Abstract or Concrete Examples in Learning Mathematics?
A Replication and Elaboration of Kaminski, Sloutsky, and Heckler's Study

Dirk De Bock
Hogeschool-Universiteit Brussel, Brussels, Belgium,
and Katholieke Universiteit Leuven, Leuven, Belgium

Johan Deprez
Hogeschool-Universiteit Brussel, Brussels, Belgium;
Katholieke Universiteit Leuven, Leuven, Belgium;
and University of Antwerp, Antwerp, Belgium

Wim Van Dooren
Katholieke Universiteit Leuven, Leuven, Belgium

Michel Roelens
Katholieke Hogeschool Limburg, Diepenbeek, Belgium

Lieven Verschaffel
Katholieke Universiteit Leuven, Leuven, Belgium

Kaminski, Sloutsky, and Heckler (2008a) published in *Science* a study on "The advantage of abstract examples in learning math," in which they claim that students may benefit more from learning mathematics through a single abstract, symbolic representation than from multiple concrete examples. This publication elicited both enthusiastic and critical comments by mathematicians, mathematics educators, and policymakers worldwide. The current empirical study involves a partial replication—but also an important validation and extension—of this widely noticed study. The study's results confirm Kaminski et al.'s findings, but the accompanying qualitative data raise serious questions about their interpretation of what students actually learned from the abstract concept exemplification. Moreover, whereas Kaminski et al. showed that abstract learners transferred what they had learned to a similar abstract context, this study shows also that students who learned from concrete examples transferred their knowledge into a similar concrete context.

Key words: Algebra; College/university; Learning; Research issues; Testing

Since the end of the "new math" era, many mathematics educators take for granted that mathematics should be taught "from concrete to abstract" and that a series of well-chosen examples can facilitate students' understanding of an underlying, or more general, mathematical idea. In most countries, this guiding design principle is currently implemented in textbooks and other educational resources in which examples (and counterexamples) precede formal definitions of concepts and statements of theorems. In Realistic Mathematics Education (RME), this idea is reflected in the principle of "conceptual mathematization" (de Lange, 1987; Gravemeijer, 1994): A general concept (e.g., the derivative of a function as the limit of a difference quotient) is extracted from several more concrete (and often realistic) instantiations (e.g., instantaneous velocity, marginal cost, slope of a curve, or growth rate). According to the RME philosophy, this approach significantly increases the chance that the new mathematical concept will be learned meaningfully and will be positively transferred to novel situations.

The issue of (positive) transfer of learned concepts, procedures, and solution methods to new situations is a persistent theme within cognitive and educational psychology. It was, and is, a major thematic issue in the three major general views of cognition, learning, and teaching (Greeno, Collins, & Resnick, 1996; Mayer & Wittrock, 1996). It was already intensively analyzed and discussed by associationist and behaviorist psychologists in the 1st decades of the previous century (see, e.g., Thorndike et al., 1924), for whom it involved the application of specific identical elements of behavior from an initially learned task to a new task. It continued to be a central theme in Gestalt and cognitive theories (see, e.g., Gick & Holyoak, 1983; Wertheimer, 1959), in which transfer was viewed as the (metacognitively driven) recognition and use of previously learned concepts, principles, or specific or general problem-solving methods in new situations. And it was problematized and reconceptualised by adherents of the situative and sociohistoric perspective, for whom it involves an "attunement" to the affordances and constraints of the material artifacts and social environments that are invariant between learning and transfer situations (Greeno, Smith, & Moore, 1993; Lobato, 2003).

Although these views differ in their descriptions of exactly what transfers, how transfer occurs, why its occurrence is so difficult, and how transfer can be optimally enhanced through instruction, they all emphasize, in one way or another, the importance of the selection of the examples to which students are exposed during the learning process. Accordingly, the role of "exemplification" (i.e., the use and analysis of examples, illustrations, occurrences, and instances of a concept as a particularly powerful tool for teaching and learning) continues to be a central research topic in the mathematics education community (Bills et al., 2006; Watson & Mason, 2002).

Recently, the discussion on the role of "practical" examples for learning mathematics was (re-)opened in several countries, both in the mathematics education communities and in society more broadly. The immediate cause was a series of papers by Kaminski, Sloutsky, and Heckler (2005, 2006a, 2006b, 2008a, 2009), among which is a much-discussed one in *Science* (Kaminski et al., 2008a). These

papers were mostly based on Kaminski's (2006) dissertation, in which students' need for concrete instantiations to learn abstract concepts was explicitly questioned. Based on controlled experiments with undergraduate students, Kaminski et al. (2008a) came to a conclusion, which goes against what is now often taken for granted in the mathematics education community:

> If a goal of teaching mathematics is to produce knowledge that students can apply to multiple situations, then presenting mathematical concepts through generic instantiations, such as traditional symbolic notation, may be more effective than a series of "good examples." This is not to say that educational design should not incorporate contextualized examples. What we are suggesting is that grounding mathematics deeply in concrete contexts can potentially limit its applicability. Students might be better able to generalize mathematical concepts to various situations if the concepts have been introduced with the use of generic instantiations. (p. 455)

In a "math wars" climate, the publication in *Science* received widespread attention in newspaper articles (e.g., Chang, 2008; "Les exemples," 2008; "Abstracte wiskunde," 2008), in our view, dramatically overstating its relevance to K–12 educational practice (cf. infra) but boosting the public debate on how mathematics should be taught and learned. In the specialized scientific circuit, several critical comments on Kaminski's work were published (e.g., Jones, 2009a, 2009b; Podolefsky & Finkelstein, 2009), but, as far as we know, these critiques were never supported by new empirical data. In this article, we elaborate on two main elements of critique, and we provide empirical evidence to substantiate them. First, we argue for why the comparison made by Kaminski et al. (2008a) is basically unfair. Second, we query about what the students actually learned of the abstract examples. To support these two elements of critique empirically, we set up a replication and extension study.

