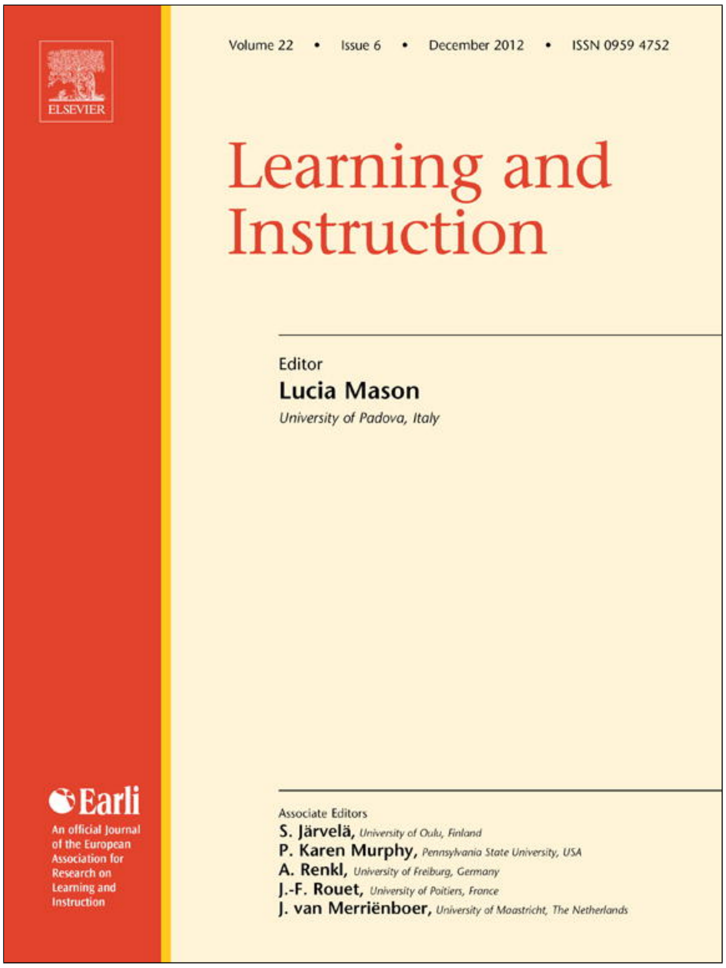
Volume 22 • Issue 6 • December 2012 • ISSN 0959 4752
ELSEVIER
Learning and Instruction
Editor
Lucia Mason
University of Padova, Italy
Earli
An official Journal of the European Association for Research on Learning and Instruction
Associate Editors
S. Järvelä, University of Oulu, Finland
P. Karen Murphy, Pennsylvania State University, USA
A. Renkl, University of Freiburg, Germany
J.-F. Rouet, University of Poitiers, France
J. van Merriënboer, University of Maastricht, The Netherlands

Learning and Instruction 22 (2012) 440–448

Contents lists available at SciVerse ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

"Concreteness fading" promotes transfer of mathematical knowledge

Nicole M. McNeil*, Emily R. Fyfe [1]

Department of Psychology, University of Notre Dame, 118 Haggar Hall, Notre Dame, IN 46556, United States

A R T I C L E I N F O

Article history:
Received 29 September 2011
Received in revised form
26 April 2012
Accepted 2 May 2012

Keywords:
Mathematics learning
Transfer
Symbols
Concreteness fading

A B S T R A C T

Recent studies have suggested that educators should avoid concrete instantiations when the goal is to promote transfer. However, concrete instantiations may benefit transfer in the long run, particularly if they are "faded" into more abstract instantiations. Undergraduates were randomly assigned to learn a mathematical concept in one of three conditions: *generic*, in which the concept was instantiated using abstract symbols, *concrete* in which it was instantiated using meaningful images, or *fading*, in which it was instantiated using meaningful images that were "faded" into abstract symbols. After learning, undergraduates completed a transfer test immediately, one week later, and three weeks later. Undergraduates in the fading condition exhibited the best transfer performance. Additionally, undergraduates in the generic condition exhibited somewhat better transfer than those in the concrete condition, but this advantage was not robust. Results suggest that concrete instantiations should be included in the educator's toolbox.

1. Introduction

Many educators use concrete materials during instruction. For example, a preschool teacher might use toy animals in a lesson on counting and sorting, and an algebra teacher might use a balance scale in a lesson on equations. We use "concrete" here to refer to materials that are grounded in previous perceptual and/or motor experiences and have identifiable correspondences between their form and referents. Contrast this with "abstract" materials, which eliminate detailed perceptual properties and are more arbitrarily linked to referents. Learning materials vary in their level of concreteness, and the notion that high levels of concreteness benefit learning has a long history in psychology and education, dating back to Montessori (1917), Bruner (1966) and Piaget (1970). Such materials are theorized to benefit learning by activating real-world knowledge (Kotovsky, Hayes, & Simon, 1985; Schliemann & Carraher, 2002), inducing physical or imagined action (Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004; Martin & Schwartz, 2005), and enabling students to construct their own knowledge of abstract concepts (Brown, McNeil, & Glenberg, 2009; Martin, 2009; Smith, 1996). There are also practical reasons to use concrete materials. They are widely available in teaching supply stores, and they seem to increase students' motivation and interest in learning (Burns, 1996).

Despite the theoretical and practical reasons for using concrete materials, there are also reasons to caution against the use of concrete materials. For example, concrete materials often contain extraneous perceptual details that can distract learners from the to-be-learned information (Kaminski, Sloutsky, & Heckler, 2008). Such materials may be so interesting in their own right that they draw learners' attention to themselves, rather than to their referents (DeLoache, 2000; Uttal, Scudder, & DeLoache, 1997). In these ways, concrete materials may "seduce" learners' attention away from the concepts educators want to convey, similar to what happens when a passage of text contains interesting or entertaining details that are irrelevant to the main ideas in the passage (a.k.a., "the seductive details effect," Garner, Gillingham, & White, 1989; Harp & Mayer, 1998; Sanchez & Wiley, 2006). In support of this view, several studies have shown little to no benefits of using concrete materials (Ball, 1992; Baranes, Perry, & Stigler, 1989; Resnick & Omanson, 1987; Thompson, 1992; Uttal et al., 1997). Some studies have even shown negative effects (Goldstone & Sakamoto, 2003; Kaminski, Sloutsky, & Heckler, 2006, 2008; McNeil, Uttal, Jarvin, & Sternberg, 2009; Sloutsky, Kaminski, & Heckler, 2005; Son, Smith, & Goldstone, 2008), particularly for transfer to novel problems. For example, Goldstone and Sakamoto (2003) found that low-performing undergraduates who learned a scientific principle (competitive specialization) from a computer simulation involving concrete images (realistic looking ants and fruit) exhibited worse transfer to a superficially different domain than did those who

* Corresponding author. Tel.: +1 574 631 5678.
E-mail address: nmcneil@nd.edu (N.M. McNeil).
[1] Department of Psychology and Human Development, Vanderbilt University, United States.

