Educ Stud Math (2008) 67:255–276
DOI 10.1007/s10649-007-9104-2

A research framework for creative and imitative reasoning

Johan Lithner

Published online: 5 December 2007

Abstract This conceptual research framework addresses the problem of rote learning by characterising key aspects of the dominating imitative reasoning and the lack of creative mathematical reasoning found in empirical data. By relating reasoning to thinking processes, student competencies, and the learning milieu it explains origins and consequences of different reasoning types.

Keywords Mathematical reasoning · Creativity · Rote learning

1 Introduction

A central problem in mathematics education is that we want students to become problem solvers, but even after 20 years of research and reform many students still do rote thinking (Hiebert 2003). This thinking is a main factor behind learning and achievement difficulties and has been analysed in a series of empirical studies (e.g. Lithner 2000b, 2003, 2004). This paper presents the framework that has gradually been formed to specify and communicate the findings of these studies.

There exist several frameworks for analyses of aspects related to rote learning. Some are not primarily designed as research tools and encompass all central goals of mathematics education (National Council of Teachers of Mathematics 2000; Niss 2003). Some propose stages of understanding. Skemp (1978) distinguishes between instrumental and relational understanding, Sfard (1991), Asiala et al. (1996) and Tall (2004) analyse how development of understanding of objects and processes interact, and Pirie and Kieren (1994) describe levels of understanding. Some define categories of cognitive achievement, like the Bloom (1956) and van Hiele (1986) taxonomies. Other frameworks structure particular competencies, e.g. for non-routine problem solving Schoenfeld (1985). However, there are not many that aim at characterising

J. Lithner (✉)
Department of Mathematics, Umeaa University, S90187 Umeaa, Sweden
e-mail: Johan.Lithner@math.umu.se

Fig. 1 The origin of reasoning

the reasoning itself. One example is given by Balacheff (1988), but as in most reasoning frameworks it concerns proofs, and that is too narrow for the scope of this paper, even if proof is considered in a wider sense as justification (Ball and Bass 2003). The purpose of this framework is to characterise imitative (rote learnt) reasoning and creative reasoning, and explain the origins and consequences of different reasoning types.

A *research framework* is "a basic structure of the ideas (i.e., abstractions and relationships) that serve as the basis for a phenomenon that is to be investigated" (Lester 2005, p. 458). It provides a structure for conceptualising and designing research studies, makes it possible to make sense of data, and allows us to transcend common sense. There are different kinds of research frameworks. While a theoretical framework guides research activities by reference to formal theory, this one is a conceptual framework that is "an argument that the concepts chosen for investigation, and any anticipated relationships among them, will be appropriate and useful given the research problem under investigation" (ibid. p. 460). Such frameworks are built from an array of possibly far-ranging sources and may be based on different theories and various aspects of practitioner knowledge. This framework is intended for use-inspired basic research (ibid.), which – relating to the tension between basic and applied science – aims at both increased fundamental understanding and at contributing to developing teaching. The underlying characterisations of students' reasoning emanate from a cognitive psychology perspective, but extend into socio-cultural considerations when addressing potential causes and consequences.

A basic idea of this framework is that rote learning reasoning is *imitative*, while the opposite type of reasoning is *creative*. When trying to explain origins of different reasoning types, the approach (Fig. 1) is to separate the *reasoning sequence* from the *thinking process* that created it. In order to understand why different thinking processes are activated or not, this framework sees them as guided and limited by the *student competencies* that are formed in a sociocultural *milieu*.

2 Reasoning

The series of studies mentioned above started as an attempt to understand what made students fail or succeed in common practice and test tasks. Soon the focus came to be on the disadvantages of being restricted to imitative reasoning. Consider the following example.

The task is to find the largest and smallest values of $y = 7 + 3x - x^2$ on the interval $[-1, 5]$. Ben, a skilled tenth-grade natural science student, solves it by a familiar method: Maxima and minima can be found where $y'(x) = 0$, where $y'(x)$ is undefined, and at the endpoints of the interval. Ben knows that for polynomials $y'(x)$ is defined for all x, and differentiation and calculating function values are trivial for him.

This example of reasoning is structurally simple, but raises some fundamental questions. In the situation above (a laboratory test), what are the reasons that Ben succeeds and others fail? If this was a practice task in a textbook, what is learnt by solving it? If it was included in a high-stake assessment, what is tested? In order to address such questions, one aim is to be relatively specific in defining both general characteristics of reasoning and particular reasoning types.

2.1 Reasoning sequences

"Mathematical reasoning is no less than a basic skill" (Ball and Bass 2003, p. 28). Still, the term 'reasoning' is mostly used among mathematics educators without defining it, under the implicit assumption that there is a universal agreement on its meaning (Yackel and Hanna 2003). In this paper, *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 not restricted to proof, and may even be incorrect as long as there are some kinds of sensible (to the reasoner) reasons backing it. *Task* includes most of the work requested from students in classrooms, such as exercises, tests, group work, etc. An *answer* provides the information requested. A *solution* is an answer and a motivation why the answer is correct. This solution often does not display the actual reasoning used to reach the answer, but an idealised summary of it. The term 'problem' has been used in the literature with many different meanings but denotes in this paper a task that is intellectually difficult for an individual (Schoenfeld 1985).

Reasoning can be seen as thinking processes, as the product of these processes, or both. What we mainly can see in data is behaviour while we can only speculate about the underlying thought processes (Vinner 1997). Since one purpose of this framework is to characterise data, the choice is to see reasoning as a *product* that appears in the form of a *sequence of reasoning* that starts in a task and ends in an answer. It may contain more data than an ordinary written solution, which may be supplemented by, for example, think-aloud protocols and interviews.

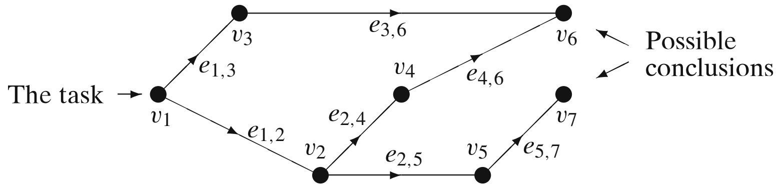
2.2 Reasoning structure

Solving a task can be seen as carrying out the following four steps.

1) A (sub) task is met, which is denoted *problematic situation* if it is not obvious how to proceed.
2) A *strategy choice* is made, where 'strategy' ranges from local procedures to general approaches and 'choice' is seen in a wide sense (choose, recall, construct, discover, guess, etc.). It can be supported by *predictive argumentation*: Why will the strategy solve the task?
3) The *strategy* is *implemented*, which can be supported by *verificative argumentation*: Why did the strategy solve the task?
4) A *conclusion* is obtained.

The reasoning sequence can be represented by a path in a directed graph (Fig. 2). A vertex v_n represents both a momentary state of knowledge and of the (sub)task. The reasoner makes a strategy choice among the edges leading from v_n. The strategy implementation is represented by a transition edge $e_{n,m}$. Here knowledge not already accessed in v_n is recalled or constructed and added up to form the new knowledge state in v_m, where the task is partially resolved and therefore a new task state is formulated. A *reason* is the motivation supporting transitions between vertices.

