Journal of Educational Psychology
2013, Vol. 105, No. 3, 666–682

0022-0663/13/$12.00 DOI: 10.1037/a0032095

Integrating Formal and Grounded Representations in Combinatorics Learning

David W. Braithwaite and Robert L. Goldstone
Indiana University Bloomington

The terms *concreteness fading* and *progressive formalization* have been used to describe instructional approaches to science and mathematics that use grounded representations to introduce concepts and later transition to more formal representations of the same concepts. There are both theoretical and empirical reasons to believe that such an approach may improve learning outcomes relative to instruction employing only grounded or only formal representations (Freudenthal, 1991; Goldstone & Son, 2005; McNeil & Fyfe, 2012; but see Kaminski, Sloutsky, & Heckler, 2008). Two experiments tested the effectiveness of this approach to instruction in the mathematical domain of combinatorics, using outcome listing and numerical calculation as examples of grounded and formal representations, respectively. The study employed a pretest-training, posttest design. Transfer performance, that is, participants' improvement from pretest to posttest, was used to assess the effectiveness of instruction received during training. In Experiment 1, transfer performance was compared for 4 types of instruction, which differed only in the types of representation they employed: pure listing (i.e., listing only), pure formalism (i.e., numerical calculation only), list fading (i.e., listing followed by numerical calculation), and formalism-first (i.e., listing introduced after numerical calculation). List fading instruction led to transfer performance on par with pure formalism instruction and higher than formalism-first and pure listing instruction. In Experiment 2, an enhanced version of list fading training was again compared to pure formalism. However, no difference in transfer performance due to training was found. The results suggest that combining grounded and formal representations can be an effective approach to combinatorics instruction but is not necessarily preferable to using formal representations alone. If both grounded and formal representations are employed, the former should precede rather than follow the latter in the instructional sequence.

Keywords: mathematics instruction, grounded and formal representations, concreteness fading, progressive formalization, transfer

Supplemental materials: http://dx.doi.org/10.1037/a0032095.supp

Many mathematical concepts may be represented in multiple formats (Lesh, Post, & Behr, 1987). For example, numbers may be represented as points on a line, strings of numerals, arrays of dots (for natural numbers), or pie charts (for proper fractions); functions as graphs or equations; and combinatorial relations as lists of outcomes, tree diagrams, or algebraic formulas. Mathematical representations may be classified as grounded or formal. Grounded representations are those that can be partially or completely understood on the basis of everyday human experience or knowledge. Such representations may be perceptually grounded (e.g., photos or realistic drawings) or contextually grounded (e.g., verbal descriptions of realistic situations). Formal representations are those whose meaning comes not from everyday experience or knowledge, but from systems of explicit rules. Numerals and algebraic equations are examples of formal representations. Our distinction between formal and grounded representations corresponds closely to Stenning's (2002) distinction between indirect and direct representations, and also broadly to Schnotz's (2005) distinction between descriptive and depictive representations.

Much research has investigated the relative advantages of learning mathematical concepts via grounded or formal representations, with mixed results. There is some evidence that contextual grounding can facilitate problem solving and reduce conceptual error (McNeil, Uttal, Jarvin, & Sternberg, 2009; Nuñes, Schliemann, & Carraher, 1993). Problem context appears to enable access to informal solution strategies that, for some problems, are more efficient and less error prone than formal strategies (Koedinger & Nathan, 2004). On the other hand, learning mathematical principles in context has the potential to limit their generality. For example, Bassok and Holyoak (1989) found that students learning about constant rate problems in physics showed less transfer of learning than those who learned in a decontextualized algebraic format. Similarly, Ross (1987, 1989) demonstrated that story problem context in which students learn about probability principles can have a negative influence on their ability to correctly apply those principles in novel contexts.

This article was published Online First March 25, 2013.

David W. Braithwaite and Robert L. Goldstone, Department of Psychological and Brain Sciences, Indiana University Bloomington.

This research was in part supported by a Research and Evaluation on Education in Science and Engineering grant (0910218) from the National Science Foundation. Portions of the data and analyses presented in this article were presented during the poster session of the 2011 annual meeting of the Cognitive Science Society in Boston, Massachusetts.

Correspondence concerning this article should be addressed to David W. Braithwaite, Department of Psychological and Brain Sciences, Indiana University, 1101 East 10th Street, Bloomington, IN 47408. E-mail: dwbraith@indiana.edu

With regard to perceptual grounding, results have been similarly mixed. Several studies have found better learning outcomes following instruction combining images with text relative to text alone (Glenberg, Willford, Gibson, Goldberg, & Zhu, 2012; Mayer, 2002; Scheiter, Gerjets, & Catrambone, 2006). Mayer (2002) attributes this effect to deeper processing by learners when they receive information through multiple sensory channels. However, some studies have failed to find such an effect (e.g., Schnotz & Bannert, 2003; see also Shneiderman, Mayer, McKay, & Heller, 1977). There is considerable evidence that perceptually rich images or animations may distract learners from the mathematical concepts they are meant to be learning (Brown, McNeil, & Glenberg, 2009; Goldstone & Sakamoto, 2003; Mayer, Hegarty, Mayer, & Campbell, 2005; McNeil et al., 2009; Scheiter et al., 2006; Sloutsky, Kaminski, & Heckler, 2005). By contrast, visual representations that emphasize structure over perceptual detail, such as idealized schematics and animations, appear to benefit learning (Goldstone & Sakamoto, 2003; Hegarty & Kozhevnikov, 1999; Scheiter et al., 2006). However, just as in the case of contextual grounding, even structural visual representations, relative to amodal symbol systems, may limit learners' ability to transfer mathematical principles to novel contexts (Kaminski et al., 2008).

Regarding transfer, two qualifications are necessary. First, although previous research (Bassok & Holyoak, 1989; Kaminski et al., 2008) has shown advantages of formal or abstract presentation for transfer in some situations (e.g., for textbook-style algebra story problems), transfer may work differently for other mathematical domains and in other contexts. Further research across a wider variety of concepts and content is needed to understand how general these advantages may be. Second, it is important to distinguish transfer involving direct application of learned concepts, here termed *near* transfer, from *adaptive* or *far* transfer (Barnett & Ceci, 2002; Hatano & Inagaki, 1986; Schwartz, Chase, & Bransford, 2012), in which learned concepts or methods must be modified or adapted. Although studies demonstrating advantages of abstract or formal instruction have often involved near transfer, grounded instruction may be relatively more effective for adaptive or far transfer. In particular, grounded presentation of mathematical concepts appears to foster conceptual understanding (McNeil et al., 2009; Schwartz, Chase, Oppezzo, & Chin, 2011), which in turn is prerequisite to far transfer (Rittle-Johnson & Alibali, 1999). Also, solution methods may be more easily modified or adapted when represented in a grounded rather than formal way (Schwartz & Black, 1996; Schwartz & Martin, 2004). Finally, perhaps due precisely to the compactness and efficiency of formal methods, instruction therein may encourage learners to apply such methods automatically and by rote, rather than think about them and adapt them when necessary (Nathan, 2012; Schwartz et al., 2012, 2011; Schwartz & Martin, 2004).