In the first part of this article, we describe Kaminski et al.'s (2008a) central experiment and its basic conclusions. Next, we elaborate on two major elements of critique, given by other commentators, which motivated our replication and extension study. Then, we report the design and main results of our empirical study. Finally, in the light of the expressed critiques and of our new empirical data, we discuss the limited generalizability and utility of Kaminski's—and thus also our—study for mathematics education practice. In particular, we re-examine the validity of Kaminski's main claim that students may benefit more from learning mathematics through a single abstract, symbolic representation than from multiple concrete examples of a to-be-learned concept.

KAMINSKI'S CENTRAL EXPERIMENT

Kaminski et al. (2008a) addressed the question of whether learning a mathematical concept starting from multiple concrete instantiations is the most efficient route to promoting transfer of knowledge. They doubt the "taken for granted" belief that well-chosen concrete contexts can facilitate students' understanding of an underlying abstract mathematical concept, because "instantiating an abstract concept in concrete contexts places the additional demand on the learner of ignoring irrelevant,

salient superficial information, making the process of abstracting common structure more difficult than if a generic instantiation were considered" (Kaminski, 2006, p. 114). More concretely, they tested the "hypothesis that learning a single generic instantiation (that is, one that communicates minimal extraneous information) may result in better knowledge transfer than learning multiple concrete, contextualized instantiations" (Kaminski et al., 2008a, p. 454). Therefore, a series of experiments was conducted.

Four of these experiments are reported in Kaminski et al. (2008a), but here, we focus on the first—and central—experiment, which is the subject of the commentators' main critiques. This experiment consisted of two phases. In the first phase, 80 undergraduate students from an introductory psychology course learned either an abstract (or "generic") instantiation of a mathematical concept, or one or more concrete instantiations of that concept. At the end of this phase, a learning test was administered. In the second phase, which took place immediately after the first one, participants were tested on a (mathematically) isomorphic transfer domain. The targeted mathematical concept was that of a group of order 3: a set of three elements with an operation that satisfies the properties (or axioms) of closure, associativity, existence of identity, and existence of inverses. All groups with three elements are isomorphic, which means that, from a mathematical point of view, there exists only one such group (and this group is commutative and cyclic). This fact was not explicitly taught to the participants (and it is not likely that this or other facts about finite groups were part of their prior knowledge). In the abstract-instantiation condition, participants learned to combine arbitrary symbols (flag, diamond, and circle) that were inscribed on a tablet discovered on an archaeological dig (e.g., see Figure 1, first row: Circle followed by diamond yields flag). In one of the concrete-instantiation conditions, participants learned to combine 1/3, 2/3, and full liquid cups in a way that the result is the left-over or "remainder" (e.g., see Figure 1, second row: a 2/3-filled cup followed by a 2/3-filled cup yields a 1/3-filled cup). Immediately after the learning phase, during which some students worked with the abstract instantiation and others with one or more concrete instantiations, all participants were engaged in the same transfer task about a fictitious children's
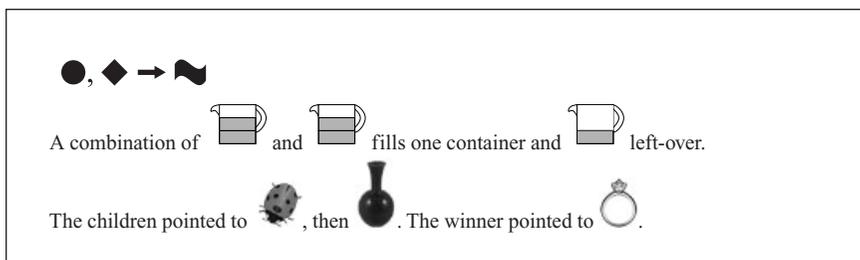
Figure 1. Combination rules presented in the generic- (first row), in one of the concrete- (second row), and in the transfer-instantiation (third row) conditions by Kaminski et al. (2008b, pp. 9, 15, 26).

game. In this game, one child pointed successively to two or more different objects from a set of three (ring, vase, and ladybug) and another child had to point to the resulting object (e.g., see Figure 1, third row: Vase followed by ladybug yields ring). After a short introduction, participants were told that "the game rules were like the rules of the system(s) they just learned and they need to figure them out by using their newly acquired knowledge (i.e., transfer)" (Kaminski, Sloutsky, & Heckler, 2008b, p. 3). Kaminski et al. (2008a) found that, on the transfer task, participants from the abstract-instantiation condition outperformed participants from the concrete-instantiation condition and concluded: "If transfer from multiple instantiations depends on whether the learner abstracts and aligns the common structure from the learned instantiations . . . then transfer failure suggests that participants may have been unable to recognize and align the underlying structure" (p. 455). This conclusion was strengthened by one of Kaminski et al.'s (2008a) follow-up experiments, in which students learned either an abstract instantiation or a concrete instantiation that was followed by an abstract instantiation. The results of that follow-up experiment showed that, on a transfer task, participants who learned only the abstract instantiation outperformed those who learned both a concrete and an abstract instantiation.

CRITICAL COMMENTS ON KAMINSKI'S CENTRAL EXPERIMENT

We summarize the critical comments that inspired us to set up an elaboration and extension study. A first group of related elements of critique refers to the unfairness of Kaminski et al.'s (2008a) comparison due to the role of uncontrolled variables. A second critique, which received less attention by other commentators so far, refers to what students actually learned about the to-be learned concept (a group of order 3). Critiques that are not directly relevant for the rationale of our empirical study are not developed in this part (but some of them will be addressed in our final discussion).