Fig. 2 Task solving as a path in a graph



There is always a reason for attempting a particular transition, even if it is vague or superficial.

The analyses of empirical reasoning data, that will not be presented in detail below since they are published elsewhere (Lithner 2000b, 2002, 2003, 2004; Bergqvist et al. 2007), were typically done in several steps, including transcription of the reasoning and interpretation (a completion, e.g. by interviews, giving a more coherent picture of the actual reasoning); identification of the main problematic situations; characterisation of the argumentation.

3 Imitative reasoning

A key characteristic of Ben's solution is that the path is laid out from the start. He imitates a solution procedure memorised from the textbook. The empirical studies behind this framework have identified two main types of imitative reasoning, memorised and algorithmic.

3.1 Memorised reasoning (MR)

Memorised reasoning fulfils the following conditions.

1. The strategy choice is founded on recalling a complete answer.
2. The strategy implementation consists only of writing it down.

All task solving builds partly on recollection, but as an overall strategy it is useful only in a few task types like those asking for facts ("How many cm^3 is a litre?"), definitions ("What is a polynomial?"), and proofs. An undergraduate examination task was "State and prove the Fundamental Theorem of Calculus". The answers by almost all of the 50% of the 150 students who got full credit were copies of the two-page textbook proof. Most of the faulty answers contained large parts of the same proof, but with some parts missing or in reversed order resulting in broken logical chains. A post-test asked the students to explain a sequence of six equalities included in the proof, most of them elementary in relation to the proof's complexity:

$$F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{1}{h}\left(\int_a^{x+h} f(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t\right)$$

$$= \lim_{h\to 0} \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t = \lim_{h\to 0}\frac{1}{h} hf(c) = \lim_{c\to x} f(c) = f(x)$$

Most students were able to explain only a few of them, thus they did not understand the long proof they managed to memorise.

Another type of MR influence is through established experiences from the learning environment, including apprehensions of facts, concept images, and beliefs. One example is when Jed calculates $k = 6$ as the correct value of the slope of a linear function, but dismisses it by the argument that "It shouldn't be that big, k, the line is too steep. [...] It is always a smaller number like 1 or -2." Such arguments often override more mathematically founded reasoning.

3.2 Algorithmic reasoning (AR)

School tasks normally ask for calculations where it is more appropriate to recall not the answer but an algorithm. "An algorithm is a finite sequence of executable instructions which allows one to find a definite result for a given class of problems" (Brousseau 1997, p. 129). Its importance is that it can be determined in advance. The nth transition does not depend on any circumstance unforeseen in the n-1st transition; not on finding new information, any new decision, any interpretation, and thus on any meaning that one could attribute to them.

In this framework 'algorithm' includes all pre-specified procedures (not only calculations), e.g. to find the zeros of a function by zooming in on its intersections with the x-axis on a graphing calculator. *Algorithmic reasoning* (AR) fulfils the following two conditions.

1. The strategy choice is to recall a solution algorithm. The predictive argumentation may be of different kinds (see below for examples), but there is no need to create a new solution.
2. The remaining reasoning parts of the strategy implementation are trivial for the reasoner, only a careless mistake can prevent an answer from being reached.

This notion of 'algorithm' is much wider than the tendency (which seems fairly common) to see an algorithm only as something that is explicitly taught in the form of a chain of executable instructions, for example the division algorithm. The main point with this wider notion is that all parts that are conceptually difficult are taken care of by the algorithm and only the easy parts are left to the student, which may limit the learning effect. I have easily taught normally talented 7-year old children to differentiate simple polynomials. They had of course absolutely no conceptual insight into what they were doing, but their performance would grant a few credit points in common upper secondary mathematics examinations. Thus AR may be carried out with limited as well as full understanding of the procedure.

Ben's implicit argument was that he knew that the task type could be solved by a particular algorithm. He was in no problematic situation, and neither did he have to consider the mathematical meaning of the components of the task nor provide any predictive or verificative arguments. He understood the mathematical foundation of the solution method, but his reasoning did not refer to this understanding and did in essence not go beyond the reasoning of the 7-year old children differentiating polynomials.

One may question whether the term 'reasoning' is proper for a solution without explicit arguments, but there are always at least implicit reasons for the strategy choices. In order to clarify what these reasons may be it is necessary to first clarify the concept of 'argument'.

4 Argumentation

Reasoning can have many functions in mathematics, including verification, explanation, systematisation, discovery, communication, construction of theory, and exploration (Yackel and Hanna 2003). Thus there are several potential aspects of argumentation to consider. This framework proposes a focus relating to the phases of problem solving formulated by Pólya and elaborated by Schoenfeld (1985): reading the task (including noting conditions and goals), analysis (to understand, select perspective, and perhaps reformulate), exploration (a broader and less structured search for information), planning, implementation (including evaluation of progress), and verification. I suggest that analysis, exploration and planning are primarily supported by predictive argumentation, and implementing and verification by verificative argumentation. Analysis argues why properties of the task components have certain consequences. Exploration considers why some of the outcomes may be useful. Planning concerns why certain approaches may better lead towards a solution. Implementing has metacognitive arguments concerning why it proceeds in a proper way, and perhaps why the strategy needs to be reconsidered. Verification is an explanation of why a solution is actually reached. If understanding of the task is immediate in the reading phase, no argumentation is needed. Otherwise understanding is sought by analysis or exploration.

4.1 What is an argument?

In Toulmin's model (Krummheuer 1995) the argument contains four components: conclusion, data, warrant and backing. Data are facts that constitute the starting point of the task. The warrant supports the conclusion by referring to the data, to register the legitimacy of the step (the transition in Fig. 2) involved by referring to a class of steps whose legitimacy is being presupposed. The warrant is a specific reference to the data, and its authority can be supported by a more general backing that refers to "global convictions and primary strategies that can be expressed in the form of categorical statements" (ibid. p. 244). The data, the thing one is reasoning about, is discussed in Section 4.2 and is in Fig. 2 included in the task and knowledge states.

I see the quality of an argument in task solving as determined by three factors: validity, ability to convince, and constructiveness. There is a distinction between the more general 'argument' and 'valid reasoning', where the latter is based on the organisation of several propositions into a deductive step and the organisation of several steps into a proof (Duval 2002). Harel (2006) argues that if this distinction is not understood we may advance mostly argumentation skills and no mathematical reasoning. The meaning of a statement is based on its content, status (premise, conclusion, theorem), logical value and epistemic value, where the latter is the degree of trust (absurd, unreal, possible, likely, obvious) that a person has in a statement as soon as she understands its content (Duval 2002). Logical value is independent of the person and characterises a proof. Outside mathematics, valid reasoning is not the main concern but the strength of an argument in the sense that it is convincing. Regardless if it is valid or not, reasoning that does not convince fails (ibid.).