Another factor that may affect the relative utility of grounded and formal representations is the complexity of the domain and task. Grounded representations may help learners to adapt previously learned formal methods for more complex tasks (Schwartz & Black, 1996). Grounded representations may also facilitate application of informal solution strategies that, in some situations, may be less error prone than formal ones (Koedinger & Nathan, 2004; Nuñes et al., 1993). However, the latter advantage may apply only to relatively simple problems. For more complex problems, formal solution strategies appear to be more effective in some domains, and application of these can be facilitated by decontextualized formal representations (Koedinger, Alibali, & Nathan, 2008). Thus, both grounded and formal representations may help learners in dealing with complexity.

The preceding discussion suggests that grounded and formal representations may have complementary strengths. Contextually or perceptually grounded representations appear to allow learners to bootstrap everyday knowledge and skills to support conceptual understanding and flexible adaptation. Formal representations, by contrast, can be useful tools for generalization to novel contexts, and can support powerful and/or efficient operations, especially for complex problems. Given these complementary sets of benefits, including both grounded and formal representations into mathematics learning might provide the best of both worlds. This conclusion is consistent with several theories of learning via multiple representations, according to which learners exposed to multiple representations of a single concept may integrate those representations to construct a unified conceptual understanding (Ainsworth, 1999, 2006; Mayer, 2002; Schnotz, 2005). However, multiple representations could offer complementary benefits to learners even if no unified concept is formed on their basis. In this case, understanding how to translate between representations when necessary is likely to be critical to their effective use (Brenner et al., 1997; Lesh et al., 1987; Pape & Tchoshanov, 2001).

One approach to combining grounded and formal representations in mathematics instruction is to begin with grounded representations and subsequently transition to formal ones, an approach variously termed *progressive formalization* (Freudenthal, 1991; Nathan, 2012; van Reeuwijk, 2001) and *concreteness fading* (suggesting that concrete details are initially present but then fade away; Goldstone & Son, 2005; Landy & Goldstone, 2007; McNeil & Fyfe, 2012; Son & Goldstone, 2009). Freudenthal (1991) associated progressive formalization with a more general process he called "mathematizing," that is, identifying mathematical structure in specific situations, objects, processes, etc., and developing notation to represent that structure. As such, he perceived progressive formalization as essential for understanding formal mathematical notations, which only have meaning by reference to concrete situations. Similarly, Son and Goldstone (2009) pointed out that when concrete representations are encountered before more idealized, less grounded ones, learners may leverage their intuitive understanding of the former to understand the latter as well, while still benefiting from the power and flexibility of the latter. Additionally, by first dealing with concrete instantiations of mathematical structures, learners are better able to understand why formal notations are needed, that is, to generalize across superficially different situations and to facilitate efficient manipulation (Freudenthal, 1991).

Concreteness fading/progressive formalization is not the only way in which grounded and formal representations might be combined in mathematics instruction. Another approach, termed *formalisms first*, is to begin by presenting formal representations and only later introduce grounded ones, much as one might first describe a principle and later provide an example or specific application. This approach is prevalent among educators in technical domains including mathematics (Nathan, 2012; Nathan, Long, & Alibali, 2002), apparently reflecting a belief that formal knowledge is a prerequisite for concrete applications. For example, Nathan and Koedinger (2000) found that mathematics educators perceived algebra story problems to be more difficult for students than equivalent equation problems, presumably because story problems could only be solved by first translating them into equations. There is, however, some evidence against the above belief: People appear to be capable of solving at least basic mathematics problems in concrete, realistic contexts, without employing formal representations (Koedinger & Nathan, 2004; Nuñes et al., 1993).

Of course, evidence that learners *can* handle concrete applications without using formal methods is not evidence that educators *ought* to present the former before the latter. One consideration favoring the latter, prescriptive claim is that learners who encounter formal representations before grounded ones may be less likely to engage with or deeply process the latter (Schwartz et al., 2012, 2011). For example, Schwartz et al. (2011) found that students who had learned a formula for calculating ratios, when later given related story problems, focused on plugging numbers into the formula, and tended not to represent the problems in terms of their ratio structure. A second consideration relates to the tendency of formal representations to specify less detail, leaving more room for interpretation, than grounded ones (Stenning, 2002). If formal representations are presented first, learners may interpret them in ways inconsistent with subsequently presented grounded representations, causing confusion that is less likely if the more detailed grounded representations are encountered first (Schnotz, 2005). Finally, as mentioned above, grounded representations can facilitate understanding of problem structure, serving as scaffolding for less grounded representations presented subsequently. Goldstone and Son (2005) found evidence for this view in a study of complex systems learning. Participants who were exposed to a perceptually rich simulation that later "faded" to a perceptually idealized simulation showed both superior learning and superior transfer, relative to participants who saw the idealized simulation before the perceptually rich one.

Even the idealized representations used by Goldstone and Son (2005) were depictive graphics and therefore perceptually grounded. It is uncertain whether the benefits they found for a fading approach also apply to fading from concrete to fully formal representations: an essential issue, considering the extensive use of formalisms such as arithmetic and algebraic notation in mathematics education and practice. Two more recent studies (Kaminski et al., 2008; McNeil & Fyfe, 2012) bear on this question. Both investigated transfer of knowledge of an abstract algebra concept following instruction that employed either abstract symbols only, a type of formalism, or meaningful icons followed by abstract symbols. Kaminski et al. (2008) found the latter approach, an implementation of concreteness fading, to be less effective than the former, formalisms-only approach, contrary to the findings of Goldstone and Son (2005). However, McNeil and Fyfe (2012) found superior learning outcomes for the concreteness fading approach when an intermediate representation was added between the grounded and formal ones. This intermediate representation involved formal symbols that were also perceptually grounded,[1] and so may have aided participants in making correspondences between the purely grounded and purely formal representations. Thus, the utility of grounded representations to learners may depend on whether they are able to map between grounded and formal representations. In line with this interpretation, Kaminski (2006) showed that instruction using only grounded representations was as effective as that using only formalisms, provided that participants were given the correspondences between elements of the representation used during instruction and elements of the transfer domain.

The studies just described (Goldstone & Son, 2005; Kaminski et al., 2008; McNeil & Fyfe, 2012) assessed the effectiveness of concreteness fading/progressive formalization with respect to near transfer as defined above, that is, application of learned concepts to situations isomorphic to those encountered during instruction. If, as suggested earlier, grounding is particularly beneficial with respect to adaptive or far transfer (i.e., adaptation of learned concepts in situations that are structurally related but not identical to those encountered during instruction), then concreteness fading/progressive formalization is likely to show a particular advantage over formalisms-first or formalisms-only approaches on measures of adaptive or far transfer. This possibility awaits direct empirical verification.