0959-4752/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.learninstruc.2012.05.001

learned the principle from a simulation involving generalized images (black dots and green patches).

Similarly, a recent study by Kaminski et al. (2008) provided systematic evidence that concreteness can be detrimental to transfer in the domain of mathematics. Participants learned the rules of a commutative group of order 3 either through a generic instantiation using abstract symbols (e.g., ●, ◆, ⚑), or through concrete instantiations using meaningful images such as measuring cups (e.g., , ,). All participants learned the material equally well, but a significant advantage emerged for the generic instantiation in terms of transfer to novel problems. Participants who learned through the generic instantiation even outperformed those who learned through *multiple* concrete instantiations (i.e., measuring cups, pizza slices, and tennis balls in a container). These results were controversial because researchers and educators have long assumed that multiple concrete instantiations of a concept are beneficial for learning and transfer (e.g., Bruner, 1966; Gick & Holyoak, 1983). A follow-up experiment produced even more striking results—participants who learned *only* through a generic instantiation outperformed those who learned through a concrete instantiation that was explicitly linked to a generic instantiation. Thus, learning from a single generic instantiation seemed to promote transfer better than learning from multiple instantiations when at least one was concrete, arguably because the generic instantiation is less distracting, less context specific, and has greater representational status than the concrete instantiation(s) (Kaminski et al., 2008, see also Goldstone & Son, 2005). Some reporters in the popular press interpreted these results as evidence that teachers should stop using concrete examples when teaching mathematics concepts to students (Chang, 2008).

Although the aforementioned studies have provided valuable information about the potential limitations of concrete materials, recent critiques suggest that the movement against concrete examples may be premature (e.g., De Bock, Deprez, Van Dooren, Roelens, & Verschaffel, 2011; Jones, 2009). In the present study, we focus on two specific points of contention that make it difficult to accept the conclusion that teachers should scrap concrete materials. First, previous studies have examined participants learning from concrete and generic instantiations either presented alone or one after another *in isolation*. Although Kaminski et al. (2008) explicitly told participants in their "Concrete-then-Generic" condition that the concrete and generic instantiations followed similar rules, many proponents of concrete materials have specifically recommended beginning with concrete instantiations and then explicitly "decontextualizing" or "fading" away to the more abstract (Bruner, 1966; Goldstone & Son, 2005; Gravemeijer, 2002; Lehrer & Schauble, 2002; Lesh, 1979). For example, a teacher might teach "three" by presenting the following sequence of representations: three apples → a picture of three apples → a picture of three red dots → a picture of three black tally marks → the Arabic numeral 3. This explicit, gradual fading process may be necessary to provide benefits over and above abstract materials.

Surprisingly, no study to date has experimentally tested the effects of explicit, gradual fading from concrete-to-abstract representations. Some studies have examined the effects of fading more general forms of instructional support (e.g., Wecker & Fischer, 2011), but none have tested the effects of explicitly and gradually fading the symbolic representations themselves. The closest exception is Goldstone and Son (2005), who found that undergraduates exhibited better transfer performance when learning a science concept after being taught with materials that switched from concrete to abstract (in a single step) than after being taught with materials that remained concrete, remained idealized, or switched from abstract to concrete. However, their study only involved two steps—concrete then abstract—as opposed to the three steps recommended by Bruner (1966). They also did not fully explore the scope of abstractness because their abstract materials were still fairly concrete representations of the target concept (competitive specialization). The abstract and concrete materials were both animations of ants foraging for food. The abstract version was simply stripped of the perceptual details that made it look specifically like ants and pieces of food. Further, extra measures were taken to ensure participants interpreted the small black dots as ants and the larger green patches as food, including labeling the images as ants and food. Thus, the abstract materials were not arbitrarily linked to their referents, as most mathematical symbols are.

Another fading study by Scheiter, Gerjets, and Schuh (2010) found that undergraduates exhibited better transfer performance after learning via worked-out examples that were accompanied by animations that gradually faded from concrete to abstract than after learning via worked-out examples that were not accompanied by any animations. However, they did not compare "faded" animations to non-faded animations, so it is unclear if the positive effects were due to the fading process or to the presence of animations. In the present study, we sought to fill this gap in evidence by testing participants who learned through a concrete instantiation that was explicitly and gradually faded into a more abstract instantiation. We predicted that learning through this "concreteness fading" method would be better than learning through either concrete or abstract instantiations alone (Hypothesis 1).

Another potential limitation of previous studies is that participants have been tested immediately after learning, with little to no time lapse between learning and transfer. There are at least two reasons to examine performance beyond a short time lapse. First, a short time lapse between learning and transfer does not match standard classroom testing conditions, where major assessments and high stakes tests are given several weeks after initial learning (Jones, 2009). In order to justify the time and effort it takes to teach students important concepts, transfer needs to withstand weeks, months, and even years between learning and testing (Barnett & Ceci, 2002). Second, a short time lapse between learning and transfer does not give consolidation a chance to occur (McGaugh, 2000). Consolidation refers to a set of processes that stabilize memory for a fact or event after it has been first encoded, and it takes two distinct forms: general reduction in fragility of a memory trace and off-line enhancement of the memory (Robertson, Pascual-Leone, & Miall, 2004). The second of these is thought to depend on neurophysiological changes that occur during sleep (Robertson et al., 2004). Prior to consolidation, knowledge that is learned through concrete materials may be at an unfair disadvantage relative to knowledge that is learned through abstract materials. This is because consolidation processes—especially those involving sleep—alter memory traces in ways that lead to more abstract, flexible knowledge that supports generalization and creative problem solving (Gomez, Bootzin, & Nadel, 2006; Stickgold & Walker, 2004; Wagner, Gals, Halder, Verleger, & Born, 2004). Thus, although knowledge gained from concrete instantiations may start out tied to the specific learning context, it may go through a process of abstraction in which it becomes more generalized and transferrable, particularly if combined with additional opportunities to engage the material (e.g., homework, quizzes, tests). In the present study, we addressed this issue by examining performance at two time points beyond the immediate transfer test: one week later and three weeks later. We predicted that learning through generic instantiations would initially be superior to learning through concrete instantiations (replication of Kaminski et al., 2008; Hypothesis 2), but that the advantage would dissipate over time and possibly even reverse after processes such as consolidation have occurred (Hypothesis 3).

2. Method

2.1. Participants

Participants were 80 undergraduates from a highly selective, mid-sized private university in the midwestern United States. The middle 50% of students admitted to this university score 680–760 on the SAT Math (compared to 580–690 for students admitted to the university at which the Kaminski et al. (2008) study was conducted). Six participants were excluded from the analysis because their scores in the learning phase were not above chance, and one additional participant was excluded due to experimenter error. Thus, the final sample contained 73 undergraduates (M age = 18 years, 11 months; 43 women, 30 men). Participants identified their own race/ethnicity as follows: 3% "African American" or "black", 19% "Asian", 8% "Hispanic" or "Latina/o", and 70% "white". Participants received either $5 or extra credit in a psychology course.