In order to avoid that differences in learning outcomes could be attributed to variables different from the type of instantiation, Kaminski (2006) controlled for superficial similarity between the transfer task and the abstract or concrete learning task by asking a sample of participants (not those in the central experiment) to read descriptions of the learning and transfer domains and then indicate a level of similarity. For both learning conditions, average ratings of similarity were rather low (around 3 on a scale from 1 to 5), and no significant differences in similarity were found. Although agreeing on the absence of *superficial* similarity between the learning domains and the transfer domain, several authors (cf. infra) point to substantial differences in similarity at a *deeper* level.

One of those substantial differences refers to the presence or absence of a physical referent. In the concrete learning domain, the symbols represent physical quantities and the operation of combining these symbols has a physical connotation as well. Both the transfer and abstract learning domain lack this physical interpretation: Here symbols are arbitrary and are combined in ways purely governed by a set of formal rules (Jones, 2009a, 2009b). Hence, only abstract learners had the

experience of performing purely formal operations with meaningless symbols, necessary for success in the transfer test. As Podolefsky and Finkelstein (2009) suggest, those students have learned to ignore potential links to prior knowledge, whereas concrete learners got the message that prior knowledge counts: The quantities that appeared in the learning task could be used as a sense-making aid. So, plausibly, concrete learners struggled when they had to discern patterns in a transfer domain in which any physical interpretation was impossible. A related critique (Cutrona, 2008; Deprez, 2008; Deprez et al., 2010; McCallum, 2008; Mourrat, 2008) refers to the supporting role of numbers in the concrete learning domain, whereas numbers are absent in the transfer and abstract learning domains.

The link with numbers has an important repercussion in terms of *what* is probably learned in the different learning domains. Whereas in the transfer task and in the abstract learning domain the targeted concept truly is that of a mathematical group, the examples in the concrete learning domains primarily aim at modular addition. As Cutrona (2008) remarks, these "examples seem geared to teaching the broader, generalized concept that includes mod 4, mod 5, and so on" (p. 1632). Modular addition is a richer concept (i.e., less general, having more structure) than that of a group. For systems of order 2 and 3, both concepts coincide, but already for order 4, two different groups exist: the group based on addition modulo 4 (in symbolic, additive notation: the set $\{n, a, b, c\}$ in which n denotes the identity or neutral element, equipped with a commutative and associative binary operation, $+$, satisfying $a + a = b$, $a + a + a = c$,[1] and $a + a + a + a = n$) and the group of symmetries of a (nonsquare) rectangle (the set $\{n, a, b, c\}$ with a commutative and associative binary operation, $+$, satisfying $a + a = n$, $b + b = n$, $a + b = b + a = c$). With increasing order, these concepts depart further and further from each other (except for prime orders). Even if modular addition and groups coincide for order 3, they have a different "look and feel," or, as argued by Cutrona (2008), the examples in the concrete and abstract learning domains serve different goals.

According to McCallum (2008), this difference in targeted object can be translated into an uncontrolled structural difference. In the transfer and abstract learning domain, the underlying structure can be described as follows. One of the elements in the group is the identity and the other two behave symmetrically: Adding each to itself yields the other and adding the two of them yields the identity. In symbolic notation, we have a group $\{n, a, b\}$, for which a and b satisfy the following relations:

$$a + a = b \tag{1.1}$$

$$b + b = a, \text{ and} \tag{1.2}$$

$$a + b = b + a = n. \tag{1.3}$$

Here, the group is presented by two generating elements a and b. Relations 1.1 to 1.3 are exactly the rules in Kaminski's (2006) summary at the end of the learning phase. In the concrete learning domains, the same rules are given. Hence,

---

[1]The associative law guarantees that expressions with three or more terms also are well defined.

superficially, the structure is exactly the same. However, the quantities or numbers in these domains implicitly communicate a *different* underlying structure, in which the symmetry between the elements *a* and *b* is broken. Now we have the group $\{n, a, b\}$, for which

$$a + a = b \text{ and} \tag{2.1}$$

$$a + a + a = n. \tag{2.2}$$

This alternative structure stresses the cyclic nature of the group, that is, it is generated by only one of its elements (which is related to addition modulo 4 discussed in the previous paragraph: A finite group is cyclic if and only if it results from modular addition). From a mathematical point of view, the two structures are equivalent (for systems of order 3, not for higher order): It is easy to prove that (2.2) also holds in the first system and that (1.2) and (1.3) are also valid in the second one. The structural difference does not lie in the *total* set of relations but in those chosen to act as the *basic* ones from which the other can be deduced. Plausibly, participants in Kaminski et al. (2008a) focused on these basic relations and failed to notice the other ones. Hence, we have different dominant mathematical patterns in the two domains. Probably, the cyclic nature of the group was not noticed by abstract learners, but this did not harm them in the transfer test, because the transfer and abstract learning domain share the same dominant pattern. Similarly, concrete learners may have failed to observe relations (1.2) or (1.3), but now, the consequences were more serious: It may account for their difficulties to recognize the dominant mathematical pattern in the transfer test.

The baseline in these elements of critique is that Kaminski's comparison is not fair and the results are not surprising, given the greater deep level similarity between the transfer and abstract learning domain (Jones, 2009a, 2009b; Mourrat, 2008). Apart from the concrete–abstract distinction, other, uncontrolled, variables may have played an important role. A change in learning and/or transfer domains without altering their concrete–abstract nature may give different results. According to Jones (2009a, 2009b), a number of possible alternative instantiations of a group of order 3 (e.g., rotations of an equilateral triangle, permutations of a set of three elements) could have been chosen for a transfer task, some of which might have led to better performance for the concrete-instantiation group than for the abstract-instantiation one. Another "addition modulo 3" context seems to be an obvious candidate here.