Considering the relative paucity of empirical studies directly comparing concreteness fading approaches to plausible alternatives, as well as the open questions just described, further research is needed to better understand the utility of such approaches to mathematics education. The present study extends existing research in three main respects. First, a concreteness fading approach was applied to a novel domain, combinatorics, thus affording a test of whether previous findings (Goldstone & Son, 2005; Kaminski et al., 2008; McNeil & Fyfe, 2012) generalize to other mathematical domains. Second, this approach was employed to fade from rather abstract grounded representations (i.e., lists of possible outcomes) to purely formal representations (i.e., numerical calculations) without any intermediate representation. Thus, the study examines the effectiveness of fading to abstract formalisms, rather than idealized graphics like those employed by Goldstone and Son (2005), when no intermediate representation such as that employed by McNeil and Fyfe (2012) is provided. Third, in contrast to previous studies that have focused mainly on near transfer, learning was assessed with measures of both near and far transfer, allowing a test of the above hypothesis that fading would especially facilitate far transfer, relative to formalisms-first or formalisms-only approaches.

The next section contains a brief review of recent psychological research regarding the combinatorics domain, with an emphasis on the role of different representations in learning and problem solving. In the following section, a specific implementation of concreteness fading in this domain is described, and the goals of the present research are laid out in more detail.

The Combinatorics Domain

Combinatorics, defined informally as the "science of counting," involves calculation of the number of possible joint outcomes of several separate events. Combinatorics is an important topic for psychological research for two reasons. First, combinatorial reasoning involves systematic consideration of possibilities, deliber-

[1] The Roman numerals I, II, and III. These Roman numerals obviously belong to a mathematical symbol system, but are also perceptually grounded in that they use numerosity of vertical lines to represent number. The availability of such an intermediate representation may be idiosyncratic to the domain used by McNeil and Fyfe (2010), and such "grounded symbols" may not be available in all domains. Even for the Roman numerals, their perceptual grounding breaks down starting at IV.

ately separated from the actual state of the world. As such, development of combinatorial reasoning is considered an important foundation of abstract thought more generally (Inhelder & Piaget, 1958). Second, combinatorics plays an important role in the study of probability and statistics, mathematical domains with many practical applications in a wide variety of fields.

Combinatorics is also an ideal domain for investigating the issues discussed above, due to the wide variety of representations employed therein. Students of combinatorics learn algebraic formulas for various types of problems, such as permutations, combinations, and sampling with replacement. Besides such formulas, representations such as verbal descriptions, pictures, concrete manipulatives, tree diagrams, and outcome lists may also be used to conceptualize and solve combinatorics problems. Figure 1 shows examples of numerical calculation and outcome listing in the context of a permutations problem.

In the terminology of the preceding section, numerical calculations (see Figure 1A) constitute formal representations. Although outcome lists (see Figure 1B) appear at first glance rather abstract, they are in fact perceptually grounded in two respects. First, they represent numbers by numerosities. That is, the six possible outcomes of the race are represented by the six rows of the list, rather than by the numeral 6, the three horses by three letters rather than by the numeral 3, and so on. Second, temporal sequence is represented by spatial sequence. That is, the order in which horses finish the race is represented by the order of letters on the page from left to right. In general, outcome lists are grounded because they use spatial information to encode structural features (i.e., number and order) in nonarbitrary ways.

A number of studies have investigated the roles of different representations on learning and problem solving in combinatorics and related domains such as probability. One main finding is that visual representations, such as tree diagrams, may serve as bridges between contextualized descriptions and formal algebraic or numerical expressions (Corter & Zahner, 2007; Figueiras & Cañadas, 2010; Zahner & Corter, 2010). These representations are often used in the earlier stages of problem solution (e.g., to help select solution strategies) rather than in the actual calculation of solutions. Zahner and Corter's (2010) study of adult statistics learners concluded that participants employed visual representations mainly to help conceptualize the mathematical structure of the problems, especially for nonroutine problems for which formal strategies were not immediately obvious. Figueiras & Cañadas (2010) reached a similar conclusion in the domain of combinatorics, with regard to younger (11- to 12-year-old) learners without prior exposure to formal methods.

A second important result, surprising in light of the above findings, is that instructional approaches incorporating grounded representations often produce no better learning outcomes than approaches using only formal representations (Belenky & Nokes, 2009; Kolloffel, Eysink, de Jong, & Wilhelm, 2009; Scheiter et al., 2006). For example, Kolloffel et al. (2009) exposed high school students to computer-based instruction in probability principles using tree diagrams only, arithmetic only, text only, text plus arithmetic, or diagrams plus arithmetic. On a posttraining assessment, participants in the two conditions involving tree diagrams displayed worse problem-solving skill, and only equal conceptual knowledge, compared to those in the text plus arithmetic condition. To the contrary, however, participants in a study of Berthold and Renkl (2009) showed superior conceptual understanding of probability principles after receiving instruction using both tree diagrams and numerical expressions, though only when the instruction included aids to highlight the connections between representations. This result suggests that probability/combinatorics learners may only benefit from grounded representations if they are also exposed to formal representations and explicitly shown the connections between the two.

Goals of the Present Study

The present study investigated the effectiveness of applying concreteness fading/progressive formalization to instruction in combinatorics. In combinatorics and probability, instruction involving grounded representations alone has shown little or no advantage over instruction involving formal representations alone. However, the instructional effectiveness of fading from one representation to the other has not been tested in these domains. It was hoped that fading would increase the benefits of grounded representations, relative to the findings of previous studies. The study focused on one type of combinatorics problem, permutations, and used outcome listing and numerical calculations (see Figure 1) as representatives of grounded and formal representations, respectively.

Suppose three horses, Amber, Beryl, and Crystal, run in a race. Assuming all 3 finish the race, one will come in first, one second, and one third. How many different arrangements of first, second, and third place are possible?

A

$$3! = 3 \times 2 \times 1 = 6$$

B

A B C	B A C	C A B
A C B	B C A	C B A

Figure 1. A permutations problem. (A) A solution of the problem using a formal representation, numerical calculation. (B) A solution of the same problem using a grounded representation, outcome listing.

It was proposed above that grounded and formal representations offer complementary benefits to mathematics learners. Applying that proposal to the specific representations just mentioned leads to several implications. First, as a type of grounded representation, outcome listing ought to aid conceptual understanding, and to be especially helpful when learners need to conceptualize novel problems for which no formal solution method is known. These implications are consistent with previous findings regarding the use of visual representations such as outcome lists by combinatorics/probability learners (Berthold & Renkl, 2009; Corter & Zahner, 2007; Figueiras & Cañadas, 2010; Zahner & Corter, 2010). Second, numerical calculations are one way to capture generalities about combinatorial structures across a variety of specific problems, and so might be expected to facilitate transfer from problems encountered during learning to superficially different but structurally isomorphic problems, as has been observed for formal representations in similar domains such as algebra (Bassok & Holyoak, 1989; Kaminski et al., 2008). Finally, combinatorics problems are more complex than the arithmetic and algebra problems employed in the aforementioned studies. Both grounded and formal representations may offer different benefits with respect to complex mathematics problems (Koedinger et al., 2008; Schwartz & Black, 1996).

The preceding discussion of concreteness fading/progressive formalization suggested that these approaches are most likely to be effective when learners are able to make connections between the grounded or concrete representations seen initially and the formal or idealized ones seen later (Kaminski, 2006; McNeil & Fyfe, 2012). Outcome listing and numerical calculation are amenable to making such connections due to their structural correspondences. For example, the numeral 3 in the numerical calculation of Figure 1A corresponds to the letters *A*, *B*, and *C* appearing in the far left position of the outcome list of Figure 1B, the numeral 2 to the sets of different letters appearing in the second position of the outcome list, and so on. These correspondences suggest that outcome lists and numerical calculations are good candidates for application of a fading approach, provided that learners are able to notice and understand the correspondences.