2.2. Design and materials

The experiment consisted of a learning phase followed by three distinct transfer phases. In the learning phase, participants learned one or more instantiations of a simple mathematical concept. Participants were randomly assigned to one of three conditions, which specified the type of instantiation(s) learned: generic, concrete, or faded from concrete to generic. Although the instantiations differed across conditions, the total number of questions given during the learning phase was equated across conditions (see learning phase below). Additionally, participants across conditions were well matched in terms of gender, $\chi^2(2, N = 72) = 1.22$, $p = .54$; year in school, $F(2, 70) = 0.62$, $p = .54$; and enrollment in a STEM (science, technology, engineering, or mathematics) major, $\chi^2(2, N = 72) = 1.55$, $p = .46$. After the learning phase, all participants were presented with a transfer test that was a novel instantiation of the learned concept at three distinct time points: immediately following the learning phase, one week later, and three weeks later. The learning phase and immediate transfer phase were the same as those in Kaminski et al.'s (2008) original study. The one- and three-week transfer phases were included to assess performance over time.

2.2.1. Learning phase

The to-be-learned concept was that of a commutative mathematical group of order 3, which is defined by a set of rules (see Table 1). Specifically, this concept is a set of three elements and an operation with the associative and commutative properties, an identity element, and inverses for each element. The goal of the learning phase was to learn the rules of the system and apply those rules to a series of problems. Participants learned the rules in one of three conditions: (a) generic, (b) concrete, or (c) fading, which differed only in the type of instantiation(s) presented during the learning phase. Fig. 1 displays the instantiation of the rules in each condition. For each instantiation, the rules were presented one at a time through specific examples, and questions with feedback were also provided. For instance, in the generic instantiation the commutative rule was presented as follows: "The order of the two symbols on the left does not change the result. For example, , , , is the same thing as , , , ." Similarly, in the first concrete instantiation the commutative rule was presented as follows: "The order by which two cups of solution are combined does not change the left-over result. For example, combining and has a leftover quantity of . And combining and has a leftover quantity of ." Participants who learned more than one instantiation (concrete and fading conditions) were explicitly told that the instantiations followed similar rules. In line with Kaminski et al. (2008), they were told, "Now you will learn about a new system. This system works the same ways as the last system you learned. The rules of the last system are like the rules of this new one." After learning the rules, participants were given an opportunity to practice applying the rules on novel problems, which were presented via a set of multiple-choice questions. Overall, participants answered 24 of these multiple-choice questions during the learning phase (see Fig. 2 for examples of these questions). Learning was equated across conditions, so that the same rules, questions with feedback, and practice questions were spread across the learning instantiations. Thus, participants who learned one instantiation (generic condition) were given practice with 24 questions. Participants who learned three instantiations (concrete and fading conditions) were given 8 questions over each instantiation. Thus, all participants received the same total number of questions (i.e., 24) during the learning phase.

In the *generic* condition, the instantiation was described as a symbolic language in which three types of symbols are combined to yield a resulting symbol (see Fig. 1). The combinations were expressed as a set of six "rules" written in the form *symbol 1*, *symbol 2* → *resulting symbol*. Three of these rules were accompanied by verbal descriptions (e.g., Rule 1: The order of the two symbols on the left does not change the result. Rule 2: When any symbol combines with , the result will always be the other symbol. Rule 6: The result does not depend on which two symbols combine first.), while the other three were not (i.e., Rule 3: → , Rule 4: → , Rule 5: →). Participants were presented with the rules and examples and then completed 24 practice questions with generic symbols. To approximate the effect of presenting the rules in three different instantiations (as in the concrete and fading conditions), additional summaries of the rules were given after the 8th and 16th questions. This condition was analogous to the "Generic 1" condition in Kaminski et al. (2008).

In the *concrete* condition, there were three instantiations, each with contextualized and meaningful elements (see Fig. 1). The first instantiation included three images of measuring cups containing varying levels of liquid; participants needed to determine the remaining amount after different cups were combined. For example, and fill one container and have remaining. The second and third instantiations contained images of pizza slices and tennis balls in a container respectively, and involved similar story lines. The task for the second instantiation was to determine the amount of burned pizza served from a restaurant where the cook systematically burned a portion of every order. For example, when an order for and was placed, then would be burned. For the third instantiation, a tennis ball machine was producing batches of 0, 1, or 2 balls so batches had to be combined. The task was to determine the extra balls resulting from combining batches. For example, if and were combined, then would be extra. After learning the rules of each instantiation, participants

Table 1
Rules of a commutative group.

Rules	Definition
Associativity	For any elements x, y, z: $[(x + y) + z = [x + (y + z)]$
Commutativity	For any elements x, y: $x + y = y + x$
Identity	There is an element, $\mathbf{I}$, such that for any element x: $x + \mathbf{I} = x$
Inverses	For any element, x, there exists another element y, such that $x + y = \mathbf{I}$

