooks; thus, it is possible, with relatively high certainty, to predict whether students will attempt AR or CMR in a particular task.

5.1.2 The CMR task-design principle

If the goal is to design a task such that a student will use CMR to solve it, the creativity, justification and conceptual challenges must be suitable.

Argument: TDS focuses on the students' *responsibility* to solve the task themselves. But how can one provide for the *possibility* for the student to solve a CMR task, to ensure that it is of suitable difficulty? Earlier analyses (Lithner 2004) have shown that there are at least two types of difficulties involved: the 'creative' challenge and the 'conceptual' challenge. These are complemented in this paper by a third challenge, 'justification'. None of these are necessary in AR. One could add other challenges, such as 'technical' (e.g., complex calculations) or 'linguistic', but these would have the same relevance for AR and are not included here.

The creative challenge concerns the level of ingenuity required. For example, it is more likely that students can solve a task that can be solved by stepwise progress rather than requiring a single far-fetched trick. The conceptual challenge determines how the advanced mathematical properties (such as in representations and connections) of the task need to be understood in order to construct a solution. The justification challenge concerns how difficult it is to use arguments to predict the outcome of a hypothetical solution idea and/or to verify that an implemented solution is correct. An extreme example is the four-colour theorem; it is fairly easy to hypothesise through empirical arguments, but extremely difficult to prove, that four colours are sufficient for any map. Another example is that when interacting with dynamic software, it is important that students can formulate predictive argumentation to fully utilise the feedback obtained when solving CMR tasks (Olsson 2017b).

The creative challenge may be relatively low but still require substantial justification and/or conceptual considerations. Thus, if the purpose is to learn some aspects of a new concept or notion, it is not necessary to have very difficult tasks (as most CMR textbook tasks do; Jäder et al. 2015). In AR, it is not usually necessary to consider the mathematical properties of the components in the reasoning, but even in simple CMR, it is necessary to understand the relevant properties. For example, many students have difficulties in learning the power rules, and thus they try to memorise them as separate rules, even though the rules are based on a few ideas that are valid for all of the rules. A common way to teach these rules is first to describe and explain them and then to let students apply them to large numbers of tasks of the type 'simplify a^5a^3 ' (see almost any algebra textbook). This can be done by following the rule (add the exponents) without considering any basic properties of powers. An alternative that does not include

difficult conceptual challenges is to give students tasks in which they construct at least some of these rules directly from the basic definitions of powers, for example, by first finding out that $a^5a^3 = aaaaa \cdot aaa = a^{5+3}$ and then generalising this idea to $a^m a^n = a^{m+n}$.

5.2 Task-design principles related to fluency and understanding

These principles concern how task design is related to learning in terms of task-solving fluency and understanding (as defined above).

5.2.1 The task-solving fluency principles

- (a) If the goal is for a student to solve a specific task successfully, an AR task design is more efficient than a CMR task design.

Argument: In all experimental studies in which students have been given corresponding AR and CMR tasks, the proportion of solved AR practice tasks is significantly higher (Fig. 4). This conclusion is also theoretically expected, and is almost self-evident, because the solution templates included in these tasks reduce the task difficulty (see Brousseau's characterisation of algorithms above). There are also indications that having a high proportion of solved tasks enhances students' self-confidence, at least in the short term. However, it may be questioned whether this type of "efficiency" is desirable (see the Discussion section below).

- (b) If the goal is for students to develop task-solving fluency, learning by CMR tasks is more efficient than learning by AR tasks.

Argument: Theoretical arguments are presented in the design propositions above. With practice tasks such as those in Figs. 2 and 3, the proportion of solved post-test tasks is higher for students practicing by CMR tasks in all LICR studies to date (Fig. 4). Examples of post-test tasks are those with short time recall formulas (such as $y=3x+1$ in Fig. 2); those with short time recall and applied solution methods (for example, finding the number of matches to get 100 squares in a row); and those with more time available to reconstruct solution methods (Jonsson et al. 2014). Jonsson et al. (2016) have shown that post-test effects are caused by the effortful struggle related to CMR and not to the transfer of appropriate processing (i.e., similarities between how information is encoded and retrieved). In addition, in a brain imaging study by Karlsson et al. (2015), students who practiced with CMR tasks scored more highly on post-test tasks and

exerted less effort in terms of brain activity. Ongoing pilot studies in ordinary classrooms indicate that CMR task design may also be more efficient in this setting.