Our specific instantiation of the fading approach presented worked examples solved with outcome listing, followed by other worked examples solved with numerical calculation. This approach, termed *list fading*, was compared to three other approaches, all of which used the same example problems but employed different representations for the solutions. The *pure formalism* approach used numerical calculation for all solutions. The *pure listing* approach using outcome listing for all solutions. The *formalism-first* approach used numerical calculation for the earlier examples and outcome listing for the later examples. Importantly, this approach involved the same instructional materials as the list fading approach, differing only in order of presentation, so comparing these two approaches would allow us to determine whether any advantage found for the list fading approach was due to fading per se, or simply to the use of multiple representations.

Two predictions were made regarding the differential effects of the four instructional approaches on learning outcomes. First, the list fading condition was predicted to lead to better learning outcomes than both pure conditions. If, as suggested above, outcome lists and numerical calculations offer learners complementary benefits, then combining both types of representation should offer the best of both worlds. Second, the list fading condition was also predicted to be superior to the formalism-first condition. Premature presentation of numerical calculation in the formalism-first condition might prevent learners from deeply processing the subsequently encountered grounded representation (Schwartz et al., 2012, 2011), whereas beginning with the grounded representation should allow learners later to use it as scaffolding to aid in understanding the formal representation (Goldstone & Son, 2005).

A third prediction was made regarding the effectiveness of different training conditions with respect to near and far transfer problems. The advantages of list fading instruction were expected to be greatest for far transfer problems, on the basis that grounded representations support generation/adaptation of formal solution methods for novel problem types in combinatorics and probability (Corter & Zahner, 2007; Figueiras & Cañadas, 2010; Zahner & Corter, 2010) and in mathematics more generally (Schwartz & Black, 1996; Schwartz & Martin, 2004). Also, we expected participants who had been exposed to grounded as well as formal representations to display greater conceptual understanding of the latter (Berthold & Renkl, 2009; McNeil et al., 2009; Schwartz et al., 2011), and thus greater ability to adapt their solution methods to solve far transfer problems (Rittle-Johnson & Alibali, 1999; Rittle-Johnson, Siegler, & Alibali, 2001).

Experiment 1

Method

Participants. One hundred eleven undergraduate students from Indiana University participated in the experiment in partial fulfillment of a course requirement. Of these, 28 were assigned to the list fading condition, 29 to formalism-first, 27 to pure formalism, and 27 to pure listing, as described below.

Materials. Two sets of permutations problems were developed to serve as pretest and posttest. The problems in one set involved a farmer rotating crops, whereas those in the other set involved a groom arranging a standing order for his groomsmen. Each set included four problems. The first two problems in each set, referred to as near transfer problems, were standard permutations problems requesting the number of permutations of a given set and could be solved by direct application of the methods shown during instruction. For example, one such problem requested the number of possible crop rotation sequences that use each of four crops for exactly 1 year over a 4-year period. The third and fourth problems, referred to as far transfer problems, required adaptation of the methods shown during instruction. For example, one such problem requested the number of possible crop rotation sequences that use four from among six crops, and use each of these for exactly 1 year over a 4-year period. Each problem in one set corresponded to a problem in the other set that involved the same numbers and mathematical structure and hence had the same correct answer. The specific order of corresponding problems was identical between sets, with near transfer problems appearing before far transfer problems. A list of all test problems is provided in the supplemental materials (Section A).

Another set of permutations problems was developed to serve as worked examples during training. This set consisted of two pairs of permutation problems. Both problems within each pair used the same back story, which involved either horses running in a race or a traveler planning to visit several cities in sequence. The two pairs, however, used different back stories. The problems within each pair involved the same mathematical structure but different numbers (i.e., one problem involved permuting three items and the other four items), whereas the two pairs were mathematically isomorphic to each other. A list of all training examples is provided in the supplemental materials (Section B).

Several standard solutions were developed for each of the problems in the training problem set (see Figure 2). The solutions were presented as slide shows, in which a solution representation and a corresponding verbal explanation were built up incrementally. The explanations were shown in a pane at the left side of each slide, and the representations were shown in a pane at the right. At the beginning of each solution, only the first paragraph in the left pane and the corresponding parts of the representation in the right pane were displayed. Each subsequent slide added one paragraph to the right pane and the corresponding representation elements to the left pane. Figures 2A, 2B, and 2C show the final slides from their respective slideshows, and Figure 2D shows an intermediate slide.

The solutions belonged to one of three types: formalism based (Figure 2A), listing based (Figure 2B), or mixed (Figure 2C). The verbal explanations were identical between the three types of solution, except that the listing-based solutions did not include any reference to "factorials," which appeared at the end of the verbal explanations for the formalism-based and mixed solutions (Figures 2A and 2C). The solution representations were either numerical calculations for the formalism-based solutions (Figure 2A), outcome lists for the listing-based solutions (Figure 2B), or both for the mixed solutions (Figure 2C). For the solutions involving lists (i.e., both listing based and mixed solutions), branching tree diagrams were used each time a new column was added to the list, in order to illustrate the relations between the letters in each column and those in the column to the left (Figure 2D). However, the tree branches disappeared on the slides following their presentation, so that only a list without tree branches was visible on the final slide (Figures 2B and 2C). For the mixed solutions, the numbers in the numerical calculations were aligned with the columns of the outcome lists to highlight the structural analogy between them.

Procedure. Participants were first presented one of the test problem sets, as a pretest. Which test problem set served as pretest was selected randomly; in subsequent analyses, this selection is referred to as *pretest–posttest rotation*. The problems were presented one at a time through a computer interface. Below each problem were two text entry fields in which participants were to show, respectively, their work and their final answers. Participants were told that they could show their work by writing numerical calculations, listing possible outcomes, explaining reasoning, or any combination of these. This instruction included an explanation of the standard keyboard symbols for exponentiation and factorial (i.e., ^ and *!*, respectively). Participants were provided with an on-screen calculator that they could use freely. Participants could not proceed if the field for work or answer was left blank. Once a problem was completed, it was not possible to return to it later. There was no time limit for completing the problems.