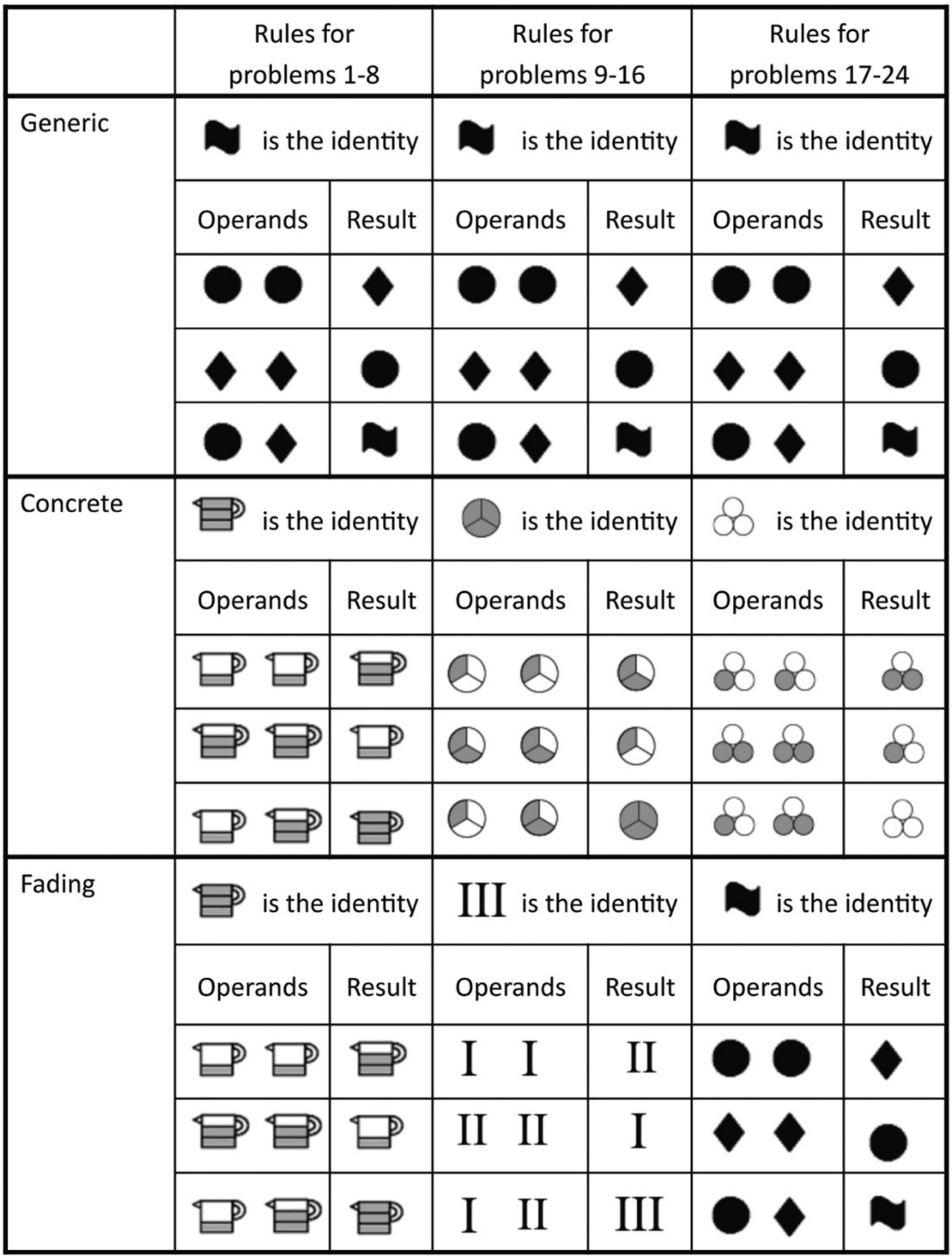

Fig. 1. Instantiations of a mathematical group of order 3 used in the generic, concrete, and fading conditions.

completed an 8 practice questions with the concrete images. Thus, participants completed a total of 24 multiple-choice questions (8 with measuring cups, 8 with pizza slices, and 8 with tennis balls). This condition was analogous to the "Concrete 3" condition in Kaminski et al. (2008).

In the *fading* condition, participants started with a concrete instantiation, which was "faded" into a generic instantiation (see-Fig. 1). This condition was similar to the "Concrete-then-Generic" condition used in Kaminski et al. (2008, experiment 4) with one important distinction. In Kaminski et al.'s original Concrete-then-Generic condition, participants first learned the concrete measuring cup instantiation followed by the generic instantiation (each in isolation). These participants were told that the generic instantiation followed the same rules as the concrete instantiation, but the images and symbols themselves were never linked. In our *fading* condition, the concrete measuring cups and the generic symbols were explicitly linked through fading with an intermediate instantiation in between to bridge the two instantiations. Participants first learned the rules in the concrete measuring cup instantiation. We selected the measuring cup instantiation because that is the instantiation used in Kaminski et al.'s original "Concrete-then-Generic" condition. After learning the concrete measuring cup

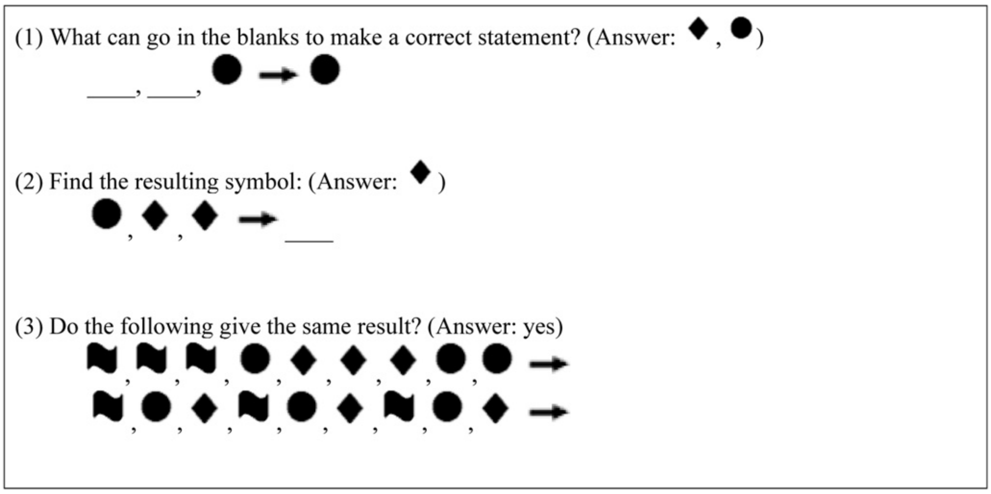

Fig. 2. Example of multiple-choice questions given in the learning phase in the generic condition.

instantiation, participants were told that the measuring cups would now be represented by simpler symbols: I, II, III. Roman numerals were used because they retain the identifiable correspondence between their forms and referents (i.e., one line, two lines, or three lines correspond to the level of liquid in a measuring cup), while being stripped of all extraneous perceptual detail (e.g., shape of the measuring cup, shading, etc.). After learning via Roman numerals, participants were told that any symbols could be used, and they were presented with the generic instantiation. After learning the rules of each instantiation, participants completed an 8 practice questions with the learned images. Thus, participants completed a total of 24 multiple-choice questions (8 with measuring cups, 8 with Roman numerals, 8 with generic symbols).

2.2.2. Transfer phases

After completing the learning phase, all participants completed the same transfer tests. There were three distinct transfer phases: immediate, one-week, and three-week. The *immediate* transfer test was presented immediately following the learning phase. The transfer test included a single instantiation, which was described as a children's game involving three objects (see Fig. 3 for elements used in the transfer phases). In this game, children pointed to objects, and a child who was labeled the "winner" pointed to a final object. The final object was specified by the rules of the game, which were the rules of the mathematical group. Participants were not explicitly taught the rules, but were told that the rules of the game were like the rules they had just learned (in the learning phase). Specifically, participants were told, "The rules of the last system you learned are like the rules of this game. So use what you know about the last system to help you figure out the rules of this game" (Kaminski et al., 2008). They were also shown nine examples from which the rules could be induced (e.g., if children point to these objects: &then the winner points to this object: ). After studying the examples, participants completed a 24-question multiple-choice transfer test. The questions on this test were isomorphic to those given during the learning phase. Questions were presented individually on the computer screen along with four key examples at the bottom of the screen (for reference). Participants did not receive any feedback about the correctness of their answers.

The *one-week* transfer phase occurred approximately one week after the learning phase. It was identical to the immediate transfer phase in all respects except a new transfer domain was presented (see Fig. 3). The novel transfer domain was described as a one-player video game, in which objects appeared sequentially on the screen, and the player's goal was to press the appropriate button corresponding to the correct resulting object.