Validation is a mental process of determining the correctness of a sequence of reasoning, which is only partly conscious and can include asking and answering questions, assenting to claims, constructing subproofs, remembering and interpreting other theorems and definitions, complying with instructions, and feelings of rightness or wrongness (Selden and Selden 2003). The process is social (Krummheuer 1995) and "comprises a set of practises and norms that are collective" (Ball and Bass 2003, p. 29). The validity of an argument is determined by sociomathematical norms (Yackel and Cobb 1996). For example, that students are expected to justify solutions is a social norm but what counts as an acceptable justification is a sociomathematical norm. The negotiation of these norms provides learning opportunities while a lack of such negotiation is a lack of devolution of problem (Section 9) and can therefore hinder insights into the characters of these norms.

4.2 Anchoring in mathematical properties

An acceptable justification is based on mathematics rather than social status, e.g. the authority of the teacher or the smartness of a peer (Yackel and Cobb 1996). But what does it mean for the warrant to be based on mathematics? Schoenfeld (1985) found that novices used naive empiricism to judge that geometrical constructions are correct if they 'look good' and experts used more relevant properties (e.g. congruence). Thus content is important: What are the arguments about?

In order to address this question I introduce the notion of *anchoring*, which refers not to the logical coherence of the warrant but to its fastening in data in relevant mathematical properties of the *components* one is reasoning about: objects, transformations, and concepts. The *object* is the fundamental entity, the 'thing' that one is doing something with, e.g. numbers, variables, functions, diagrams, etc. A *transformation* is what is being done to an object, and the outcome is another object. A sequence of transformations, e.g. finding polynomial maxima, is a *procedure*. A *concept* is a central mathematical idea built on a set of objects, transformations, and their properties, e.g. the function or infinity concept. The status of a component depends on the situation. $f(x) = x^3$ can be seen as a transformation of the input object 2 into the output object 8. If f is differentiated then the differentiation is the transformation, $f(x)$ is encapsulated (Tall 1991) into an input object, and $f'(x)$ is the output object.

The relevancy of a mathematical property depends on the context. In deciding if $9/15$ or $2/3$ is larger, the size of the numbers (9, 15, 2, 3) is a *surface* property that is insufficient to consider while the quotient captures the *intrinsic* property. Another example (Lithner 2003) is a student trying to determine if the same test for absolute convergence is applicable to the series $\sum_{n=1}^{\infty} \frac{\cos n\pi}{(n+1)\ln(n+1)}$ and $\sum_{n=1}^{\infty} \frac{n\cos n\pi}{2^n}$. He decides it can, based on the surface property that $\cos n\pi$ appears in both numerators, while the intrinsic property in this comparison lies in other parts of the fractions. The intrinsic - surface notion was introduced since one of the reasons behind students' difficulties was the anchoring of arguments in surface properties (ibid.).

5 Variants of imitative reasoning

In AR, the main difficulty is to identify a suitable algorithm. Three common ways are found, which may be based on surface property considerations only and are then mathematically unstable.

5.1 Familiar AR

The elementary Keyword strategy is determined by the appearance of keywords, e.g. 'more' or 'less' that correspond to the addition and subtraction algorithms respectively (Hegarty et al. 1995). It may be generalised to more complex situations, and the reasoning in a task solution attempt fulfilling the following is called *familiar AR*.

1. The reason for the strategy choice is that the task is seen as being of a familiar type that can be solved by a corresponding known algorithm.
2. The algorithm is implemented.

The (implicit) argument that convinces the reasoner of the strategy choice is often based, through similar practice tasks, on the established experience that a task with certain textual, graphical and/or symbolical features is related to a corresponding algorithm, e.g. as Ben connects a differentiation algorithm to a task type asking for the largest value of a function without considering the meaning of differentiation. Since an algorithm is "designed to avoid meaning" (Brousseau 1997, p. 130) there is no need for any argumentation once the algorithm is identified. However, the validity of superficial familiar AR is not based on any anchoring in intrinsic mathematical properties and is therefore not reliable in problematic situations.

Familiar AR is common. For example, Lithner (2000a) found that 33 of 46 university mathematics students made a superficial interpretation of the task of finding the number of produced units x that maximises the profit if the production cost is $f(x) = -2x^2 + 2{,}000x - 420{,}000$ kr/unit, the sale price is $g(x) = -x^2 + 700x$ kr/unit, and $400 \leq x \leq 600$. They reached a clear but faulty image that it was of a more familiar type, where the difference $g(x) - f(x)$ is maximised (only 5 of 46 made a correct interpretation, to maximise $x(g(x) - f(x))$).

5.2 Delimiting AR

Sally's approach to the same task as Ben (Section 2) is also algorithmic, but in a quite different way (Bergqvist et al. 2007). She differentiates y, solves $y'(x) = 0$ ($x = 1.5$), and evaluates $y(1.5) = 9.25$. She hesitates: "I think I should have got two values (meaning max and min), and I don't know what I did wrong." She abandons this method only because the algorithm gave one value instead of two, she does not reflect on why. Sally instead tries her calculator's minimum function and in order to use it she has to enter a rough guess of the minimum's location. She does not understand why she cannot see the minimum on the graph (y has only one critical point, a maximum) to estimate its location, and again abandons the method without reflection. Sally then uses the calculator's table-function to evaluate y for $x = -1, 0, ...5$. She finds min $= -3$ and max $= 9$, but is troubled since she found a larger value ($y = 9.25$) earlier. She does not understand why the table method does not work and abandons it without reflection. She finally solves $7 + 3x - x^2 = 0$ and gets two values, $x_1 \approx 4.54$ and $x_2 \approx -1.54$. This method is completely wrong but she sees it as the solution.

In familiar AR, the vertices and edges of Fig. 2 are all apprehended (wrongly or rightly) to be specified by a familiar algorithm, but the present task is not sufficiently familiar to Sally. Nevertheless, though there is no attempt at planning or at analytically arguing from the intrinsic properties of the task, her choice of algorithm is not random. This would be unrealistic since she knows hundreds of algorithms, so Sally delimits the set of potentially useful ones by trying only those that have some connection to the task. 'Largest and smallest' is related to the first two methods, functions (represented by the polynomial) to tables, and the polynomial to the quadratic equation algorithm. The initial drawback with this approach is that these connections are based only on surface property considerations, so she is unable to predict where the transition will take her. There is nothing wrong in exploring different algorithms with weak predictive argumentation, but in order to be constructive this has to be compensated by valid evaluations of the outcomes. Sally's evaluations are based on surface considerations that do not take her far enough. She could for example have tried to understand why the first method only gave one value, why she could not see a minimum on the graphing calculator, why the table gave a smaller maximum, or how equation solutions are related to maxima and minima.

Sally's reasoning is an example of *delimiting AR*, defined as follows.

(1) An algorithm is chosen from a set that is delimited by the reasoner through the algorithms' surface relations to the task. The outcome is not predicted.

(2) The verificative argumentation is based on surface considerations that are related only to the reasoner's expectations of the requested answer or solution. If the implementation does not lead to a (to the reasoner) reasonable conclusion it is simply terminated without evaluation and another algorithm may be chosen from the delimited set.

Delimiting AR is the main AR approach in problematic situations where familiar AR does not work and no guidance source (see below) is at hand. Often it is done in one step, when the reasoner only knows one related algorithm or when the first attempt yields an acceptable conclusion. This can be seen in the perspective that justifying a successful solution by simply describing the algorithm is an accepted sociomathematical norm in most practice and test situations studied.