- (c) If the goal is for students to develop task-solving fluency, adding justifications that explain the solution methods given in AR tasks does not improve learning compared to AR tasks in which the given solution method is not explained.

Argument: Brousseau's (1997) theory can be interpreted to be rather categorical in claiming that students can truly learn only when they construct the knowledge themselves (1997 p. 30). Nonetheless, it can reasonably be hypothesised that if the task information (or teacher) not only presents *how* the solution method to a task works but also explains *why* it works, then students will learn better. However, Norqvist (2016) found no learning gains after experienced teachers added written explanations to AR practice tasks. Although that study supports the categorical interpretation of TDS, this design principle needs to be further analysed to be better confirmed, modified or refuted—for example, by not only providing explanations but also by ensuring that students actively try to use them in effective ways.

5.2.2 The task-solving understanding principle

If the goal is for students to develop task-solving understanding, learning by CMR tasks is more efficient than learning by AR tasks.

Argument: Task-solving understanding is substantially more difficult to measure than task-solving fluency, particularly through comparisons of the effects of the two task designs, and LICR studies have only begun such attempts. Theoretically, one line of argumentation is presented in the design propositions above. Empirically, the findings by Olsson (2017a) indicate that students practicing by addressing CMR tasks are better in the post-test at explaining and justifying their choices and claims. An ongoing study indicates that students addressing such tasks (for example, as in Fig. 3) become more proficient in solving certain transfer post-test tasks. For example, in a task similar to that in Fig. 3, but with the 1×1 row of squares replaced by a 1×2 row of rectangles, the same solution idea but not the same algorithm ($y = 3x + 1$) can be used. This transfer ability is taken as an indication that students addressing CMR tasks better understand their solution. Another ongoing eye-tracking study indicates that students learning by CMR tasks are more focussed on the parts of the task that are judged necessary to consider in order to develop task-solving understanding.

5.2.3 The cognitive capacity principle

The design principles above are valid for students of varying cognitive capacities.

In this broad formulation, this principle is still largely hypothetical. The reason to include it is that it appears that students with mathematics learning difficulties are sometimes assigned imitative tasks more than average students are, and that teachers sometimes claim that students with a lower capacity cannot solve creative tasks. In contrast, Jonsson et al. (2014) found that students with lower cognitive capacity have more to gain from learning by CMR tasks than do students with higher cognitive capacity.

5.3 Teaching design principles

Although this paper primarily concerns task design, and LICR research on teaching design (Fig. 1) is in the early phases, the present teaching design principles are briefly presented.

5.3.1 The AR teaching design principle

If the goal is for students to use AR to solve tasks, this can be achieved by giving the students solution methods either before or during the task-solving attempts. In this situation, it is not necessary for the teacher to understand the students' specific needs, because regardless of what they are, the teacher's action can always be the same: to describe the solution method.

Argument: In this theoretical version of "AR teaching", it is not necessary for the teacher to do anything apart from describing task-solution methods for students to apply and learn (perhaps by rote learning). Although this may not be common in a strict sense, the general idea of providing solution templates is common (Hiebert 2003; Boesen et al. 2014). Teaching in line with an AR design is easier to plan and requires less teacher competence and less teacher and student effort, at least in a short-term perspective.

5.3.2 The CMR teaching design principle

If the goal is for students to use CMR to solve tasks, first let the students try to construct their own solutions. If this fails, then, in line with characteristics of formative assessment, diagnose the students' task-specific difficulties and provide feedback that supports students' ability and responsibility to construct solutions.