The *three-week* transfer phase occurred approximately three weeks after the learning phase. It was identical to the one-week transfer phase, except the examples were no longer present at the bottom of the screen for reference. We removed the examples to help prevent ceiling effects on the transfer test.

2.3. Procedure

At each time point, participants were tested individually in a quiet room, with the experimenter seated just outside to field any technical questions. All stimuli were presented on a computer screen in a self-paced manner. Participants recorded their responses to all learning and transfer questions on a sheet of paper provided by the experimenter. In an attempt to prevent participants from talking to one another about the experiment, the experimenter explicitly asked them to refrain from discussing the tasks with other students.

3. Results

Attrition was low, with one participant dropping out at one week, and an additional six participants dropping out at three weeks. Thus, total attrition was seven participants (2 abstract, 4 concrete, 1 fading). We included all possible data relevant for the individual analyses below. Conclusions were unchanged when we limited all analyses to participants with complete data.

Performance during the learning phase was high ($M = 21.96$ out of 24, $SD = 2.60$). Consistent with Kaminski et al. (2008), learning was similar across conditions (generic $M = 21.32$, $SD = 3.70$; concrete $M = 22.11$, $SD = 1.93$; fading $M = 22.50$, $SD = 1.50$), $F(2, 70) = 1.29$, $p = .28$. Fig. 4 presents the average number correct on the transfer test as a function of condition and transfer phase. We first conducted the omnibus 3 (condition: generic, concrete, fading) $\times$ 3 (transfer phase: immediate, one-week, three-week) mixed-factor ANOVA with repeated measures on transfer phase and number correct (out of 24) as the dependent measure. Sphericity could not be assumed, $\chi^2(\text{df} = 2) = 33.76$, so we multiplied the numerator and denominator degrees of freedom by the relevant Huyhn–Feldt estimate of epsilon for tests involving the repeated factor, and we did not assume homogeneity of variance for follow-

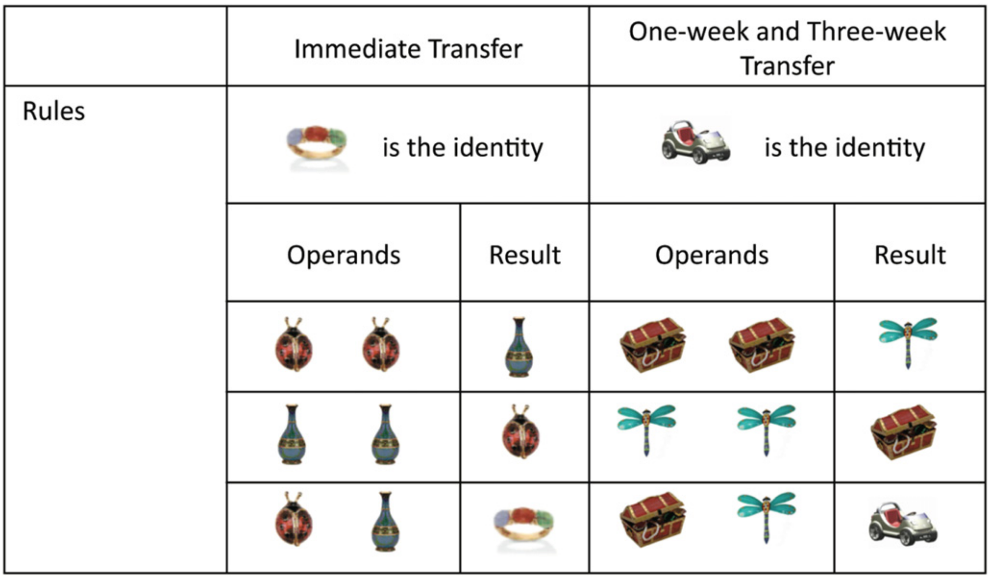

Fig. 3. Instantiations of a mathematical group of order 3 used in the immediate, one-week, and three-week transfer phases.

up comparisons. Results revealed main effects of condition, $F(2, 62) = 5.53$, $p = .006$, $\eta_p^2 = .15$ and transfer phase, $F(1.47, 91.29) = 6.56$, $p = .005$, $\eta_p^2 = .10$, but these main effects were qualified by the interaction of condition and transfer phase, $F(2.95, 91.29) = 2.95$, $p = .04$, $\eta_p^2 = .09$.

To test our prediction that the fading condition would be best (Hypothesis 1), we performed a planned contrast comparing the performance of participants in the fading condition to the performance of participants in the other two conditions averaged over the three transfer phases. As predicted, participants in the fading condition ($M = 22.85$, $SD = 1.40$) performed significantly better than participants in the other two conditions ($M = 20.03$, $SD = 4.62$), $t(51.07) = 3.80$, $p < .001$, $d = 0.83$. As shown in Fig. 4, the superiority of the fading condition was evident at all three transfer phases, even when the alpha level was adjusted using the conservative Bonferroni procedure ($\alpha_i = .02$): immediate transfer, $t(63.29) = 3.85$, $p < .001$, $d = 0.81$; one week transfer, $t(63.26) = 2.70$, $p = .009$, $d = 0.60$; three week transfer, $t(49.47) = 3.83$, $p < .01$, $d = 0.83$. However, the effect at the immediate transfer phase was driven primarily by poor performance in the concrete condition ($M = 16.65$, $SD = 5.77$).

To test our prediction that the generic condition would initially be superior to the concrete condition (Hypothesis 2), we performed a planned contrast comparing the performance of participants in the concrete condition at the immediate transfer phase to the performance of participants in the generic condition at the immediate transfer phase. As predicted, participants in the concrete condition performed significantly worse than participants in the generic condition ($M = 21.04$, $SD = 3.71$), $t(42.89) = -3.24$, $p = .002$, $d = 0.91$. Thus, although participants in our concrete and generic conditions performed about 15% better than did participants in the Kaminski et al. (2008) study, the difference between the two conditions was similar.

Finally, to test our prediction that the generic advantage over the concrete would dissipate over time and possibly even reverse (Hypothesis 3), we tested whether the difference in performance between the generic and concrete conditions was the same at all transfer phases (partial interaction). As predicted, the difference in performance between the generic and concrete conditions differed as a function of transfer phase, $F(1.38, 59.25) = 4.40$, $p = .03$. In contrast to the generic advantage seen at the immediate transfer (shown above), there was not a statistically significant difference in performance at the one week transfer phase, concrete $M = 19.12$, $SD = 5.40$, generic $M = 21.21$, $SD = 3.51$, $t(43.30) = -1.64$, $p = .11$, nor at the three-week transfer phase, concrete $M = 19.82$, $SD = 5.24$, generic $M = 21.04$, $SD = 5.07$, $t(42.74) = -0.80$, $p = .43$.