5.3 Guided AR

When familiar or delimiting AR does not work, one may aim for external guidance. In *text-guided AR* the following conditions hold.

1. The strategy choice concerns identifying surface similarities between the task and an example, definition, theorem, rule, or some other situation in a text source.
2. The algorithm is implemented without verificative argumentation.

70% of the exercises in some common American calculus textbooks may be done with text-guided AR, 20% by essentially copying a solution but with a minor modification by local creative reasoning (Section 6), and 10% require global creative reasoning (Lithner 2004). The latter are generally the most difficult exercises at the end of each exercise section, which few students ever attempt. Similar distributions are found in ongoing studies of Swedish grade 5–12 textbooks. Text-guided AR is the dominating type of student reasoning found in the individual or small-group learning situations (Lithner 2003; Långström and Lithner 2007). Ben's and Sally's task can be solved by copying the following standard textbook example.

Find the maxima and minima of $f(x) = x^2 + 4x + 1$ if $x \in [-3, 0]$.

Solution: $f'(x) = 2x + 4$. Solve $f'(x) = 0$. $2x + 4 = 0 \;\Rightarrow\; x = -2$. Since $f'(x)$ exists for all x and $x = -2$ is its only zero the function is evaluated there and at the endpoints of the interval. $f(-3) = -2$, $f(-2) = -3$, and $f(0) = 1$. Thus we have maximum $f(0) = 1$ and minimum $f(-2) = -3$.

The first step is to identify this as a proper example to copy. In the school textbooks analysed most exercises are preceded by only one example, and then the identification is trivial. At higher educational levels, the theory sections preceding the exercises become longer and may in a university calculus textbook contain ten pages including some ten solved examples. However, the identification in the 70% text-guided AR tasks mentioned above can still be based on surface considerations alone. It is sufficient to know that the expression 'largest and smallest value' in Ben's task is equivalent to 'maxima and minima' in the example above, and to note that both tasks contain quadratic equations on closed intervals. In the strategy implementation (consisting of sub-algorithms), it is required to differentiate and evaluate a quadratic polynomial, and solve a linear equation. Thus a task that is supposed to concern extreme values is reduced to a practice in basic algebra. The reasoning has low logical value since there are no arguments anchored in intrinsic properties of the function, the derivative, or extreme values. Still, the epistemic value is normally regarded as high since, by the didactic contract (Section 9), the role of examples is seen as providing complete templates and the authority of the book is seen as guaranteeing that the algorithm will provide a correct solution.

The other common version of guidance is denoted *person-guided AR*.

1. All strategy choices that are problematic for the reasoner are made by a guide, who provides no predictive argumentation.
2. The strategy implementation follows the guidance and executes the remaining routine transformations without verificative argumentation.

The ninth grade student Moa asks the teacher for help to calculate 15 % of 90. The teacher starts by writing 0.15 multiplied by 90 in Moa's notebook and the complete dialogue is as follows (Lithner 2002).

T:	"What is $5 \cdot 0$?" **M:** "0." [T writes 0]	90
T:	"What is $5 \cdot 9$?" **M:** "45." [T writes 45]	$\cdot$ 0.15
T:	"What is $1 \cdot 0$?" **M:** "0." [T writes 0]	———
T:	"What is $1 \cdot 9$?" **M:** "9." [T writes 9]	450
T:	"What is $5 + 0$?" **M:** "5." [T writes 5]	+ 90
T:	"What is $4 + 9$?" **M:** "13." [T writes 13]	———
T:	"Where shall the decimal point be placed?" **M:** [Silence]	13.50
T:	Marks the decimal point at the correct place and leaves.	

Moa participates in no mathematical activities beyond one-digit arithmetic, which is not her difficulty. There is no argumentation and no intrinsic properties of the problematic components (percentage, the multiplication algorithm, and its relation to the task) or strategy choices are considered. In the graph analogy, the teacher takes the student by the hand and chooses the transition edge at each vertex. As in text-guided AR, validity is not an issue if the guidance is provided by an authority that is seen as guaranteeing a correct solution.

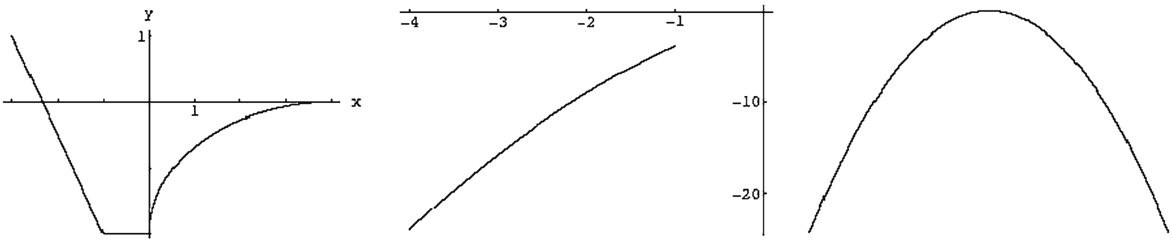
A study of students' collaborative reasoning in their ordinary classroom setting included person-guided AR by peers that had approximately the same level of knowledge as the guided student (Långström and Lithner 2007). The guidance was often unstable and erroneous, but predictive or verificative argumentation was still rarely requested or provided. The students' reasoning was extremely focused on text-guided AR, mainly based on fragmentary searches for surface clues, and their dialogues were seldom mathematically meaningful.

Teacher presentations of task solutions rarely contain true problematic situations, since the teacher normally has prepared them. The framework was used to study if teachers simulate creative reasoning (Section 6) in 23 ordinary teaching situations from grade 7 to university (Bergqvist and Lithner 2005). The teaching mainly focused on algorithmic guidance in three versions, of which the first two are person-guided AR: without comments (as above), with descriptions of the strategy choices, and with explanations of the strategy choices.

6 Creative reasoning

Per is given the graph $f(x)$ in Fig. 3a (Lithner 2000b). His task is to draw a primitive function $g(x)$ to $f(x)$. Per's strategy choice, that correctly implemented would solve the task, is to combine sequences of familiar AR: to find an algebraic expression for the leftmost linear part of $f(x)$ on the interval $[-3, -1]$, integrate it to obtain an algebraic expression for $g(x)$, and finally draw $g(x)$. However, he makes several careless mistakes in the strategy implementation. He reaches after 20 min of laborious work the faulty graph $g(x)$ in Fig. 3b and considers himself finished with the interval $[-3, -1]$. He immediately turns to the next interval, $[-1, 0]$, but the observer interrupts.

JL: "Is the derivative of your $g(x)$ described by the graph of $f(x)$?"
Per: "No, it isn't. My answer is wrong. The derivative of $g(x)$ is positive all the way, but $f(x)$ is not."
JL: "Can you from this reasoning make a rough sketch of what the function should look like?"
Per: "The derivative is first positive, then zero, and then negative."



a b c

Fig. 3 **a** $f(x)$. **b** Per's first, faulty, $g(x)$. **c** Per's second, correct, $g(x)$

Per skilfully makes a rough but reasonable sketch (Fig. 3c) by reading the slope of $g(x)$ point-wise as the function value of $f(x)$.