Next, the training example problems were presented. The order of the two training problem pairs was determined randomly, so that

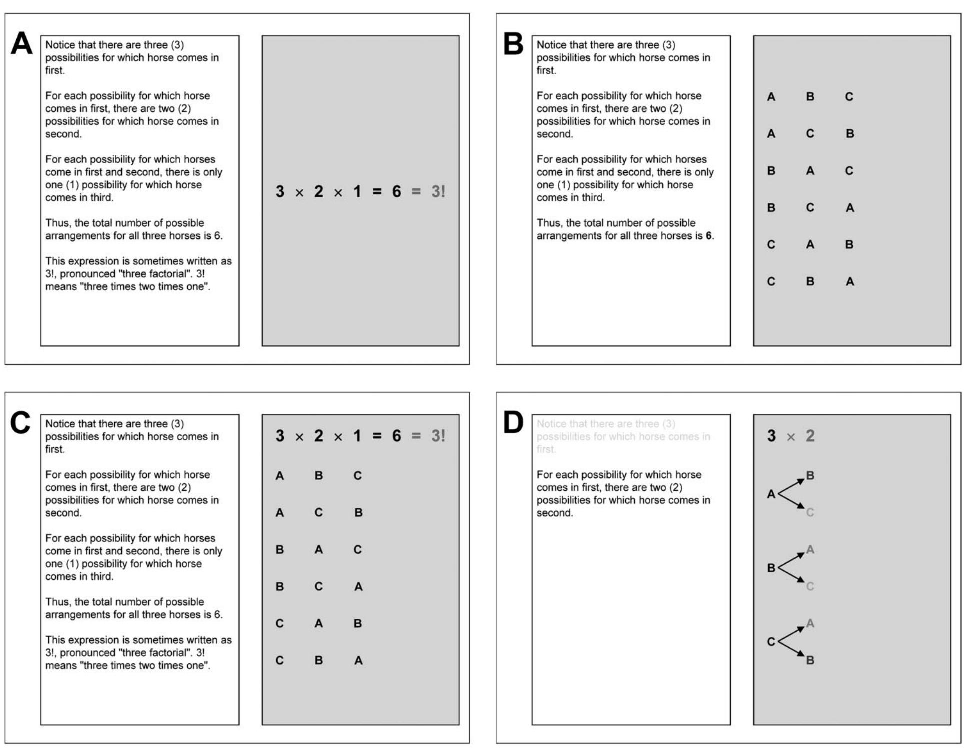

Figure 2. Standard solutions to a permutation problem from the training problem set (Experiment 1). (A) Final slide from formalism-based solution. (B) Final slide from listing-based solution. (C) Final slide from mixed solution. (D) Intermediate slide from mixed solution, illustrating the use of a tree diagram.

some participants saw the horse problems first and the traveler problems second, whereas other participants saw the reverse. This assignment is referred to as *training problem sequence*. The training problems were presented in the same way as in the pretest, except that participants were explicitly instructed to solve each problem using a specific method, either numerical calculation or outcome listing. A text description of the designated method, including either an algorithm for systematically listing outcomes or a procedure for calculating the number of outcomes arithmetically, was displayed below each problem. No general formula was presented for the calculation procedure, and no illustrations or examples were provided for either method. After each problem, participants were shown the slideshow for a standard solution before they proceeded to the next problem.

The designated solution methods and the types of solutions shown depended on training condition. Participants were assigned randomly to one of four conditions: pure formalism, pure listing, list fading, or formalism-first. The assigned methods and the solutions shown for each problem in each condition are summarized in Table 1. In the pure formalism condition, participants were instructed to use numerical calculation, and were shown the formalism-based solution, for each problem. In the pure listing condition, participants were instructed to use outcome listing, and were shown the listing-based solution, for each problem. In the list fading condition, participants were instructed to use outcome listing for the first two problems and numerical calculation for the second two. They were shown the listing-based solution for the first problem, the mixed solutions for the second and third, and the formalism-based condition for the fourth. In the formalism-first condition, participants were instructed to use numerical calculation for the first two problems and outcome listing for the second two. They were shown the formalism-based solution for the first problem, the mixed solutions for the second and third, and the listing-based solution for the fourth. Thus, the mixed solutions only appeared for the second and third problems in the list fading and formalism-first conditions. They were intended to assist participants in transitioning from one representation to the other by highlighting their structural correspondences.

After the training, the posttest problem set (i.e., whichever test problem set had not been used as pretest) was presented in the same way as the pretest problem set. Finally, participants were asked several multiple-choice comprehension questions. These questions tested understanding of combinatorics principles, such as the fundamental principle of counting, which were not explicitly mentioned during the training phase. The full text of these questions appears in the supplemental materials (Section C).

Coding. Responses to the test problems assigned a score of 1 for correct answers and 0 for incorrect answers. Before assigning these scores, responses were corrected for calculation errors based on participants' shown work, changing incorrect answers to correct ones if the corresponding work contained an arithmetic expression that yielded the correct answer and contained no other arithmetic expression. The reason for this correction is that our interest was in whether participants understood how to solve the problems rather than in whether they conducted the relevant calculations correctly. Fourteen responses, eight on pretest and six on posttest, from twelve participants were changed for this reason.

Participants' shown work for each problem was coded according to the solution methods used. A complete list of the coding standard used is provided in the supplemental materials (Section D). The primary codes of interest are (a) "numerical calculation," signifying presence of an arithmetic expression; (b) "factorial," signifying use of factorials; and (c) "outcome listing," signifying listing of at least two specific outcomes, usually represented by strings of letters. More than one code could be assigned to the same shown work. All shown work was coded independently by two coders, one of whom was the first author. All discrepancies in coding were resolved through discussion.

Results

Pretest performance. Participants' average score on the four pretest problems was .55 (the maximum possible score being 1.0), indicating that the test problems were challenging but not impossible for them. A 2 × 4 × 2 × 2 mixed analysis of variance was performed on the pretest scores, with transfer distance (near or far) as a within-subjects factor and training condition (pure listing, pure formalism, list fading, or formalism-first), pretest–posttest rotation, and training problem sequence as between-subjects factors. There was no effect of training condition, $F(3, 95) = 1.54$, $p = .209$, indicating that participants in the different training conditions did not differ in problem-solving performance before training. There was a significant main effect of transfer distance, such that scores were higher on near (.68) than on far (.41) transfer prob-

Table 1
Assigned and Shown Solution Methods by Training Condition and Training Problem

Condition	Problem 1	Problem 2	Problem 3	Problem 4
Pure formalism	Assigned: calculation Shown: formalism-based			
Pure listing	Assigned: listing Shown: listing-based			
List fading	Assigned: listing Shown: listing-based	Assigned: listing Shown: mixed	Assigned: calculation Shown: mixed	Assigned: calculation Shown: formalism-based
Formalism-first	Assigned: calculation Shown: formalism-based	Assigned: calculation Shown mixed	Assigned: listing Shown: mixed	Assigned: listing Shown: listing-based

lems, $F(1, 95) = 64.06$, $p < .001$. There was also a significant main effect of training problem sequence on pretest score, $F(1, 95) = 8.35$, $p = .005$. Because nothing shown before or during pretest was influenced by training problem sequence, this effect must be attributed to random variation. No other main effects or interactions reached significance. In the analysis of transfer performance below, pretest score is included as a covariate to account for any possible impact of pretest differences on subsequent transfer.

Transfer performance. Average transfer performance on near transfer and far transfer problems (i.e., posttest score minus pretest score) is shown by condition in Figure 3. Note that transfer performance for a given transfer distance constituted an average over two problems and thus could assume five possible values: −1.0, −0.5, 0.0, 0.5, and 1.0. Participants' scores improved on average by .17 from pretest to posttest. One-sample two-tailed *t* tests conducted for each transfer distance and training condition found that transfer performance was significantly higher than 0 for all conditions except formalism-first for near transfer and for all conditions except pure listing for far transfer, using the criterion $\alpha = .05$.