Given that scores were near ceiling and not normally distributed, we also performed a nonparametric analysis to ensure that the observed effects did not depend on the method of analysis. We first classified participants according to whether they achieved the maximum score (24) on the transfer test. Across conditions, 27% of participants achieved the maximum score at time 1, 33% at time 2, and 48% at time 3. We then used binomial logistic regression to predict the odds of achieving the maximum score at each time point (see Agresti, 1996). To parallel the results from the ANOVAs, two Helmert contrast codes were used to represent the three levels of condition (1) fading condition versus the other two conditions, and (2) concrete condition versus generic condition. Results were consistent with the ANOVAs. Participants in the fading condition were more likely than participants in the other two conditions to achieve the maximum score at time 1 (10 of 22 [45%] versus 10 of 51 [20%], $p = .03$) and at time 3 (14 of 20 [70%] versus 17 of 45 [38%], $p = .02$) (Hypothesis 1), but the difference did not reach significance at time 2 (10 of 22 [45%] versus 14 of 50 [28%], $p = .15$). Importantly, the effect sizes at times 1 and 3 were large. The model for time 1 estimates that the odds of achieving the maximum score are more

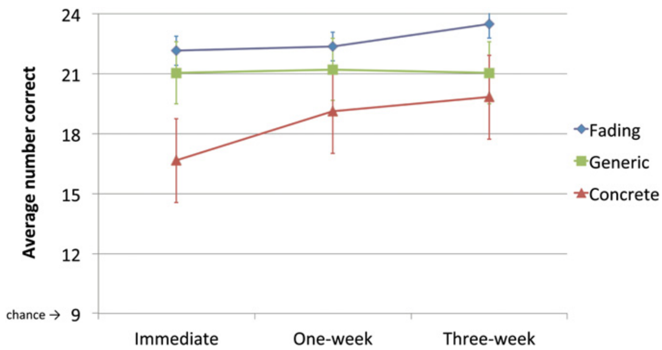

Fig. 4. Accuracy on the transfer test as a function of condition and transfer phase.

than three times higher after participating in the fading condition than after participating in one of the other conditions, and the model for time 3 estimates that the odds are six and a half times higher. The differences between the concrete and generic conditions were not significant in the nonparametric analysis at any time point (inconsistent with Hypothesis 2), suggesting that the superior performance of the generic condition at time 1, which was revealed in the ANOVA, may not be robust (Hypothesis 3).

4. Discussion

Recent evidence has suggested that concrete instantiations may be inferior to more generic, abstract instantiations, particularly when the goal is to promote transfer (Goldstone & Sakamoto, 2003; Kaminski et al., 2008, 2009; Sloutsky et al., 2005). This evidence has led some to conclude that educators should avoid concrete instantiations when teaching mathematics and science concepts (Chang, 2008). However, results of the present study and others (e.g., De Bock et al., 2011) suggest that such conclusions are premature. Although participants in the generic condition initially exhibited better transfer performance than those in the concrete condition when transfer score was analyzed as a continuous outcome using ANOVA (Hypothesis 2), this advantage was not present at any other time point, nor was it present in the nonparametric analysis at any time point (Hypothesis 3). Performance in the concrete and generic conditions was indistinguishable just three weeks after initial learning, primarily due to increased performance in the concrete condition. More importantly, participants in the fading condition exhibited the best transfer performance overall, suggesting that concrete instantiations may have some benefits for transfer in the long run, particularly if they are systematically "faded" into more abstract instantiations (Hypothesis 1). These results force us to take another look at the potential benefits of concrete instantiations.

Several decades ago, Bruner (1966) touted the benefits of concreteness fading in the teaching and learning of mathematics. According to his theory, mathematical concepts should be presented first in concrete, recognizable forms. Then, these forms should be "varyingly refined" to strip away irrelevant details until they are finally presented in the most economic, abstract symbolic form. This sequence is ideal, according to Bruner, because it increases learners' ability to understand, apply, and transfer what they learn. Bruner argued that it would be risky to skip the concrete form because learners who learn only through the abstract, symbolic form do not have a store of images in long-term memory to "fall back on" when they either forget, or are unable to directly apply the abstract, symbolic transformations they have learned. Since Bruner's time, many researchers and educators have supported the idea of starting with concrete instantiations then explicitly "decontextualizing" or "fading" away to the more abstract (Goldstone & Son, 2005; Gravemeijer, 2002; Lehrer & Schauble, 2002; Lesh, 1979; see also Koedinger & Anderson, 1998). Given this rich history of support, it is surprising that the present study was the first to experimentally test the benefits of gradually fading instantiations of mathematical concepts from concrete, recognizable forms to abstract forms that are only arbitrarily related to their referents.

We found that the fading condition produced the best transfer performance, particularly three weeks after the initial learning phase. Additionally, the concrete condition produced transfer performance that was ultimately just as good as that in the generic condition. Taken together, these results suggest that concrete instantiations may be beneficial for learning and transfer, particularly when they are "faded" into more abstract instantiations. Concrete materials may be particularly advantageous when introducing learners to novel concepts because they allow such concepts to be grounded in easily understood, real-world scenarios (Baranes et al., 1989; Goldstone & Son, 2005; Kotovsky et al., 1985). For example, in the present study, participants in the concrete condition were able to learn the rules by drawing on their knowledge of liquid measurements and remainders. Without these initial, grounded representations, participants would have just been learning to manipulate meaningless symbols (Glenberg et al., 2004). Equally important, if these initial, grounded representations had never been "faded" into more generic, abstract representations, then participants' knowledge of the rules would likely have remained context specific and tied to the measuring cups. Through fading, participants were able to "empty" the learned concept of its specific sensory and perceptual properties, so they could grasp its formal, abstract properties (Bruner, 1966).

Results suggest that concrete representations can be "faded" into more generalizable forms both during and after instruction. During instruction, educators can design lessons and activities that explicitly and gradually fade from concrete to abstract representations. For example, the fading condition in the present study first introduced the concept with representations that activated real-world knowledge (i.e., measuring cups). These concrete representations were then explicitly linked to an intermediate set of representations that were stripped of extraneous details, but maintained some level of iconicity (i.e., Roman numerals). Finally, the intermediate representations were faded into more flexible, abstract representations (i.e., arbitrary symbols). This instruction-based fading allowed learners to take advantage of the benefits of the grounded, recognizable concrete context, while also encouraging them to generalize beyond that specific context (Bruner, 1966; Goldstone & Son, 2005).

Another key aspect of the fading condition that may have contributed to its success is that it provided learners with three different instantiations. Researchers have long known that multiple representations have the potential to support learning and transfer (e.g., Ainsworth, 1999; Mayer, 2003; Scheiter et al., 2010; Sternberg, Toroff, & Grigorenko, 1998), particularly when accompanied by explicit prompts to compare those representations (Gick & Holyoak, 1983). Multiple representations not only encourage learners to extract the underlying structure of problems (Gick & Holyoak, 1983; Jitendra, Star, Rodriguez, Lindell, & Someki, 2011), but also support remembering because they prevent the learning environment from becoming too constant and predictable (Bjork & Bjork, 2011). In the present study, we tried to control for the number of representations by using a concrete condition that included three concrete instantiations that differed in surface form. However, all three of these concrete instantiations embodied the mathematical group of order 3 with remainders, so they may have been too similar to one another for learners to extract the more general structure (Jones, 2009). Future work should examine if other conditions with the same number of representations (e.g., three random instantiations varying in abstractness, or even three different abstract instantiations) would produce comparable results to the three-step fading sequence.