After Per's initial unsuccessful AR attempt, when 'mildly guided', his reasoning is non-imitative in three central aspects.

1. Per constructs novel reasoning sequences, first when evaluating the graph b and then when forming the graph c. He has not practised any tasks that are similar to the questions I ask him, and though his basic knowledge is used it does not equip him with complete answers or solution algorithms as in MR or AR.
2. There are arguments supporting the logical validity of the reasoning. It is motivated why the first solution (b) is inconsistent with the premises and then what the graph should look like. These arguments are not just explaining something he knew beforehand, but guide the strategy choices and implementations and therefore the constructiveness of the reasoning that results in new knowledge (to Per).
3. Per's arguments are mathematically founded and provide validity through anchoring in intrinsic properties of the components in the reasoning: the relation between the function value and the graph of the function, the relation between the slope and the graph, and that point-wise the slope of $f(x)$ at $x = a$ is equal to the derivative $f'(a)$.

Before defining the reasoning illustrated by Per's solution, a short discussion on logical rigour is required. To some authors 'mathematical reasoning' seems equivalent to strict proof (Duval 2002; Harel 2006) while others include other types of less strict reasoning (NCTM 2000; Ball and Bass 2003). School tasks normally differ from the tasks met by mathematicians, engineers and economists, etc. Within the didactic contract (Section 9) of school it is allowed, and sometimes encouraged, to guess, take chances, and use reasoning with considerably reduced requirements of logical rigour. Even in examinations it is accepted to have only, for example, 50% of the answers correct, while it is absurd if mathematicians, engineers, or economists are correct only in 50% of their conclusions. This framework proposes a wider conception of mathematical reasoning inspired by Pólya (1954): "In strict reasoning the principal thing is to distinguish a proof from a guess, [...] In plausible reasoning the principal thing is to distinguish a guess from a guess, a more reasonable guess from a less reasonable guess." In proof, the value of the reasoning relies on its correctness. Analogously, plausible reasoning is in this framework not necessarily logically strict but constructive through the support of plausible arguments. The stronger the logical value, the more plausible it is.

Creative mathematically founded reasoning (CMR) fulfils all of the following criteria.

1. Novelty. A new (to the reasoner) reasoning sequence is created, or a forgotten one is re-created.
2. Plausibility. There are arguments supporting the strategy choice and/or strategy implementation motivating why the conclusions are true or plausible.
3. Mathematical foundation. The arguments are anchored in intrinsic mathematical properties of the components involved in the reasoning.

CMR does not, as problem solving, have to be a challenge. The definition also includes elementary reasoning. Still, in most studies CMR is rare and AR is dominating. The only exceptions so far are two studies on reasoning requirements in assessment. The official national tests had large proportions of tasks judged to require CMR while such tasks were scarce in teacher-made tests (Palm et al. 2005). CMR was used extensively and successfully in those tasks that required simple CMR, but in difficult CMR tasks unsuccessful delimiting AR was more common (Boesen et al. 2005).

Summarising by relating back to Fig. 2, the transition in CMR is created. In AR it follows a known path, and in MR it is immediate through recollection. Furthermore, in CMR the epistemic value lies in the plausibility and logical value of the reasoning, in MR and AR it is determined by the authority of the source of the imitated information.

7 Thinking processes

Sections 7–9 address the origins and consequences of the reasoning types in the foregoing. The reasoning sequence is in this framework seen as a product of *thinking processes* that are fundamentally different in CMR and in imitative reasoning.

7.1 Creative mathematical thinking

In the context of this framework, the antonym to imitation is creation. 'Creative' is a classic example of a hooray word alongside rational, freedom, etc. "The meanings of hooray words are typically imprecise; they often have contestable definitions, yet all are heavily value-laden" (Huckstep and Rowland 2000, p. 83). Originally associated with ideas in arts, it is nowadays used in a vast variety of settings (e.g. 'creative cooking'). There is a large body of proof literature referring to the creativity of great mathematicians like Poincaré and Hadamard. Sriraman (2004) concluded that research mathematicians followed the Gestalt model of preparation–incubation–illumination–verification. Using similar models (Section 4) in problem solving Schoenfeld (1985) showed that experts are not only more proficient in each stage but also much more flexible in shifting between suitable stages than novices. However, since this framework treats ordinary students' thinking it is insufficient to impute creativity only to experts.

> Although creativity is being associated with the notion of 'genius' or exceptional ability, it can be productive for mathematics educators to view creativity instead as an orientation or disposition toward mathematical activity that can be fostered broadly in the general school population (Silver 1997, p. 75).

There is no single definition of 'creative' generally used in research (Haylock 1997), and there are two major uses of the term: (1) a thinking *process* that is divergent and overcomes fixation and (2) a *product* that is perceived as creative for some reason, e.g. works of art. Haylock sees two types of process fixation. Content universe fixation limits the range of elements seen as appropriate for applications to a given problem. Algorithmic fixation is the repeated use of an initially successful algorithm that becomes an inappropriate fixation. Silver (1997) suggests a view of creativity in which the thinking processes are related to deep, flexible knowledge in content domains and associated with long periods of work and reflection rather than rapid and exceptional insights. Silver sees fluency, flexibility and novelty as key qualities of creativity. Haylock emphasises originality and flexibility, but argues that a limited notion of fluency is less useful. If asked to generate questions with the answer 4, the pupil might start with '5–1', '6–2', '7–3', and continue this indefinitely. This may score high on fluency if it is only seen as the number of acceptable responses but not on flexibility, which is the number of different responses.

Ordinary students' task solving creativity is thus based on thinking processes that are flexible, fluently admit different approaches and adaptations to the situation, and are not hindered by fixation. One may note that the plausibility condition in the CMR definition excludes products of purely impulsive (e.g. brainstorming) or affective (e.g. creating art based on feelings) creative thinking processes.

As in Sally's solution, in Per's first algorithmic attempt his thinking is too narrow and lacks flexibility even though he clearly notes that his reasoning is awkward. The guidance initiates a thinking process better adapted to the situation, the creation of a qualitative evaluation of the task components that can support the validity of the reasoning.

7.2 Superficial and imitative thinking

Vinner (1997) proposes a diagrammatic model for analytical routine task solving where one needs a pool of algorithms, mental schemes that identify the task type and its structure, and mental schemes that assign a suitable algorithm to the task type. Two cognitive faculties are involved, the abilities to identify similarities and to imitate. These correspond in AR to the strategy choice and implementation respectively. A main drawback with imitative reasoning is that the analytical and conceptual thinking processes may be missing, which is not possible in CMR. The effect is that the strategy choice and implementation lack analytical support and become haphazard. Furthermore, monitoring, control, and correction of careless mistakes are impossible if the thinking is strictly imitative.