The transfer performance data were entered into a linear mixed model with average pretest score as a covariate, transfer distance as a within-subjects factor, and training condition, pretest–posttest rotation, and training problem sequence as between-subjects factors. The covariate had a significant and large effect on transfer, $F(1, 191.07) = 100.59$, $p < .001$, $\eta_p^2 = .344$,[2] indicating that participants showed more improvement on posttest for problems on which they had done poorly on pretest.[3] The main effect of transfer distance was also significant, indicating greater transfer for near (.21) than for far (.12) transfer problems, despite the higher pretest scores on the former, $F(1, 133.48) = 49.14$, $p < .001$, $\eta_p^2 = .143$. The main effect of training condition was not significant, $F(3, 105.27) = 2.11$, $p = .104$, $\eta_p^2 = .039$, but there was a significant interaction between condition and distance, $F(3, 107.34) = 4.35$, $p = .006$, $\eta_p^2 = .040$. The effects of rotation and sequence were not significant (*p*s > .55.)[4]

To characterize the Condition × Distance interaction, we analyzed the data for near and far transfer problems separately using the above model except with the transfer distance factor removed. These analyses showed a significant effect of training condition for near transfer problems, $F(3, 104.00) = 2.78$, $p = .045$, $\eta_p^2 = .074$, but not for far transfer problems, $F(3, 104.00) = 2.15$, $p = .099$, $\eta_p^2 = .058$. Considering the data for near transfer only, pairwise comparisons between conditions showed that transfer in both the pure formalism and list fading conditions (.30 and .32, respectively) was marginally higher than in the pure listing condition (.19), $p = .055$ for pure formalism and $p = .089$ for list fading, and significantly higher than in the formalism-first condition (.05), $p = .022$ for pure formalism and $p = .039$ for list fading. The pure formalism and list fading conditions did not differ from each other ($p = .813$), nor did the pure listing and formalism-first conditions differ from each other ($p = .711$). The corresponding comparisons for far transfer are not reported because of the absence of a significant main effect of condition at that transfer distance.

Comprehension performance. The effect of training condition on performance on the comprehension questions did not reach significance in an analysis of covariance with condition as a factor and pretest score as a covariate, $F(3, 106) = 1.73$, $p = .164$. A more detailed analysis of these data is presented in the supplemental materials (Section C).

Training compliance. On the basis of their shown work, participants were classified as either having complied or not with instructions regarding which solution method to use for the training problems. First, they were classified as having complied for a given pair of training problems if they received the code corresponding to the designated solution method, either numerical calculation or outcome listing, for at least one problem in the pair. Next, they were classified as having complied overall if they complied for both pairs of problems. Note that these classifications reflect whether participants attempted to use the designated methods, not whether they did so correctly or successfully.

Compliance was maximal in the pure formalism condition (100%), rather low in the formalism-first condition (28%), and intermediate in the pure listing (52%) and list fading (57%) conditions. The difference between conditions was significant, $\chi^2(3) = 31.09$, $p \approx .000$. Pairwise comparisons showed that compliance was significantly higher in the pure formalism condition than in all other conditions, and lower in the formalism-first condition than in all other conditions except the pure listing condition, for which the difference was not significant ($p = .113$). Compliance in the pure listing and list fading conditions did not differ significantly. Noncompliance was primarily driven by some participants' refusal to use outcome listing when instructed to do so, rather than by failure to use numerical calculation. For example, in the formalism-first condition, 100% of participants used numerical calculation on the first pair of training problems, but only 28% of these used listing on the second pair.

To determine whether compliance influenced the effectiveness of the training, we compared transfer performance of complying and noncomplying participants in each training condition. In the list fading condition, transfer was higher among participants who complied (.36) than among those who did not (.08). The reverse was true in the pure listing (.09 vs. .15) and formalism-first (.09 vs. .10)[5] conditions. The linear mixed model described above was run again with compliance added as a factor. However, neither the main effect of compliance nor its interaction with training condition reached significance: $F(1, 74.87) = 1.57$, $p = .214$, for the main effect and $F(2, 75.36) = 2.39$, $p = .099$, for the interaction.

Individual differences. Participants' shown work for the pretest problems was also analyzed to see whether individual participants differed in their usage of formulas and outcome lists before training. Using the coding of their shown work, we classified participants as having used listing if they received the outcome listing code for any pretest problem and as having used factorials if they received the factorials code for any pretest problem. On the basis of these stan-

[2] Here and elsewhere, effect sizes reported for linear mixed models were estimated with partial eta-squared values obtained from general linear models treating transfer distance as a between-subjects factor. Partial eta-squared could not be calculated directly from the linear mixed model treating transfer distance as a within-subjects factor because sums of squares cannot be calculated for that model.

[3] This effect is virtually necessitated by the way in which the data was calculated. For example, for pretest score = 1.0, transfer performance could only range from −1.0 to 0.0, whereas for pretest score = 0.0, transfer performance could range from 0.0 to 1.0.

[4] Interaction terms involving pretest score, rotation, and sequence were not included in the model.

[5] No comparison was possible in the pure formalism condition because all participants in this condition complied. For the same reason, the pure formalism condition was excluded from the subsequent analysis described below.

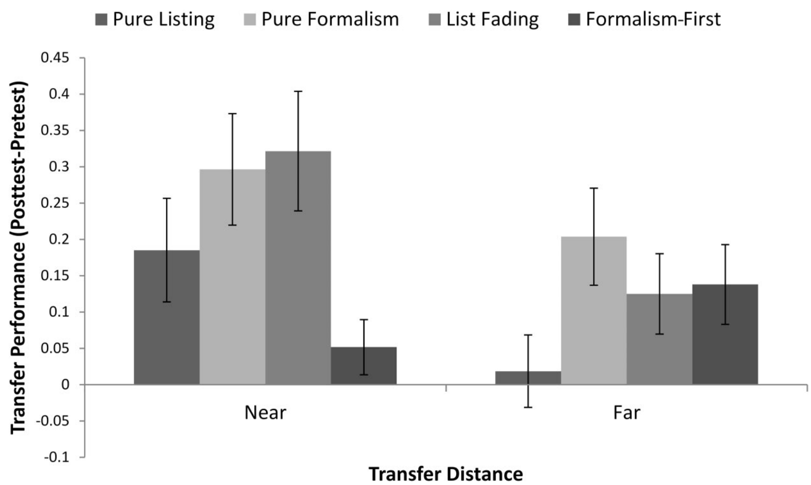

Figure 3. Transfer performance by training condition and transfer distance (Experiment 1). Error bars indicate standard errors.

dards, 68% of participants used factorials and 29% used outcome listing on pretest. The rates of pretest use of factorial and listing did not differ significantly between training conditions: $\chi^2(3) = 1.32$, $p = .725$, for factorials and $\chi^2(3) = 0.85$, $p = .840$, for listing.

Next, rates of training compliance were compared between participants displaying different strategy use during pretest. Descriptive data for this comparison are shown in Figure 4. In the list fading condition, compliance was marginally significantly higher for participants who used listing in pretest than for those who did not, $\chi^2(1) = 3.71$, $p = .054$, and significantly lower for participants who used factorials in pretest than for those who did not, $\chi^2(1) = 4.93$, $p = .026$. Recall that noncompliance with instruction in the list fading condition was driven almost exclusively by failure to use outcome listing when instructed to do so. We may now add that such failure to use outcome listing was more likely among participants with a preexisting preference for numerical calculation. No significant effects of pretest strategy use on compliance were found in any other training condition.