Besides the first group of critical comments referring to the unfairness of Kaminski's comparison, a second—but related—vulnerable element in Kaminski's work inspired us to conduct a new empirical study. So far, no clear information on what participants in the two learning conditions actually learned about the to-be-learned concept (a group of order 3) has been provided. With only multiple-choice data available, it seems quite impossible to know how these participants perceived the three objects and the operations in the abstract and concrete instantiations (see also Jones, 2009a). In other words, the Kaminski et al. (2008a) study does not provide strong empirical evidence that the participants in the abstract learning condition actually learned the abstract concept of a group instead of just memorizing and mapping symbols and combination rules. The

latter is not improbable. Indeed, the number of specific rules is limited: only 3 for a group of order 3. Moreover, some of the rules defining the group structure (e.g., commutativity and associativity) have an "evident" character for participants who are accustomed to calculating with numbers (which also may have contributed to participants' relatively good scores on the abstract learning and on the transfer test). So, it is likely that these combination rules did not even attract participants' attention during the learning phase. For similar reasons, it also remains unclear what the participants in the concrete-instantiation condition actually learned: Did they learn the group concept, arithmetic modulo 3, how to apply acquired rules, or even something else?

To substantiate empirically the two previously mentioned elements of critique, we designed a replication and extension study. In order to make the comparison between the learning groups more fair, we added an "addition modulo 3" transfer domain to Kaminski et al.'s (2008a) transfer domain. This new transfer domain introduced a context with slices of pizza (about pizza's being 1/3, 2/3, or completely burned and these slices of pizza were combined in the same way as the liquids in the cups). Kaminski also used this pizza context, but as one of the learning domains. This addition enabled us to measure transfer to a new domain and, whereas Kaminski et al. discovered better transfer from an abstract learning domain, we expected the analogue would happen from a concrete learning domain to the new transfer domain. Besides the multiple-choice data, we collected data from an open-ended question in order to obtain information on what participants actually learned from the abstract- and concrete-instantiation conditions. Although we basically expected rule memorization by the abstract learners, we anticipated some learning of arithmetic modulo 3 by learners in the concrete-instantiation condition.

Before discussing the method and results of this study, we want to comment briefly on two terminological points that deserve further clarification: the nature of the observed transfer in Kaminski's (2006) study and the concrete–abstract nature of the two transfer domains.

The main focus in Kaminski's (2006) study is transfer. Her claim that abstract examples are better refers to better transfer. Jones (2009a, 2009b) argues that the more successful transfer by generic learners in Kaminski's study is near transfer, that is, the "relatively direct use of knowledge acquired during learning" (Jones, 2009a, p. 84). Moreover, the transfer is short term, because there was little or no time delay between the learning phase and the transfer test. Finally, the transfer was prompted: Participants were explicitly asked to apply the rules of the system from the learning phase, and, thus, did not have to demonstrate spontaneous transfer. These comments clarify the (limited) nature of the transfer studied by Kaminski and, hence, also the transfer study reported in this article.

Kaminski et al. (2008a) call the children's game transfer domain a *concrete* domain, based on the superficial characteristic that "perceptually rich elements" are used. However, at a deeper level it satisfies those researchers' own definition of *generic* (or abstract): "Instantiations that communicate minimal extraneous details, beyond the defining structural information, are generic instantiations, while those that communicate more extraneous information are concrete" (p. 1). Hence,

it is appropriate to call it an *abstract* transfer domain. The newly added pizza transfer domain is *concrete*, both in Kaminski's and in our interpretation.

METHOD

One hundred thirty undergraduate students in educational sciences participated in return for course credit. Like Kaminski's experiment, our study had two phases: (a) training and testing in a learning domain, and (b) testing for transfer. The training consisted of a brief introduction to the context, an explicit presentation of the rules using examples, questions with feedback, some more complex examples, and a summary of the rules. All participants were randomly assigned to one of four experimental conditions—AA, AC, CA, and CC—in which the first letter specifies the learning domain and the second letter the transfer domain, both of which could be either abstract (A) or concrete (C). The experimental materials and procedures were comparable to those described in Kaminski's (2006) dissertation.

A first important difference between the design for this study and the overall design of Kaminski's central experiment was that, in addition to the AA- and CA-conditions that were also used in Kaminski's experiment, we included AC- and CC-conditions, in which a new modulo 3 addition context was presented in the transfer phase in order to measure transfer to a new *concrete* domain. The A-learning and A-transfer conditions were operationalized by the arbitrary symbols and their combination rules in, respectively, the archaeological and children's game context as described previously. For the C-learning and C-transfer conditions, respectively, the liquid cups and pizza contexts were introduced (see supra). Just before the test administration in the learning domain, a summary of key ideas—possible combinations and properties of identity and commutativity ("order doesn't matter")—was presented in both verbal and pictorial form. In the four experimental conditions, the tests that were administered at the end of the learning phase as well as the transfer tests consisted of 24 "isomorphic" multiple-choice questions in which participants had to choose one correct answer out of two, three, or four alternatives (for an example of two isomorphic multiple-choice questions from the abstract and concrete learning tests, see Figure 2). Across the experimental
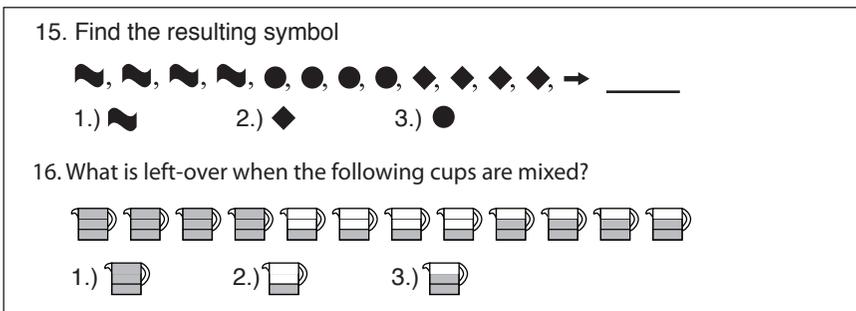

Figure 2. Examples of multiple-choice questions as used in the A- (left) and C- (right) learning tests.

conditions, training in the learning domain and the presentation of the transfer domain were similar.