A large part of the research results on reasons behind learning and achievement difficulties can be characterised as an unwarranted and too extensive reduction of ambiguity, risk and complexity of mathematical concepts and processes by teachers, textbook writers, and students trying to cope with curricular goals that are hard to reach (Doyle 1988; Schoenfeld 1991). McGinty et al. (1986) found that the number of word problems in American grade 5 textbooks from 1924, 1944, and 1984 had decreased and they had become less rich, while the number of drill tasks had increased. In Australia algebra instructions and exercises were reduced to easier arithmetic (Stacey and MacGregor 1999) and in Sweden (using American textbooks) university calculus was reduced to grade 10 algebra (Lithner 2003, 2004). In Vinner's (1997) framework two of the main notions are pseudo-conceptual and pseudo-analytical reasoning, which in routine task solving might give the impression of being conceptual/analytical. Students' difficulties are then better understood if they are interpreted within a 'non-cognitive' framework, than if they are seen as misconceptions within the domain of meaningful contexts. Because of the didactic contract (Section 9) students may, consciously or not, try to please the educator with superficial behaviour that they believe is acceptable.

The most frequent type of reduction of complexity seems to be to focus on actions in the form of routine procedures (Tall 1996; Vinner 1997) that may prevent the development of conceptual and relational understanding (Skemp 1978; Hiebert and Carpenter 1992). In a research literature survey, Hiebert (2003) finds massive amounts of converging data showing that students know some basic elementary skills but there is not much depth and understanding. "One can model their behaviour and predict the errors they will make by looking only at the symbol manipulation rules they have been taught and pretending that they are following these rules like robots with poor memories" (p. 12). Memorised operations without structural conceptions can only be stored in unstructured, sequential cognitive schemata, which are inadequate for the modest dimensions of human working memory (Sfard 1991). The reduction of complexity may thus go so far that the imitative thinking is detached from its mathematical meaning, as in the MR and AR examples in the foregoing. In this case it is not constructive to analyse students' reasoning only by characterising its mathematical coherence. The framework should thus in addition to CMR capture non-cognitive means of trying to cope: attempts to guess and to find familiar surface clues for action, and the need to meet the expectations of the teacher or researcher (Leron and Hazzan 1997). In this perspective the delimiting AR by Sally above is a quite reasonable, however unwanted, outcome of imitative thinking processes.

There are other thinking processes that may hinder CMR. For example, strategy choices may be guided by erroneous intuitions that seem self evident to the individual so that no need for argumentation is perceived (Fischbein 1999). Other aspects concern difficulties with abstraction processes (Monaghan and Ozmantar 2006) and unclear influences from tacit knowledge (Ernest 1999), but it is beyond the scope of this paper to elaborate further in these directions.

8 Student competencies

There are many possible ways to characterise the mathematical knowledge used and learnt in reasoning, but it is insufficient to describe it only in terms of mathematical content. The NCTM Principles and Standards (NCTM 2000) complements its five *content standards* (number and operations, algebra, geometry, measurement, and data analysis and probability) with five *process standards* (problem solving, reasoning and proof, connections, communication, and representation). A framework by Niss (2003) contains factors similar to the NCTM process standards but denotes them *competencies*. A competence is the ability to understand, judge, do, and use mathematics in a variety of mathematical contexts and situations. Three competencies are particularly relevant for this framework: problem solving abilities, reasoning abilities and conceptual understanding. The first two are self-evident since the framework is about task solving reasoning. Conceptual understanding is emphasised in CMR by its presence in anchoring.

The problem solving competence includes identifying, posing, and specifying different kinds of problems and solving them, if appropriate, in different ways (Niss 2003). Schoenfeld (1985) formed through a series of empirical studies the probably most cited problem solving framework based on four key competencies[1] (p. 15). *Resources* (basic knowledge), *Heuristics* (rules of thumb for non-standard problems), *Control* (metacognition: monitoring and decision-making), and *Belief Systems* (one's

[1]Schoenfeld did not use the term 'competence' but 'knowledge and behaviour'.

mathematical world view). He found that novices often had sufficient resources but were lacking in the other three competencies.

The NCTM Standards (2000) recognize reasoning and proof as fundamental aspects of mathematics. “People who reason and think analytically tend to note patterns, structure, or regularities in both real-world situations and symbolic objects; they ask if those patterns are accidental or if they occur for a reason; and they conjecture and prove” (p. 56). The reasoning competence goes beyond constructing reasoning, and includes abilities like following and assessing chains of arguments, knowing what a proof is and how it differs from other kinds of reasoning, uncovering the basic ideas in a given line of argument, and devising formal and informal arguments (Niss 2003).

The concept of understanding is very complex (Sierpinska 1996), and will not be pursued here beyond noting that several of the theoretical constructs concern relations between rote learning and deeper understanding. Skemp (1978) distinguishes between ‘instrumental understanding’ and ‘relational understanding’ of mathematical procedures. The former can be apprehended as ‘true’ (relational) understanding, but is only the mastering of a rule or procedure without any insight in the reasons that make it work. A similar distinction, between ‘action’ and ‘process’ is made by Asiala et al. (1996), and Hiebert and Lefevre (1986) distinguish between conceptual and procedural understanding.

Several competencies are central in CMR. Resources, including knowledge and understanding of the properties of the components (objects, transformations and concepts) involved in the reasoning, are required for anchoring. A battery of heuristic approaches is the basis for flexibility. A key characteristic of the common situation where different versions of AR fail in problematic situations, often due to careless mistakes, is that there are no metacognitive activities that can identify and remediate the problems. In line with the results of Schoenfeld (1985), the beliefs about mathematics are often such that students do not even attempt CMR, even in situations where it relatively easily could have led to considerable progress. For example in Per’s case the guidance does not provide the knowledge resources required for the CMR: he already has them but fails to utilise them in his first algorithmic attempt. The heuristic approach he takes after the guidance is “reformulating the problem by assuming you have a solution and determining its properties” (Schoenfeld 1985, p. 109).

9 The milieu

To understand why a particular reasoning type is used it is necessary to consider the learning environment in which the competencies are formed. Brousseau’s (1997) theory of didactical situations emphasises “the social and cultural activities which condition the creation, the practice and the communication of knowledge” (p. 23). It specifies conditions for learning through task solving and uses this characterisation to explain why and how rote learning appears.

9.1 Learning creative reasoning

The *milieu* is “everything that acts on the student or that she acts on” in a learning situation (ibid., p. 9). *Didactique* studies the communication of knowledge and one central aspect of Brousseau's didactical situations is the *devolution of problems*. The student has to take responsibility for a part of the problem solving process, but she cannot in general learn in isolation. The teacher's task is to arrange a suitable didactic situation in the form of a problem. Between when the student accepts the problem as her own and the moment when she produces her answer, the teacher refrains from interfering and suggesting the knowledge that she wants to see appear. This part of the didactic situation is called an *adidactical situation*. The student must construct the piece of new knowledge and the teacher must therefore arrange not the communication of knowledge, but the devolution of a good problem. If the student avoids or does not solve the problem, the teacher has the obligation to help. Then a relationship is formed that (mainly implicitly) determines what each party will be responsible for: the *didactic contract* that ensures the functioning of the process. Wedege and Skott (2006) note that the term 'didactic contract' is used outside France as a metaphor for the set of implicit and explicit rules of social and mathematical interaction in a particular classroom, which is an extension outside Brousseau's didactical situations and more in line with a definition by Balacheff (1990).