Argument: This proposition is in line with TDS and has general support in reviews (Hiebert 2003; Niss 2007). In contrast to AR teaching, if the teacher's goal is to support the students' CMR, then it is necessary both to diagnose what the students' difficulties with the particular task are,

and to provide feedback that is adapted to students' difficulties (but does *not* provide a solution method). Thus, such teaching by necessity is designed as *formative assessment*, which in several reviews has been shown to be one of the most effective ways to enhance student learning (Black and Wiliam 1998; Hattie 2009).

6 Discussion

6.1 The paradox: no one advocates rote learning, but it is common

No one advocates rote learning, at least not as the only way to learn. Although it is far from clear whether, when and how alternative approaches such as problem solving, modelling, and explorative learning are better, there is a growing body of evidence supporting such alternatives or complements to imitative teaching and learning. Concerning task design, Coles and Brown (2016) noted the persistence of a gap between teacher intentions and student activity in the literature. Many mathematics textbooks worldwide are dominated by imitative tasks (Jäder et al. 2015), and the relatively few creative tasks may be turned into imitative tasks if teachers provide solution templates (Stein et al. 1996). A fundamental question is why imitative tasks continue to dominate. A partial answer may be found in the conflict between short- and long-term teaching goals. Teaching based on an AR task design is more efficient in the sense that (a) it takes less time to prepare for a lesson, (b) it requires less teacher competence, because the main teaching strategy can be to describe the solution methods, which requires only that the teacher be able to solve the tasks, and (c) the students will to a larger extent know what to do and be able to solve large numbers of tasks with (d) a minimal need for help other than the given solution methods. As teachers, we may for good reason believe that giving our students solution methods is the best way to help them learn. In contrast, the main gains from CMR task design seem to be from a longer-term perspective than a single lesson in terms of, for example, students' improved understanding and problem-solving ability. The short-term/long-term perspective contrast may also be relevant for individual students in terms of AR tasks requiring less struggle during training but leading to rote learning and a weak understanding.

6.2 Is an AR task design never better?

Of course, a CMR task design is not an option when the target knowledge is too difficult for students to attain through their own constructions; and even when attaining such knowledge is possible, it may take too much time.

One elementary argument is that it took the world's best mathematicians thousands of years to construct our upper-secondary school mathematics. Thus, it seems utopian to design tasks so cleverly that students can construct all of the knowledge by themselves. It is more realistic to find a suitable balance between the two task designs.

It also seems reasonable to consider whether an AR design can be better than the relatively strict versions used in the studies above. One attempt mentioned above has failed in the sense that the added explanations did not yield better post-test results. Nonetheless, it seems difficult to accept the categorical statement of Brousseau (1997, p. 30) that students can truly learn *only* by constructing their own solutions. Perhaps richer explanations are required, such as explanations from a teacher or other types of social interactions. Another possibility may be *practice testing learning*, which in recent studies has been shown to strengthen learning (Dunlosky et al. 2013) and could probably be used with AR tasks to enhance memorisation but perhaps not understanding. There are, of course, also numerous other potential possibilities.

6.3 How, when and why does CMR task design lead to better learning?

As argued above, it seems reasonable that the reason that CMR practice leads to better task-solving understanding than AR practice is that intrinsic mathematical properties must be considered when solving CMR tasks. But why does CMR lead to better task-solving fluency? From the main starting points of this paper (Schoenfeld 1985; Brousseau 1997), it can be argued that CMR enhances all learning. But what are the more specific mechanisms? LICR researchers, in mathematics education, psychology and neuroscience, agree that the effect is likely to be caused by some type of productive struggle (Niss 2007), but these researchers have somewhat different specific hypotheses explaining the effect. One hypothesis is that the struggle itself leads to more efficient memory consolidation of the task-solving methods learned. Another is that the increased understanding reached by using CMR enhances task-solving fluency and transfer. This is one of the questions being further pursued.