As a second important difference between Kaminski's and our procedures, we included an open-ended question immediately after the learning phase and before the transfer test in order to get a better idea of what students report to have learned from the abstract or concrete examples. A sequence of four elements was presented (diamond, circle, circle, circle, and 2/3, 1/3, 1/3, 1/3 filled cups in, respectively, the A- and C-learning conditions), followed by an arrow and a question mark (see Figure 3). Participants were asked what should come on the place of the question mark and had to explain in their own words how they had arrived at their answer.

As in Kaminski's (2006) research, training and testing happened individually on a computer, and the testing occurred immediately after the training. Participants proceeded through training and testing at their own pace (but all of them finished the experiment in less than 2 hours); their responses to the multiple-choice tests and to the open-ended question were digitally registered.

Using exactly the same procedure as in Kaminski (2006), 20 participants were removed from this analysis because they failed to learn: Their learning scores were lower than 11 (of 24). Again following Kaminski's procedure, 5 outliers were removed because their learning or transfer scores were more than 2 standard deviations from the mean of their respective conditions. For the AA-, AC-, CA-, and CC-conditions, the number of participants' scores ultimately involved in the ANOVA was 23, 30, 28, and 24, respectively.

Participants' learning and transfer test scores were analyzed (in SPSS) using two single-factor ANOVA's (one on the learning scores and one on the transfer scores) with treatment condition as the single factor. Post hoc contrasts between conditions were conducted using corrections for multiple comparisons (Tukey–Kramer). As a consequence of removing participants (as previously stated), variation in the different conditions might no longer be homogeneous. However, Levene's test indicated homogeneity of variances ($p = .173$ and $.092$ for the learning and transfer scores, respectively). A post hoc power analysis on the transfer scores revealed a statistical power of 72.9% (using a confidence level of 5%).
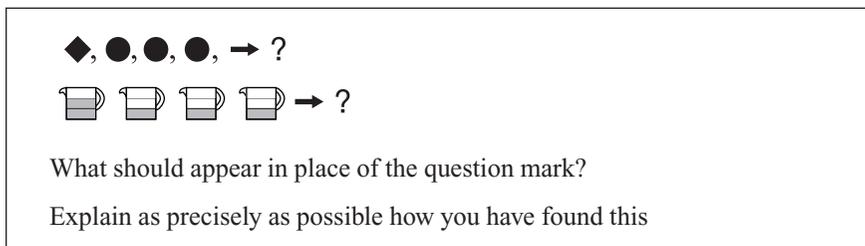

What should appear in place of the question mark?

Explain as precisely as possible how you have found this

Figure 3. Open-ended question included after the learning domain test in the A- (first row) and C- (second row) learning conditions.

To analyze participants' explanations of their answers to the open-ended question, a scoring system was developed that was applicable irrespective of the specific learning domain (see Table 1). Each participant explanation was categorized in one or more (sub)categories of the system. A score of 2 was given for an explanation at a general level (even if formulated awkwardly). For the abstract learning domain, examples of such formulations are "if you combine a flag with another symbol, you get that other symbol" (role of identity) or "order doesn't matter" (commutativity). For the concrete learning domain, it means that explanations are stated in terms of modulo 3 calculation, making abstraction of the concrete context of liquid cups that are poured together (e.g., "$2 + 2 = 4 - 3 = 1$"). A score of 1 was assigned when a rule was correctly applied in at least one step of the participant's reasoning, but not formulated at a more general level. Two independent raters applied this system to all participants' explanations and indicated that they were unsure about only 4% of the cases. After a careful comparison of these cases to the detailed descriptions in the scoring system, full agreement between raters was reached.

RESULTS

Table 2 lists the means and standard deviations of the learning and transfer scores in the four conditions. With respect to the learning scores, there was a significant difference between groups, $F(3, 101) = 5.144$, $p = .002$. The AC-group scored significantly lower than the CA-group ($p = .004$) and the CC-group ($p = .010$). The

Table 1
Scoring System Applied to Explanations Given for Answer to Open-Ended Question

Category	Description	Subcategory	Score
G (Group)	Axioms of commutative group applied/application of these axioms formally or informally formulated	G1: associativity G2: role of identity G3: role of inverses G4: commutativity	2: formulation at general level 1: unambiguous application 0: else
M (Modulo)	Properties of "modulo 3" arithmetic applied/application of these properties formally or informally formulated	M1: via whole numbers M2: via fractions	2: formulation at general level 1: unambiguous application 0: else
R (Rules)	One or more combination rules (almost) literally repeated		1: yes 0: no
N (No)	No explanation/explanation irrelevant or incomprehensible		1: irrelevant or incomprehensible 0: no

Table 2
Mean and Standard Deviation (Between Brackets) of Learning and Transfer Test Scores