Temporarily incomplete or faulty conceptions in the form of *obstacles* are in Brousseau's theory not in general seen as failures but are often inevitable and constitutive of knowledge. An obstacle produces correct responses within a particular, frequently experienced context but not outside it and may withstand both occasional contradictions and the establishment of a better piece of knowledge. Clarifying obstacles helps the student see the necessity for learning, not by explaining what the obstacle is but to help her discover it. Good problems will permit her to overcome the obstacles. The teacher may (e.g. to reduce complexity, Section 7.2) try to overcome the obstacle and force learning by devolving less of the problem to the student. Brousseau exemplifies this by the Topaze effect (p. 25) when the teacher lets the teaching act collapse by taking responsibility for the student's work and letting the target knowledge disappear (as in the examples of guided AR above). Telling the student that an automatic method exists relieves her of the responsibility for her intellectual work, thus blocking the devolution of a problem. If this is the normal didactic situation the student meets then the didactical contract is formed accordingly, which may not be the teacher's intention. The teacher expects the student to learn problem solving reasoning, while the student expects that an algorithm should be provided that relieves her of the responsibility of engaging in the adidactical situation. This avoids dealing with the obstacle that can therefore become insurmountable.

9.2 Learning imitative reasoning

Vinner (1997) assumes that pseudo conceptual behaviours exist because they are easier. Another starting point is taken below, that students may learn what they are given opportunity to learn (Hiebert 2003).

There is a pressure on textbooks to be self-contained (so students do not have to ask the teacher many questions) by providing guiding examples to most exercises. Therefore students may develop strategy choice approaches where the question 'what algorithm should be applied?' is immediately asked. When the textbook is at hand, this question leads to text-guided AR. In tests, the same question leads to a memory search for earlier situations that are experienced as similar. If the student does not understand the components involved and has no real practice in creative problem solving, the only alternative may be to search for superficial familiarities. That is, an individual reduction of complexity when trying to cope in a difficult situation. The main alternative seems to be to ask the teacher for help and perhaps receive it in the form of teacher-guided AR. The help students obtain from their peers is also dominated by direct algorithm descriptions, but then often in incorrect ways. One memorises algorithms through guided AR and does not learn much beyond familiar and delimiting AR.

Thus the lack of problem devolution in the milieu leads to the situation that the obstacles of learning CMR are not addressed and therefore not overcome. This yields that relevant reasoning, problem solving, and understanding competencies are not developed. The thinking processes the student manages to activate are superficial and imitative. Thus it may be that students are neither looking for the correct nor the easiest way to solve a task, but for the most plausible within their competence. And this is, according to their knowledge and beliefs, to find 'the' algorithm.

In Per's first solution attempt he struggles with the algorithms and never considers attempting CMR. He says in the post-interview that the algorithmic methods (that failed) "feel safer" (a common view among the students) than the correct CMR he constructs when mildly guided. To Per, the epistemic value is higher in AR though the logical value is clearly higher in his CMR.

10 Concluding reflections

A fundamental issue in research is how well the claims are warranted by the evidence presented (Lester 2005; Schoenfeld 2007). Churchman (1971) classified inquiry systems into five categories distinguished by the source of evidence: reasoning, observation, representation, dialectic, and ethical values and practical consequences. In this framework, as in most research, the primary source of evidence is observation of data, complemented by reasoning and references to the research literature. The discussion below relates to three questions posed by Schoenfeld (2007): Why should one believe what the author says (trustworthiness)? What situations does the research apply to (generality)? Why should anyone care (importance)?

10.1 Trustworthiness

The two main claims are that the framework can help to (1) characterise important aspects of reasoning and (2) explain the origins and consequences of the reasoning types. The main warrant in the first case consists of direct references to data. It has been possible to relate critical incidents with respect to success and failure in task solving situations, to the argumentation about strategy choices and strategy implementations, which in this framework defines the reasoning types. Rigour in this endeavour is sought for by providing specific definitions of the notions involved, and by following specified analysis procedures. Fairly extensive descriptions of data and analyses have been presented in separate studies, so that the warrants are open to criticism. Concerning the explanations (2) there are also direct inferences from data but they are to a larger extent complemented by references to theories, frameworks and notions in the research literature. A key feature of the framework is that it can support explanations by relating data from different parts of the learning environment (e.g. teaching, textbooks, assessment) to students' use of reasoning in learning and task solving.

Trustworthiness is increased by triangulation. For example, observations of a student solving a task have been complemented by analyses of the student's written solution, practice tasks and textbooks, and pre- and post-interviews. The claims are also warranted by comparisons with other research studies.

10.2 Generality

Generality is supported by clarifying the contexts of the studies. This is impossible to do in full detail, especially in qualitative studies, but can give some ideas about which other contexts the results apply to. The framework is so far primarily developed in teaching, learning, and application situations where AR is common. Sweden's educational system is guided by official national curricula documents that contain goals similar to the NCTM Principles and Standards (NCTM 2000). However, the Swedish goals are vaguely and concisely written and intended to be implemented through the interpretations of individual teachers. Several national studies conclude that the relatively uniform mathematics textbooks in reality guide the teaching.

Representativeness has been ranging from relatively uncertain in studies where a handful of students participated, to higher in studies with statistically accepted data selections (e.g. a nation-wide random selection of teacher-made tests).

10.3 Importance

The rote learning problem is severe, widespread, and well known, but still far from resolved (Hiebert 2003). It seems that the gravity of the problem, as a main cause behind learning difficulties, is not fully apprehended by students, teachers, textbook writers, syllabus constructers, administrators, politicians, and perhaps also among many researchers. One reason may be that we lack terminology and frameworks to communicate the insights in more specific ways. Just to say that students learn by rote and need to practice creative reasoning will not help much, if we cannot specify the problem and suggest well-founded measures.

By focusing on reasoning, the framework makes the algorithmic character transparent in the analyses and provides evidence that the students' AR focus is to a large extent caused by insufficiencies in the learning environment (under the assumption that learning is strongly affected by the opportunity to learn). The AR focus will lead to short term gains like managing textbook exercises and passing examinations that are adapted to AR. It is shown how it is possible to get really far by superficial AR strategies in exercises and in tests. But this leads to long term losses: The search for algorithms *becomes* mathematics instead of being a part of it. The aspects that can make mathematics truly meaningful to the students, such as conceptual understanding, creative reasoning, and insights in the central roles of mathematics in our society are not enhanced by rote learning.

In the introduction the framework was labelled 'use-inspired basic research'. Research has been in focus, but it has also been applied in some ways that are directly useful in practice. For example in exercise and assessment task construction it is possible to judge what kind of reasoning is required, which correlates strongly with the reasoning that students actually will use. There are many development projects including ones utilising frameworks like NCTM Principles and Standards (NCTM 2000). However, due to its character and vast scope the NCTM document lacks specific definitions of many concepts, which exemplifies the need for local specialised frameworks. The framework of this paper has not (yet) been used in curriculum development, but conceptual frameworks can be valuable, together with instructional design frameworks, in guiding construction and evaluation. Lester (2005) claims that "without a framework, the researcher can speculate at best or offer no explanation at all" (p. 461). This quote concerns interpreting research findings, but the situation is similar in instructional design. Without a framework we have to rely only on intuition, experience and common sense. This can take us far, and indeed it often does. But without a framework guiding our constructions or focussing our evaluations, we will never really know exactly what we are doing and why it failed, or why it worked so well.