In addition to fading through instruction, the present results also suggest that it may be possible for concrete representations to fade into more generalizable forms without explicit instruction, but only over time and with additional engagement of the material. This more gradual fading process may rely on consolidation processes (including sleep) to alter the memory traces in ways that led to more abstract, flexible knowledge (cf., Gomez et al., 2006). It also may depend on the presence of multiple testing points, or other similar opportunities to engage the material (cf., Dean & Kuhn, 2006). We used multiple testing points in the present study because it matches standard classroom testing conditions, where students repeatedly engage the same material as they learn lessons, practice examples, take quizzes, exams, and high stakes tests. Although we did not provide participants with feedback during these tests, testing itself affects knowledge construction because it provides students with an opportunity to enact retrieval processes, which enhance learning and long-term retention (Bjork & Bjork, 2011; Karpicke & Roediger, 2007, 2008). Critics may argue that participants in the concrete condition could have benefited more from this testing effect compared to other participants because they had somewhat lower performance initially, so had more room for improvement. Although this is an important possibility to consider, it cannot fully explain the present results because participants in the fading condition had the highest performance initially and still exhibited improvements over time.

Another important caveat is that we only followed participants for three weeks. As Barnett and Ceci (2002) argued, transfer needs to withstand weeks, months, and even years between learning and testing in order to justify the time and effort it takes to teach students important concepts in school. Future studies should investigate transfer performance after months to determine the lasting effects of different instructional conditions. A longer time lapse would not only shed light on the role of consolidation in the abstraction process, but also increase the applicability to schools, where students have to retain information for several months to succeed on one high-stakes test per year. Extrapolating from our data, we predict that the advantage of the fading condition would increase over long time periods and that the concrete condition might eventually overtake the abstract condition.

Another question remains regarding the generalizability of our results. Our participants were highly educated adults. Research suggests that perceptually rich, concrete representations may not hinder the transfer performance of individuals with high knowledge as much as they hinder the transfer performance individuals with low knowledge (Goldstone & Sakamoto, 2003). Thus, it is possible that the initial difference in transfer performance between the concrete and generic conditions may have been larger and longer lasting in a "community sample." Interestingly, the present results tentatively suggest that generalizability across learners may be one of the advantages the fading condition has over the concrete and generic conditions. More specifically, the standard deviation in the fading condition was smaller than the standard deviation in the other two conditions. This pattern suggests that the fading condition benefits most learners, whereas the concrete and abstract conditions benefit some learners while failing others. Future studies should test individuals with varying levels of education to determine how the fading process affects their transfer performance.

Finally, it is important to note that our experiment was specifically designed to replicate and extend Kaminski et al. (2008), so it shares some of the limitations of the original study. For example, all conditions in the study used what Schwartz and Martin (2004) refer to as "tell-and-practice" instruction. Learning was achieved by reading rules, seeing examples of their use, and practicing them. Future studies should test to see if the fading process is equally beneficial in the context of other instructional approaches, such as the "inventing to prepare for learning" instruction (Schwartz & Martin, 2004) or the problematizing/modeling approach (Bransford, Franks, Vye, & Sherwood, 1989; Greer, 1997; Hiebert et al., 1996). Additionally, the concrete instantiations presented during learning were not the most realistic instantiations. They were schematic, colorless images of measuring cups that had minimal levels of perceptual detail. Because one of the arguments in favor of concrete materials is that they activate real-world knowledge (Kotovsky et al., 1985; Schliemann & Carraher, 2002), this lack of realism may have reduced the benefits of the concrete instantiations (but see McNeil et al., 2009 for evidence that lack of realism can improve performance).

Jones (2009) identified two additional methodological limitations of the Kaminski et al. (2008) study, and both apply to the current study. First, the transfer tasks may have been more similar to the abstract instantiation than they were to the concrete instantiations. The concrete instantiations behave like quantities that can be combined to form wholes and remainders. The generic instantiation and the transfer tasks do not share this feature, as the symbols do not represent quantities, nor are they combined. Indeed, Jones suggested that Kaminski et al.'s transfer task was simply "another version of the generic instantiation with a different contextualization" (p. 83). This may explain why participants in the abstract condition outperformed those in the concrete condition on the immediate transfer test. Second, the domain instructed (group of order 3) is relatively narrow. That is, participants were expected to learn how to manipulate three symbols according to given rules, and then apply those rules to a new context with the same number of symbols. No deeper understanding of underlying principles (e.g., commutative property) was tested.

The present study should not be taken as evidence against previous work that has revealed important limitations of concrete instantiations (e.g., Goldstone & Sakamoto, 2003; Kaminski et al., 2006, 2008; McNeil et al., 2009). Indeed, the limitations of concrete instantiations were shown in our first transfer phase. However, the present results extend previous results in important ways by showing that the abstract advantage dissipates when concrete instantiations are faded either through explicit instruction, or over time. The current evidence suggests that knowledge gained from concrete instantiations may start out being tied to the particular learning context, but through explicit fading and time or both it can become more generic in nature and thus, more transferable. Overall, the present study provides support for the educational practice of beginning with concrete instantiations of a concept and then "fading" away to the more abstract (e.g., Bruner, 1966; Goldstone & Son, 2005), and it suggests that concrete instantiations should be included in the educator's toolbox.

References

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: Wiley.

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131–152.

Ball, D. L. (1992). Magical hopes: manipulatives and the reform of mathematics education. *American Educator, 16*(14–18), 46–47.

Baranes, R., Perry, M., & Stigler, J. W. (1989). Activation of real-world knowledge in the solution of word problems. *Cognition and Instruction, 6*, 287–318. http://dx.doi.org/10.1207/s1532690xci0604_1.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612–637. http://dx.doi.org/10.1037/0033-2909.128.4.612.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth Publishers.

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). New approaches to instruction: because wisdom can't be told. In S. Vosniadou, & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470–497). New York: NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511529863.022.

Brown, M. C., McNeil, N. M., & Glenberg, A. M. (2009). Using concreteness in education: real problems, potential solutions. *Child Development Perspectives, 3*(3), 160–164. http://dx.doi.org/10.1111/j.1750-8606.2009.00098.x.

Bruner, J. S. (1966). *Towards a theory of instruction*. Cambridge, MA: Harvard University Press.

Burns, M. (1996). How to make the most of math manipulatives. *Instructor, 105*, 45–51.

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

Dean, D., & Kuhn, D. (2006). Direct instruction vs. discovery: the long view. *Science Education, 91*, 384–397.