Condition	Mean (standard deviation) test scores (Max = 24)	
	Learning test	Transfer test
AA ($N = 23$)	17.1 (3.9)	18.1 (3.8)
AC ($N = 30$)	15.3 (3.5)	17.4 (4.2)
CA ($N = 28$)	18.5 (2.9)	12.0 (4.3)
CC ($N = 24$)	18.3 (3.5)	20.2 (2.4)

difference between the AA-group and the three other groups was not significant. With respect to the transfer scores, there also was a significant between-groups difference, $F(3, 101) = 21.797$, $p < .001$. The scores of the CA-group were significantly lower than those in the other three groups ($p < .001$ for all comparisons) and the scores of the AC-group were significantly lower than those in the CC-group ($p = 0.044$).Our results with respect to the two conditions that were taken from Kaminski et al.'s (2008a) original study (AA and CA) confirm their findings: The AA-group outperformed the CA-group on the abstract transfer task. So, at least for this type of task, if transfer to a new *abstract* domain is targeted, abstract instantiations are indeed more advantageous than concrete instantiations. However, our finding that the CC-group outperformed the AC-group on the concrete transfer task shows that the opposite holds as well: Transfer to a new concrete domain is more enhanced by a concrete learning domain than by an abstract one. Finally, although the learning scores of the AC-group were significantly lower than those of the two concrete-learning-domain groups, the transfer scores of the AC-group did not significantly differ from those of the AA-group. This result shows that, to a certain extent, students who were taught in an abstract learning domain were able to deal with a "modulo 3" concrete instantiation of the concept of a group of order 3, even when the learning task provided little or no support for this. The different complexity of the two learning domains is also reflected in the learning scores themselves: Although all students learned the concept sufficiently well, the A-learning groups performed worse—and the AC-group even significantly worse—on the learning test than the C-learning groups.

Table 3 lists the ratings of the explanations participants gave for their answer to the open-ended question according to the scoring system presented in Table 1.

The ratings for the abstract learning domain provide little or no evidence that participants have learned the concept or axioms of a commutative group. Although participants were asked to explain *as precisely as possible* how they had found their answer, three of the four group axioms (associativity, commutativity, and role of the inverse) were not formulated at a general level in any of the explanations. The other group axiom (role of the identity) was formulated at a general level in only

Table 3
Ratings of Explanations Given for Answer on Open-Ended Question According to the Scoring System of Table 1

Learning domain	Score	G				M		R	N
		G1	G2	G3	G4	M1	M2		
A (N = 66)	2	0	6	0	0	0	0	–	–
	1	16	43	0	3	0	0	62	11
	0	50	17	66	63	66	66	4	55
C (N = 52)	2	0	0	0	0	7	0	–	–
	1	13	7	0	2	22	5	5	14
	0	39	45	52	50	23	47	47	38

6 of the 66 explanations, even though such formulations were provided and exemplified in the learning phase (e.g., "When any symbol combines with a flag, the result will always be the other symbol" or "The order does not change the result"). On the other hand, some explanations reflected a spontaneous, possibly unconscious, application of the properties of identity (43/66), associativity (16/66), or commutativity (3/66). The high mark for identity is not surprising because its role can also be seen—and was rated that way, too—as just one of the combination rules. With respect to associativity and commutativity, it remains unclear whether participants "learned" these properties from the preceding abstract learning phase or simply were familiar with them from their prior experiences with operations in number systems. None of the explanations reflected modular arithmetic, which agrees with our expectation that learners in the abstract-instantiation condition would not notice the modular structure implicitly present in the learning domain. Finally, 62 of 66 participants (almost) literally repeated such rules when explaining their answer to the open-ended question. It is questionable whether these participants also learned the quintessence of the abstract group concept (a set of elements with an operation that satisfies specific axioms). Rather, the results seem to point out that they merely learned by memorizing the formal combination rules that can be applied to arbitrary symbols.

The ratings for the concrete learning domain show that about half the explanations reflected an application of "modulo 3" arithmetic by using whole numbers (22/52) or fractions (5/52). Moreover, although it was not explicitly addressed in the learning phase, in another seven explanations the rules for modulo 3 addition were "formulated" without referring to the concrete liquid cups context (e.g., "The answer is a 2/3 filled cup because $(2 + 1 + 1 + 1)/3 \rightarrow 1$ remainder 2" or "A cup with two: $2 + 1 + 1 + 1 = 5$ and $5 - 3 = 2$"). Pure repetitions of combination rules were rare in the concrete learning domain (5/52). As in the abstract domain, some explanations reflected a spontaneous application of the group properties of identity (7/52), associativity (13/52), or commutativity (2/52). Compared to the abstract

domain, the identity property was reported less often, probably because this property is self-evident here and no longer just one of the combination rules. The relative frequency of the associativity and commutativity ratings were more or less equal across learning domains, suggesting that the application of these properties is probably caused by their natural, straightforward character or by participants' prior familiarity with them.

CONCLUSIONS AND DISCUSSION

Kaminski et al. (2008a) conducted a study consisting of a series of controlled experiments with undergraduate students, from which they concluded that the concept of a group of order 3 is learned better through an abstract representation than from concrete examples. They suggested that this conclusion is generalizable to concepts and methods in other areas of mathematics:

> Moreover, because the concept used in this research involved basic mathematical principles and test questions were both novel and complex, these findings could likely be generalized to other areas of mathematics. For example, solution strategies may be less likely to transfer from problems involving moving trains or changing water levels than from problems involving only variables and numbers. (Kaminski et al., 2008a, p. 455)

They also argue for generalizability to other age ranges:

> Because the difficulty of transferring knowledge acquired from concrete instantiations may stem from extraneous information diverting attention from the relevant mathematical structure, concrete instantiations are also likely to hinder transfer for young learners who are less able than adults to control their attentional focus. (Kaminski et al., 2008a, p. 455)

Kaminski et al. (2008a) attracted several critical reactions. A first recurring element of critique refers to the unfairness of the comparison, due to the role of uncontrolled variables: Differences in underlying structure, in the targeted mathematical concept, and in the learners' possibility to rely on prior knowledge about physical reality or number systems resulted in a difference in deep-level similarity to the transfer domain between the generic (or abstract) and the concrete learning domains. A second element of critique refers to what students actually learned from the abstract examples: the abstract concept of a group or simply the ability to memorize and map symbols and combination rules. It also remains unclear what the students in the concrete instantiation conditions learned.