References

Asiala, M., Brown, A., DeVries, D., Dubinsky, E., Mathews, D., & Thomas, K. (1996). A framework for research and curriculum development in undergratduate mathematics education. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education II, CBMS issues in mathematics education* (pp. 1–32). American Mathematical Society.

Balacheff, N. (1988). Aspects of proof in pupils practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216–235). London: Hodder and Stoughton.

Balacheff, N. (1990). Towards a problematique for research on mathematics teaching. *Journal for Research in Mathematics Education*, *21*(4), 258–272.

Ball, D., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27–44). Reston, VA: National Council of Teachers of Mathematics.

Bergqvist, T., & Lithner, J. (2005). *Simulating creative reasoning in mathematics teaching.* Research Reports in Mathematics Education 2, Dept. of Mathematics, Umeå University.

Bergqvist, T., Lithner, J., & Sumpter, L. (2007). Upper secondary students task reasoning. *International Journal of Mathematical Education in Science and Technology* (in press).

Bloom, B. S. (Ed.) (1956). *Taxonomy of Educational Objectives, Part 1: The Cognitive Domain*. New York: David McKay.

Boesen, J., Lithner, J., & Palm, T. (2005). *The mathematical reasoning required by national tests and the reasoning actually used by students.* Research Reports in Mathematics Education 4, Dept. of Mathematics, Umeå University.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluwer.

Churchman, C. W. (1971). *The design of inquiring systems: Basic concepts of system and organization*. New York: Basic Books.

Doyle, W. (1988). Work in mathematics classes: The context of students thinking during instruction. *Educational Psychologist*, *23*, 167–180.

Duval, R. (2002). Proof understanding in mathematics: What ways for students? *Proceedings of 2002 international conference on mathematics: Understanding proving and proving to understand* (pp. 61–77).

Ernest, P. (1999). Forms of knowledge in mathematics and mathematics education: Philosophical and rhetorical perspectives. *Educational Studies in Mathematics*, *38*(1), 67–83.

Fischbein, E. (1999). Intuitions and schemata in mathematical reasoning. *Educational Studies in Mathematics*, *38*(1), 11–50.

Harel, G. (2006). Mathematics education research, its nature and its purpose: A discussion of Lester's paper. *Zentralblatt fuer Didaktik der Mathematik*, *38*(1), 58–62.

Haylock, D. (1997). Recognising mathematical creativity in schoolchildren. *Zentralblatt fuer Didaktik der Mathematik*, *29*(3), 68–74.

Hegarty, M., Mayer, R., & Monk, C. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, *87*(1), 18–32.
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: National Council of Teachers of Mathematics.
Hiebert, J., & Carpenter, T. (1992). Learning and teaching with understanding. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 65–97). New York: Macmillan.
Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge* (pp. 1–27). Hillsdale, NJ: Erlbaum.
Huckstep, P., & Rowland, T. (2000). Creative mathematics – real or rhetoric? *Educational Studies in Mathematics*, *42*(1), 81–100.
Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ: Erlbaum.
Leron, U., & Hazzan, O. (1997). The world according to Johnny: A coping perspective in mathematics education. *Educational Studies in Mathematics*, *32*, 265–292.
Lester, F. (2005). On the theoretical, conceptual, and philosophical foundations for research in mathematics education. *Zentralblatt fuer Didaktik der Mathematik*, *37*(6), 457–467.
Lithner, J. (2000a). Mathematical reasoning and familiar procedures. *International Journal of Mathematical Education in Science and Technology*, *31*, 83–95.
Lithner, J. (2000b). Mathematical reasoning in task solving. *Educational Studies in Mathematics*, *41*, 165–190.
Lithner, J. (2002). Lusten att lära, Osby (The motivation to learn, Osby). *Skolverkets nationella kvalitetsgranskningar* (*Quality inspections of the Swedish National Agency for Education*), in Swedish.
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.
Långström, P., & Lithner, J. (2007). Students learning strategies and mathematical reasoning in textbook-related metacognitive processes (in press).
McGinty, R., VanBeynen, J., & Zalewski, D. (1986). Do our mathematics textbooks reflect what we preach? *School Science and Mathematics*, *86*, 591–596.
Monaghan, J., & Ozmantar, M. (2006). Abstraction and consolidation. *Educational Studies in Mathematics*, *62*(3), 233–258.
National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: The Council.
Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. *Third Mediterranean conference on mathematics education* (pp. 115–124).
Palm, T., Boesen, J., & Lithner, J. (2005). *The requirements of mathematical reasoning in upper secondary level assessments.* Research Reports in Mathematics Education 5, Dept. of Mathematics, Umeå University.
Pirie, S., & Kieren, T. (1994). Growth in mathematical understanding: How can we characterise it and how can we represent it? *Educational Studies in Mathematics*, *26*, 165–190.
Pólya, G. (1954). *Mathematics and plausible reasoning* (Vols. I and II). Princeton, NJ: Princeton University Press.
Schoenfeld, A. (1985). *Mathematical problem solving*. Orlando, FL: Academic.
Schoenfeld, A. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J. Voss, D. Perkins, & J. Segal (Eds.), *Informal reasoning and education* (pp. 311–344). Hillsdale, NJ: Erlbaum.
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.
Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, *34*(1), 4–36.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*, 1–36.
Sierpinska, A. (1996). *Understanding in mathematics*. Routledge: Falmer.
Silver, E. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *Zentralblatt fuer Didaktik der Mathematik*, *29*(3), 75–80.
Skemp, R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher*, *26*(3), 9–15.
Sriraman, B. (2004). The characteristics of mathematical creativity. *The Mathematics Educator*, *14*(1), 19–34.
Stacey, K., & MacGregor, M. (1999). Taking the algebraic thinking out of algebra. *Mathematics Education Research Journal*, *11*, 24–38.
Tall, D. (1991). Reflections. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 251–259). Dordrecht: Kluwer.
Tall, D. (1996). Functions and calculus. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 289–325). Dordrecht: Kluwer.
Tall, D. (2004). Thinking through three worlds of mathematics. In M. J. Høines & A. B. Fuglestad (Eds.), *28th Conference of the international group for the psychology of mathematics education.* vol. 4 (pp. 281–288). Bergen.
van Hiele, P. (1986). *Structure and insight. A theory of mathematics education*. Orlando: Academic.
Vinner, S. (1997). The pseudo-conceptual and the pseudo-analytical thought processes in mathematics learning. *Educational Studies in Mathematics*, *34*, 97–129.
Wedege, T., & Skott, J. (2006). *Changing views and practices? A study of the Kapp-Abel mathematics competition*. Trondheim: NTNU.
Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458–477.
Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 227–236). Reston, VA: National Council of Teachers of Mathematics.