Discussion

Experiment 1 explored the potential benefits to instruction in combinatorics of a concreteness fading/progressive formalization approach, here termed list fading, in which solution methods employing outcome listing are introduced first, followed by solution methods employing numerical calculation. List fading instruction was compared to instruction employing formal representations only (pure formalism condition), grounded representations only (pure listing condition), or formal followed by grounded representations (formalism-first condition). Effectiveness of instruction was evaluated by measuring transfer performance (i.e., improvement from pretest to posttest) on near and far transfer problems.

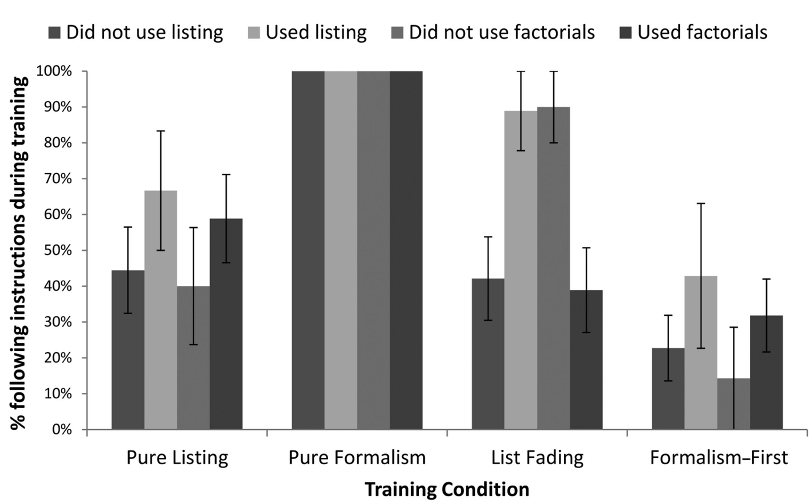

Figure 4. Rates of following instructions during training by training condition and pretest strategy use (Experiment 1). Error bars indicate standard errors. Standard errors could not be calculated for the pure formalism condition due to the absence of variance in the data from that condition.

One clear conclusion to be drawn from our findings is that formal representations are an effective tool for instruction in the combinatorics domain. The pure listing condition, which eschewed formal representations entirely, showed no advantage over any of the training conditions that employed formal representations. Two of these, the pure formalism and list fading conditions, led to marginally better transfer performance than the pure listing condition, albeit only for near transfer. Apparently, for these representations and this domain, learning with formal representations can be at least as effective, and sometimes more effective, than learning without them. These findings are consistent with the importance currently attached to formal representations in mathematics pedagogy.

Second, our findings support the general view that learners benefit more from exposure to both grounded and formal representations if grounded representations are encountered before, rather than after, formal ones (Nathan, 2012). The list fading condition was found to promote transfer better than the formalism-first condition, although this advantage was found only for near transfer and not for far transfer. This finding is consistent with previous research on concreteness fading (Goldstone & Son, 2005), and extends the conclusions of the latter from fading between perceptually rich and perceptually idealized graphical representations to fading between grounded and formal representations. Explanations for, and implications of, this result are explored at greater length in the General Discussion.

On the other hand, no significant difference was found between list fading and pure formalism instruction in their effects on either near or far transfer. This finding is consistent with the possibility that learning combinatorics through a fading approach involving outcome lists and numerical calculation offers no advantage over a purely formal approach involving numerical calculation alone. However, it is also possible that the fading approach is potentially more effective than the purely formal approach, but underperformed relative to its potential due to weaknesses in our implementation. The effectiveness of grounded representations in mathematics instruction is often sensitive to the details of how they are employed (Brown et al., 2009; Mayer, 2002; Mayer & Moreno, 2003; Schnotz, 2005; Schnotz & Bannert, 2003). Below we discuss two possible weaknesses in our implementation of list fading. Corresponding modifications intended to improve the effectiveness of that approach were tested in Experiment 2.

The first possible weakness is that the list fading instruction failed to convey the structural correspondences between the outcome lists and numerical calculations. Theories of instruction involving multiple representations suggest that the benefits of such instruction depend on learners' integrating information from the different representations to form unified concepts (Ainsworth, 2006; Mayer, 2002; Schnotz, 2005), which is unlikely to occur if learners do not understand how the different representations are related to each other. Such understanding may be relatively easy when the different representations are perceptually similar, as indeed was the case in previous studies on concreteness fading (Goldstone & Son, 2005; McNeil & Fyfe, 2012). On the other hand, such understanding may be difficult in cases where one representation is grounded and the other formal, bearing no perceptual similarity to the grounded one, as was the case in the present study. A study on probability learning, employing materials very similar to ours, provides evidence consistent with this account (Berthold & Renkl, 2009). Participants learned about probability principles using either tree diagrams, numerical calculations, or both. Simply displaying both representations together yielded no learning advantage, but superior learning resulted when an instructional aid intended to highlight the correspondences between representations was added. Similarly, the present study's fading approach might be more effective if learners were actively encouraged to notice the correspondences between representations. This prediction was tested in Experiment 2.

A second possible weakness of the list fading condition in Experiment 1 is that participants were not reminded or required actually to use the designated solution methods during instruction. Of course, this point holds true for all the training conditions, not just the list fading condition. However, in fact, compliance was low in the list fading condition compared to the pure formalism condition, and participants who did not comply in the list fading condition showed a nonsignificant trend toward less improvement on posttest. Thus, it is possible that encouraging greater compliance with training instructions would improve performance in the list fading condition. On the other hand, it is unlikely that encouraging greater compliance would improve performance in the pure formalism condition, in which compliance was already maximal. This prediction was also tested in Experiment 2.

It is an interesting finding in its own right that participants were not equally willing to employ the two representation types, that is, outcome lists and numerical calculations, to solve combinatorics problems. Instead, participants were much more willing to use numerical calculation than outcome listing, as evidenced by their higher usage of the former than the latter on pretest and their higher compliance with instructions to use the former than the latter during training. Compliance was particularly low in the formalism-first condition, in which participants were asked to use numerical calculation first and then outcome listing. Possible explanations for this result are deferred to the General Discussion.

Experiment 2

In Experiment 2, modified versions of list fading and pure formalism instruction were tested with a similar methodology to that of Experiment 1. The other two approaches, pure listing and formalism-first, were omitted, due to their relatively poor performance in Experiment 1. Experiment 2 had two goals. First, Experiment 1 found no difference in the effectiveness of list fading and pure formalism instruction. We wished to see whether this negative finding could be replicated, using an extended set of problems that included even farther far transfer problems. Second, the results of Experiment 1 suggested two ways in which list fading instruction might be made more effective: by facilitating learners' integration of outcome listing and numerical calculation and by encouraging higher levels of compliance with instructions during training. We wished to gauge the effectiveness of list fading relative to pure formalism instruction once these changes had been implemented.