To substantiate these two elements of critique empirically, we designed an investigation that was partly a replication and partly an extension of Kaminski's study. First, to re-examine the transfer potential of abstract and concrete learning domains, we extended Kaminski's design by including—in addition to the two conditions with the abstract domain as a transfer domain—two new conditions in which the transfer domain is a concrete one. Second, to better understand what students actually learned from the two learning domains, we gathered and analyzed participants' explanations for their answers to an open-ended question.

Our results confirm the basic finding by Kaminski et al. (2008a): Transfer to a new abstract domain is better enhanced by an abstract learning domain than by a concrete learning domain. However, through our extended design, we were able to show that this is only one side of the coin: Transfer in a new concrete domain is also enhanced more by concrete instantiations than by abstract instantiations. Moreover, the qualitative analysis of participants' explanations suggested that students who worked in the abstract domain learned how to apply combination rules formally to arbitrary symbols but provided little evidence that they also acquired the abstract mathematical concept of *group*. Some students who worked in the concrete domain reached a higher level of abstraction by formulating combination rules in a context-independent manner, and most students in the concrete domain clearly learned a different mathematical topic, namely, addition modulo 3. These results pose a serious challenge to Kaminski et al.'s (2008a) affirmative conclusions about "the advantage of abstract examples in learning math" (p. 454).

We agree with Jones (2009b) that it is inappropriate to extrapolate Kaminski's—and, thus, also our—findings to the broader realm of mathematics education: The mathematical topic of these experiments, namely, groups of order 3, is a very particular one, and the "examples" used to make this topic "concrete" are quite artificial and far away from what (realistic) mathematics educators generally have in mind when they think of the introduction of more general mathematical ideas by concrete instantiations. Moreover, Kaminski's (2006) and our short-term "laboratory" experiments differ on crucial points from instruction that occurs in typical educational settings, and their effects over longer periods of time remain unexplored.

Understanding the abstract mathematical concept of *group* also has an epistemological meaning: From where does this algebraic structure with its four axioms—these and no others—originate? Neither the concrete nor the abstract instantiations used by Kaminski (2006) (and by us) can shed any light on this fundamental question. Similar to instruction in calculus, in which the derivative is often introduced in the context of motion, closely related to one of the historical sources of that concept, one could consider a meaningful educational approach (at the tertiary level) of the abstract group concept starting from an instantiation used historically by mathematicians to investigate basic mathematical problems with which they were faced (e.g., permutation groups in a quest for general solutions of polynomial equations of higher degree or symmetry groups as a classification tool for lattices and patterns in geometry).[2]

We conclude with a more specific question about the generalizability of Kaminski et al.'s (2008a) results. Even if Kaminski et al. have shown that one instructional approach to the learning of a given concept leads to better transfer than another, it is still unclear what the consequences of that approach are for further learning of related and/or more advanced concepts. Applied to the concept dealt

---

[2]For examples of historical approaches to the teaching and learning of mathematics, see Fauvel and van Maanen (2000).

with here, one can think of the mathematical objects coming next in level of complexity: cyclic groups of order 4 and higher (Deprez et al., 2010). Can we still construct an abstract learning domain for these objects in Kaminski et al.'s style? In fact, a much larger number of specific rules would be necessary: Instead of the 3 rules in the case of order 3, for cyclic groups of order 4, 5, 6, 7, . . . , one would need 6, 10, 15, 21, . . . rules.[3] Knowing from our qualitative analysis that abstract learners primarily relied on the set of specific rules, we conjecture that both learning and transfer scores of abstract learners would rapidly decline with increasing order of the group. In order to cope with higher-order systems, additional structure, beyond formal combination rules, has to be introduced. So, it would even be problematic to construct a successful abstract learning phase in Kaminski et al.'s style for the concept of a cyclic group of higher order.

Ideas for future research arise from the results of Kaminski's and our study. Kaminski et al. (2008a) investigated the effect of a concrete instantiation followed by an abstract instantiation on an abstract transfer domain (cf. supra), but the effect of a combination of concrete and abstract instantiations on a new *concrete* transfer domain remains to be explored. It would also be interesting to further investigate empirically Jones' (2009a, 2009b) claim that the transfer effects obtained by Kaminski might remain restricted to near, immediate, and prompted transfer and not generalize to far, long-term, or unprompted transfer. Also, replications of parts of Kaminski's work in ecologically more valid educational settings, wherein learners' personal mathematics-related beliefs, attitudes, and emotions (McLeod, 1992) and the sociomathematical classroom norms and practices (Yackel & Cobb, 1996) play a more significant role, would be worth considering.

---

[3]In the following, we calculate the number of specific rules (the Cayley table below—for a group of order 5—illustrates our argument). No rules are needed for operations in which the identity (n) figures (see the first row and the first column in the Cayley table). For the first nonidentity element, a rule is required for each of its combinations with the other nonidentity elements, including itself (see the Xs in the remaining cells in the second row in the Cayley table). This already gives $k-1$ rules (for a group of order k). Due to the abelian (i.e., commutative) character of the group, these rules allow one to fill in the remaining cells in the second column of the Cayley table, as well. Similarly, for the second nonidentity element, $k-2$ rules are required, corresponding to the Xs in the remaining cells in the third row of the Cayley table, and allowing one to fill in the remaining cells in the third column as well. A continuation of this argument shows that the total number of rules required is given by $(k-1)+(k-2)+(k-3)+\ldots+2+1$, that is, the $(k-1)$th triangular number $k(k-1)/2$.

	n	a_1	a_2	a_3	a_4
n	—	—	—	—	—
a_1	—	X	X	X	X
a_2	—	—	X	X	X
a_3	—	—	—	X	X
a_4	—	—	—	—	X

REFERENCES

Abstracte wiskunde leert beter dan praktische voorbeelden [Abstract mathematics is learned better than practical examples]. (2008, April 30). *De Standaard*.