Another remaining question is what specific competencies, other than task-solving fluency and understanding, are enhanced by CMR. To date, most interventions have been very short, from 30 min to a few hours. The learning from short practice sessions is probably very local and restricted to, for example, task-solving fluency and understanding related to limited sets of tasks. Broader competencies, such as problem solving, modelling and communication abilities and understanding the big ideas of mathematics, take a much longer time, months and years, to develop, and the

LICR programme is beginning to engage in longer interventions. A third question is whether the findings by Jons-son et al. (2014) that students with a lower cognitive capacity have more to gain from CMR practice than do students with a higher capacity, are also valid for students with more severe learning difficulties.

It is also possible that different emphases on the three challenges of the CMR task-design principle may develop competencies differently: An emphasis on the creative challenge may enhance problem solving, on the justification challenge to enhance reasoning and on the conceptual challenge to enhance mathematical understanding. However, these possibilities remain to be disentangled as the LICR programme seeks to further develop the design principles.

Acknowledgements Funding was provided by The Swedish Research Council, The Marcus and Amalia Wallenberg Foundation, The Kempe Foundations and Umea University.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bergqvist, T., & Lithner, J. (2012). Mathematical reasoning in teachers' presentations. *Journal of Mathematical Behavior*, 31, 252–269.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student thinking*. Mahwah, NJ: Lawrence Erlbaum.
- Boesen, J., Helenius, O., Lithner, J., Bergqvist, E., Bergqvist, T., 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. (2010). The mathematical reasoning required by national tests and the reasoning used by students. *Educational studies in mathematics*, 75, 89–105.
- Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluwer Academic Publishers.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences*, 2(2), 141–178.
- Clements, D. H., Wilson, D. C., & Sarama, J. (2004). Young children's composition of geometric figures: A learning trajectory. *Mathematical Thinking and Learning*, 6, 163–184.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9–13.
- Coles, A., & Brown, L. (2016). Task design for ways of working: making distinctions in teaching and learning mathematics. *Journal of Mathematics Teacher Education*, 19, 149–168.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4–58.
- Fan, L., & Bokhove, C. (2014). Rethinking the role of algorithms in school mathematics: A conceptual model with focus on cognitive development. *ZDM—The International Journal on Mathematics Education*, 46, 481–492.
- Granberg, C., & Olsson, J. (2015). ICT-supported problem solving and collaborative creative reasoning: Exploring linear functions using dynamic mathematics. *Journal of Mathematical Behaviour*, 37, 48–62.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hiebert, J. (2003). What research says about the NCTM standards. In J. Kilpatrick, G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 5–26). Reston, Va.: NCTM.
- Hiebert, J., & Grouws, D. (2007). The effects of classroom mathematics teaching on students' learning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1293–1312). Charlotte, NC: Information Age Publishing.
- Jäder, J., Lithner, J., & Sidenvall, J. (2015). *A cross-national textbook analysis with a focus on mathematical reasoning—The opportunities to learn*. In: J. Jäder, Elevers möjligheter till lärande av matematiska resonemang, Licentiate thesis, Linköping University, 2015.
- Jonsson, B., Kulaksiz, YC. & Lithner, J. (2016) Creative and algorithmic mathematical reasoning: Effects of transfer-appropriate processing and effortful struggle. *International Journal of Mathematics Education in Science and Technology*. doi:10.1080/0020739X.2016.1192232.
- Jonsson, B., Norqvist, M., Liljekvist, Y., & Lithner, J. (2014). Learning mathematics through algorithmic and creative reasoning. *Journal of Mathematical behaviour*, 36, 20–32.
- Karlsson, L., Lithner, J., Jonsson, B., Liljekvist, Y., Norqvist, M., & Nyberg, L. (2015). Learning mathematics without a suggested solution method: Durable effects on performance and brain activity. *Trends in Neuroscience and Education*, 4, 6–14.
- Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. Washington, D.C.: National Academy Press.
- Leung, F. (2014). What can and should we learn from international studies of mathematics achievement? *Mathematics Education Research Journal*, 26, 579–605.
- Lithner, J. (2000). Mathematical reasoning in school tasks. *Educational Studies in Mathematics*, 41(2), 165–190.
- Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational Studies in Mathematics*, 52, 29–55.
- Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *Journal of Mathematical Behavior*, 23, 405–427.
- Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics*, 67(3), 255–276.
- McKenney, S., & Reeves, T. (2012). *Conducting educational design research*. London/New York: Routledge.
- NCTM (2000). *Principles and standards for school mathematics*. Reston, Va.: National Council of Teachers of Mathematics.
- Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. *Third Mediterranean Conference on Mathematics Education, Athens*, 115–124.
- Niss, M. (2007). Reactions on the state and trends in research on mathematics teaching and learning: From here to utopia. In F. Lester (Ed.), *Second handbook of research on mathematics*