De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., & Verschaffel, L. (2011). Abstract or concrete examples in learning mathematics? A replication and elaboration of Kaminski, Sloutsky, and Heckler's study. *Journal for Research in Mathematics*

Education, 42(2), 109–126, Retrieved from. http://hdl.handle.net/10067/876830151162165141.

DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development, 71*(2), 329–338. http://dx.doi.org/10.1111/1467-8624.00148.

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction, 6*, 41–57. http://dx.doi.org/10.1207/s1532690xci0601_2.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38. http://dx.doi.org/10.1016/0010-0285(83)90002-6.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology, 96*, 424–436. http://dx.doi.org/10.1037/0022-0663.96.3.424.

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46*, 414–466. http://dx.doi.org/10.1016/S0010-0285(02)00519-4.

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences, 14*, 69–110. http://dx.doi.org/10.1207/s15327809jls1401_4.

Gomez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17*, 670–674. http://dx.doi.org/10.1111/j.1467-9280.2006.01764.x.

Gravemeijer, K. (2002). Preamble: from models to modeling. In K. Gravemeijer, R. Lehrer, B. Oers, & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 7–22). Dordrecht, The Netherlands: Kluwer.

Greer, B. (1997). Modelling reality in mathematics classrooms: the case of word problems. *Learning and Instruction, 7*(4), 293–307. http://dx.doi.org/10.1016/S0959-4752(97)00006-6.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: a theory of cognitive interest in science learning. *Journal of Educational Psychology, 3*, 414–434. http://dx.doi.org/10.1037/0022-0663.90.3.414.

Hiebert, J., Thomas, C. P., Fennema, E., Fuson, K. C., Human, P., Murray, H., et al. (1996). Problem solving as a basis for reform in curriculum and instruction: the case of mathematics. *Educational Researcher, 25*, 12–21. http://dx.doi.org/10.3102/0013189X025004012.

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction, 21*, 731–745. http://dx.doi.org/10.1016/j.learninstruc.2011.04.002.

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

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2006). Do children need concrete instantiations to learn an abstract concept? In R. Sun, & N. Miyake (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 411–416) Mahwah, NJ: Lawrence Erlbaum Associates.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*, 454–455. http://dx.doi.org/10.1126/science.1154659.

Karpicke, J. D., & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57*, 151–162. http://dx.doi.org/10.1016/j.jml.2006.09.004.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968. http://dx.doi.org/10.1126/science.1152408.

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: the early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments, 5*, 161–179. http://dx.doi.org/10.1080/1049482980050111.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problem hard? Evidence from the tower of Hanoi. *Cognitive Psychology, 17*, 248–294. http://dx.doi.org/10.1016/0010-0285(85)90009-X.

Lehrer, R., & Schauble, L. (2002). Symbolic communication in mathematics and science: co-constituting inscription and thought. In E. D. Amsel, & J. P. Byrnes (Eds.), *Language, literacy, and cognitive development. The development and consequences of symbolic communication* (pp. 167–192). Mahwah, NJ: Lawrence Erlbaum Associates.

Lesh, R. (1979). Mathematical learning disabilities: considerations for identification, diagnosis, remediation. In R. Lesh, D. Mierkiewicz, & M. G. Kantowski (Eds.), *Applied mathematical problem solving* (pp. 111–180). Columbus, OH: ERIC.

Martin, T. (2009). A theory of physically distributed learning: how external environments and internal states interact in mathematics learning. *Child Development Perspectives, 3*(3), 140–144. http://dx.doi.org/10.1111/j.1750-8606.2009.00094.x.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science, 29*, 587–625. http://dx.doi.org/10.1207/s15516709cog0000_15.

Mayer, R. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction, 13*, 125–139.

McGaugh, J. L. (2000). Memory: a century of consolidation. *Science, 287*, 248–251. http://dx.doi.org/10.1126/science.287.5451.248.

McNeil, N. M., Uttal, D. H., Jarvin, L., & Sternberg, R. J. (2009). Should you show me the money? Concrete objects both hurt and help performance on mathematics problems. *Learning and Instruction, 19*, 171–184. http://dx.doi.org/10.1016/j.learninstruc.2008.03.005.

Montessori, M. (1917). *The advanced Montessori method*. New York, NY: Frederick A. Stokes.

Piaget, J. (1970). *Science of education and the psychology of the child*. New York, NY: Orion Press.

Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. InGlaser, R. (Ed.), (1987). *Advances in instructional psychology, Vol. 3* (pp. 41–96). Hillsdale, NJ: Erlbaum.

Robertson, E. M., Pascual-Leone, A., & Miall, R. C. (2004). Current concepts in procedural consolidation. *Nature Reviews Neuroscience, 5*, 1–7. http://dx.doi.org/10.1038/nrn1426.

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition, 34*, 344–355. http://dx.doi.org/10.3758/BF03193412.

Scheiter, K., Gerjets, P., & Schuh, J. (2010). The acquisition of problem-solving skills in mathematics: how animations can aid understanding of structural problem features and solution procedures. *Instruction Science, 38*, 487–502. http://dx.doi.org/10.1007/s11251-009-9114-9.

Schliemann, A. D., & Carraher, D. W. (2002). The evolution of mathematical reasoning: everyday versus idealized understandings. *Developmental Review, 22*, 242–266. http://dx.doi.org/10.1006/drev.2002.0547.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: the hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction, 22*, 129–184. http://dx.doi.org/10.1297/s1532690xci2202_1.

Sloutsky, V. M., Kaminski, J. A., & Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin and Review, 12*, 508–513. http://dx.doi.org/10.3758/BF03193796.

Smith, J. P. (1996). Efficacy and teaching mathematics by telling: a challenge for reform. *Journal for Research in Mathematics Education, 27*, 387–402, Retrieved from. http://www.jstor.org/stable/749874.

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: short-cutting abstraction in children's object categorizations. *Cognition, 108*, 626–638. http://dx.doi.org/10.1016/j.cognition.2008.05.002.

Sternberg, R. J., Toroff, B., & Grigoreko, E. L. (1998). Teaching triarchically improves school achievement. *Journal of Educational Psychology, 90*, 374–384.

Stickgold, R., & Walker, M. (2004). To sleep, perchance to gain creative insight? *Trends in Cognitive Sciences, 8*, 191–192. http://dx.doi.org/10.1016/j.tics.2004.03.003.

Thompson, P. W. (1992). Notations, conventions, and constraints: contributions to effective uses of concrete materials in elementary mathematics. *Journal for Research in Mathematics Education, 23*, 123–147, Retrieved from. http://www.jstor.org/stable/749497.

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: a new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology, 18*, 37–54. http://dx.doi.org/10.1016/S0193-3973(97)90013-7.

Wagner, U., Gals, S., Halder, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 427*, 352–355. http://dx.doi.org/10.1038/nature02223.

Wecker, C., & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: the role of peer monitoring during the fading of instructional scripts. *Learning and Instruction, 21*, 746–756. http://dx.doi.org/10.1016/j.learninstruc.2011.05.001.