Bills, L., Mason, J., Watson, A., Zaslavsky, O., Goldenberg, P., Rowland, T., et al. (2006). Exemplification: The use of examples in teaching and learning mathematics. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings of the 30th conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 125–154). Prague, Czech Republic: International Group for the Psychology of Mathematics Education.

Chang, K. (2008, April 25). Study suggests math teachers scrap balls and slices. *The New York Times*. Retrieved August 29, 2009, from http://www.nytimes.com/2008/04/25/science/25math.html

Cutrona, L. J. (2008). "Concrete" examples a fraction too abstract [Letter]. *Science, 322*, 1632. Retrieved August 29, 2009, from http://www.sciencemag.org/cgi/content/summary/320/5875/454

de Lange, J. (1987). *Mathematics, insight and meaning. Teaching, learning and testing of mathematics for the life and social sciences*. Utrecht, The Netherlands: Vakgroep Onderzoek Wiskundeonderwijs en Onderwijs Computercentrum, Rijksuniversiteit Utrecht.

Deprez, J. (2008). Concreet of abstract? [Concrete or abstract?]. *Uitwiskeling, 24*(4), 63–69.

Deprez, J., De Bock, D., Roelens, M., & Verschaffel, L. (2010). Do abstract examples *really* have advantages in learning math? In M. M. F. Pinto & T. F. Kawasaki (Eds.), *Proceedings of the 34th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 305–312). Brazil, Belo Horizonte: International Group for the Psychology of Mathematics Education.

Fauvel, J., & van Maanen, J. (Eds.). (2000). *History in mathematics education. New ICMI study series: Vol. 6. History in mathematics education: An ICMI study*. Dordrecht, The Netherlands: Kluwer.

Gick, M., & Holyoak, K. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.

Gravemeijer, K. (1994). *Developing realistic mathematics education*. Utrecht, The Netherlands: Freudenthal Institute.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 15–46). New York: Macmillan.

Greeno, J. G., Smith, D. R., & Moore, J. L. (1993). Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 99–167). Norwood, NJ: Ablex.

Jones, M. G. (2009a). Transfer, abstraction, and context. *Journal for Research in Mathematics Education, 40*, 80–89.

Jones, M. G. (2009b). Examining surface features in context. *Journal for Research in Mathematics Education, 40*, 94–96.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2005). Relevant concreteness and its effects on learning and transfer. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the Cognitive Science Society* (pp. 1090–1095). Mahwah, NJ: Erlbaum.

Kaminski, J. A. (2006). The effects of concreteness on learning, transfer, and representation of mathematical concepts. (Doctoral dissertation, Ohio State University, 2006). *Dissertation Abstracts International, 67*, 2500A.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2006a). Do children need concrete instantiations to learn an abstract concept? In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 411–416). Mahwah, NJ: Erlbaum.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2006b). Effects of concreteness on representation: An explanation for differential transfer. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 1581–1586). Mahwah, NJ: Erlbaum.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008a). The advantage of abstract examples in learning math. *Science, 320*, 454–455.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008b). Supporting online material for *The advantage of abstract examples in learning math*. Retrieved May 2, 2008, from http://www.sciencemag.org/cgi/content/full/320/5875/454/

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. (2009). Transfer of mathematical knowledge: The

portability of generic instantiations. *Child Development Perspectives, 3*, 151–155.

Les exemples sont mauvais pour l'apprentissage des mathématiques [Examples are bad for learning math]. (2008, April 25). *Le Monde*. Retrieved August 29, 2009, from http://www.lemonde.fr/planete/article/2008/04/25/les-exemples-seraient-des-mauvais-instruments-d-enseignement-des-mathematiques_1038312_3244.html

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17–20.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47–62). New York: Macmillan.

McCallum, W. (2008). *Commentary on Kaminski et al.,* The advantage of abstract examples in learning math, *Science*, April 2008. Retrieved August 29, 2009, from http://math.arizona.edu/~wmc/Research/Kaminsky_et_al_commentary.pdf

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics learning and teaching* (pp. 575–596). New York: Macmillan.

Mourrat J.-C. (2008). Testing the limits of "concrete" and "generic" [Letter]. *Science, 322*, 1632. Retrieved December 17, 2008, from http://www.sciencemag.org/cgi/content/summary/320/5875/454

Podolefsky, N. S., & Finkelstein, N. D. (2009, February). Counterexamples and concerns regarding the use of abstract examples in learning math [Letter]. Retrieved August 29, 2009, from http://www.sciencemag.org/cgi/eletters/320/5875/454#11970

Thorndike, E. L., Cobb, M. V., Orleans, J. S., Symonds, P. M., Wald, E., & Woodyard, E. (1924). *The psychology of algebra*. New York: Macmillan.

Watson, A., & Mason, J. (2002). Student-generated examples in the learning of mathematics. *Canadian Journal of Science, Mathematics and Technology Education, 2*(2), 237–249.

Wertheimer, M. (1959). *Productive thinking*. (Enlarged ed., M. Wertheimer, Ed.). New York: Harper & Brothers.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458–477.

Authors

Dirk De Bock, Faculty of Economics and Management, Hogeschool-Universiteit Brussel, Stormstraat 2, B-1000 Brussels, Belgium, dirk.debock@hubrussel.be

Johan Deprez, Faculty of Economics and Management, Hogeschool-Universiteit Brussel, Stormstraat 2, B-1000 Brussels, Belgium, johan.deprez@hubrussel.be

Wim Van Dooren, Centre for Instructional Psychology and Technology, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium, wim.vandooren@ped.kuleuven.be

Michel Roelens, Department of Teacher Education, Katholieke Hogeschool Limburg, B-3590 Diepenbeek, Belgium, michel.roelens@ler.khlim.be.

Lieven Verschaffel, Centre for Instructional Psychology and Technology, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium, lieven.verschaffel@ped.kuleuven.be