- teaching and learning* (pp. 1293–1312). Charlotte, NC: Information Age Publishing.
- Norqvist, M. (2016). *On mathematical reasoning—Being told or finding out*. PhD thesis, Umeå University.
- Olsson, J. (2017a). *GeoGebra, enhancing creative mathematical reasoning*. PhD thesis, Umeå University.
- Olsson, J. (2017b). The contribution of reasoning to the utilization of feedback from software when solving mathematical problems. *International Journal of Science and Mathematics Education*. doi:10.1007/s10763-016-9795-x (**Advance online publication**).
- Pólya, G. (1954). *Mathematics and plausible reasoning*. Princeton, N.J.: Princeton U.P.
- Ruthven, K., Laborde, C., Leach, J., & Tiberghien, A. (2009). Design tools in didactical research: Instrumenting the epistemological and cognitive aspects of the design of teaching sequences. *Educational Researcher*, 38, 329–342.
- Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
- Schoenfeld, A. (2007). Method. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 69–107). Charlotte, NC: Information Age Publishing.
- Schoenfeld, A. (2015). Thoughts on scale. *ZDM*, 47, 161–169.
- Shield, M., & Dole, S. (2013). Assessing the potential of mathematics textbooks to promote deep learning. *Educational Studies in Mathematics*, 82, 183–199.
- Sidenvall, J., Lithner, J., & Jäder, J. (2015). Students' reasoning in mathematics textbook task-solving. *International J. of Mathematics Education in Science & Technology*, 46, 533–552.
- Silver, E. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *ZDM—The International Journal on Mathematics Education*, 29(3), 75–80.
- Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26, 114–145.
- Smith, M. S., & Stein, M. K. (2011). *Five practices for orchestrating productive mathematics discussions*. Reston, VA: National Council of Teachers of Mathematics.
- Sriraman, B., Haavold, P., & Lee, K. (2013). Mathematical creativity and giftedness: A commentary on and review of theory. *ZDM—The International Journal on Mathematics Education*, 45(2), 215–225.
- Stacey, K., & Vincent, J. (2009). Modes of reasoning in explanations in Australian eighth-grade mathematics textbooks. *Educational Studies in Mathematics*, 3, 271–288.
- Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, 36, 404–411.
- Steigler, J., & Hiebert, J. (1999). *The teaching gap*. New York, NY: The Free Press.
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10, 313–340.
- Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning. *American Educational Research Journal*, 33, 455–488.
- Swan, M. (2008). Designing a multiple representation learning experience in secondary algebra. *Educational Designer*, 1, 1–17.
- Thompson, D., Senk, S., & Johnson, G. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education*, 43, 253–295.
- Van den Akker, J. (2010). Building bridges: How research may improve curriculum policies and classroom practices. In S. Stoney (Ed.), *Beyond Lisbon 2010: Perspectives from research and development for education policy in Europe Europe (CIDREE Yearbook 2010)* (pp 175–196). Slough: NFER.
- Van Steenbrugge, H. & Norqvist, M. (2016). Unraveling students' reasoning: analyzing small-group discussions during task solving. In M. Norqvist (Ed.), *On mathematical reasoning - Being told or finding out*. PhD thesis, Umeå University.