Method

Participants. Seventy-three undergraduate students from Indiana University participated in the experiment in partial fulfillment of a course requirement. Of these, 37 were assigned to the list fading condition and 36 to the pure formalism. An additional three participants began the experiment but did not complete it, in one case because the participant did not wish to continue and in two cases due to technical problems.

Materials. As in Experiment 1, two sets of permutations problems were used as pretest and posttest. Each problem set included six problems. The first four problems in each set were the near and far transfer problems from the corresponding problem set in Experiment 1. The last two problems in each set were novel "very far transfer" problems created for Experiment 2. These problems did not involve permutations and thus could be solved neither by direct application nor by simple adaptation, of the methods shown during instruction. However, they could be solved with the fundamental principle of counting, which participants could potentially induce from the instruction regarding permutations. For example, one such problem asked for the number of different ways to assign one of six crops to each of three fields, assuming that each crop could be used more than once. The very far transfer problems from both problem sets are shown in the supplemental materials (Section A).

The problems used as worked examples during training for Experiment 1 were used for the same purpose in Experiment 2. Four types of standard solution were developed: formalism based, listing based, mixed/listing first, and mixed/formula first. These solutions were presented as videos. Screenshots from the videos and Internet links to sample videos for each solution type are provided in the supplemental materials (Section E). As in Experiment 1, each solution presented a solution representation built up incrementally. However, unlike in Experiment 1, the explanations accompanying the solution representations were played as audio tracks rather than as on-screen text. The explanation was always played first, after which the corresponding representation appeared. Separation of representations and explanations into different media channels was intended to increase the effectiveness of instruction by reducing participants' cognitive load in each sensory channel (Mayer & Moreno, 2003).

The two mixed solutions, in which numerical calculations and outcome lists were shown together, were modified in several ways primarily to facilitate integrating across the two representations. First, outcome lists and numerical calculations were presented sequentially instead of simultaneously. For each step of the solutions, the relevant elements of one representation type were added to the representation already built up in previous steps, after which the corresponding elements of the other representation type appeared. The temporal separation of the two representations was intended to facilitate attending to both of them (Mayer & Moreno, 2003). Secondly, between the appearances of each element of the first representation and of the corresponding element in the other representation, an animation appeared showing the former element flying to where the latter element would appear. This animation was intended to highlight the correspondences between the two representations, like the highlighting method used by Berthold and Renkl (2009). Finally, unlike in Experiment 1, the branching tree diagrams employed in the listing-based and mixed solutions did not disappear after initial presentation, so that the final pages of these solutions displayed complete tree diagrams.

Procedure. Experiment 2 followed the same pretest-training, posttest procedure used in Experiment 1. Participants received one test problem set as pretest, then the training problem set and standard solutions, then the other test problem set as posttest. The pretest and posttest were administered in the same way as in Experiment 1, with the exception that on posttest only, when the first very far transfer problem was presented, a message appeared informing participants that the next two problems were different from those seen during training and required different solution methods.

The training problems were presented in the same way as in Experiment 1. Participants were assigned randomly to one of two conditions: pure formalism and list fading. The assigned solution methods and standard solutions shown for each training condition for each of the four problems in the training phase were in all cases the same as for the corresponding conditions of Experiment 1. However, unlike in Experiment 1, when the first and third training problems were presented, participants were shown pop-up messages instructing them to be sure to use the assigned method to solve the problem. The standard solutions were also the same as for the corresponding conditions of Experiment 2, with the exception that in the list fading condition, the mixed/listing-first and mixed/formula-first solutions were shown for training Problems 2 and 3 respectively, rather than the mixed solutions shown in Experiment 1.

After each training problem was completed, participants' shown work was automatically checked to ascertain whether they had, in fact, employed the assigned method, and if not, they were given feedback and asked to correct their work before proceeding. The feedback was sensitive to the specific errors made so that, for example, participants who did not use the designated solution method would be reminded to do so, whereas those who used this method but did not apply it correctly would be reminded of the specific algorithm to be used. Participants who committed either of these errors twice on a given problem were allowed to proceed, considering that repeated failure to fix the error might indicate inability to do so rather than noncompliance with instructions.

For each training problem, after attempting the problem themselves, participants were shown the standard solution as a video. Participants were provided with earphones in order to listen to the audio tracks of the videos, and the experiment administrator verified that they were wearing the earphones while the videos were playing. Once each video finished playing, a button appeared that participants could click to proceed to the next problem. Although a few participants attempted fast-forwarding through the videos in order to proceed more quickly, the great majority appeared to watch them from start to finish.

The comprehension questions asked at the end of Experiment 1 were not included in Experiment 2. However, after completing all test problems, participants were asked a number of background questions including sex, age, and SAT/ACT scores.

Coding. Responses to the test problems were assigned scores of 1 or 0 after correcting for arithmetic errors in the same way as in Experiment 1. Twelve responses, four on pretest and eight on posttest, from nine participants were corrected for arithmetic errors. Additionally, participants were assigned codes according to the methods they used to solve the problems, as indicated in their shown work. A simplified version of the coding scheme used in Experiment 1 was employed, using only the codes for "numerical calculation," "exponent," "factorial," and "outcome listing." The standards for applying these codes were the same as in Experiment 1. All shown work was coded independently by two coders, one of whom was the first author. All discrepancies in coding were resolved through discussion.

Results

Pretest performance. Participants' average score on the six pretest problems was .47. For the problems that were also included in Experiment 1 (i.e., the near and far transfer problems), the average pretest score was .50, very close to the corresponding average in Experiment 1 (.55), suggesting that participants' initial competence in combinatorics was similar between the two experiments. A 3×2 mixed analysis of variance was performed on the pretest scores, with transfer distance (near, far, or very far) as a within-subjects factor and training condition (pure formalism or list fading) as a between-subjects factor.[6] The main effect of transfer distance was significant, $F(2, 130) = 9.52$, $p < .001$. Performance for different transfer distances was compared via pairwise paired *t* tests with a Holm adjustment for multiple comparisons. Performance was higher on near (.61) than on far (.38) and very far (.42) transfer problems (*p*s $< .01$), whereas performance on far and very far transfer problems did not differ ($p =$ .483). Thus, the near transfer problems were initially easier for participants than the far and very far transfer problems, whereas there was no difference in difficulty between the latter two types. Neither the main effect of training condition nor its interaction with transfer distance was significant: $F(1, 65) \approx 0.00$, $p = .992$, for the main effect and $F(2, 130) = 1.46$, $p = .237$, for the interaction. Thus, participants in different training conditions did not differ in their ability to solve the test problems before receiving instruction.

Transfer performance. Average transfer performance for each transfer distance and training condition is shown in Figure 5. Participants' scores improved on average by .18 from pretest to posttest. Improvement on near and far transfer problems only was .27, slightly higher than the corresponding improvement observed in Experiment 1 (.17). One-sample, two-tailed *t* tests conducted for each transfer distance and training condition found that transfer performance was significantly higher than 0 on near and far transfer problems in both training conditions (*p*s $< .05$), as in Experiment 1, but was not significantly different from 0 on very far transfer problems in either training condition (*p*s $> .25$).

The transfer performance data were analyzed with the same model as in Experiment 1: namely, a linear mixed model with transfer distance as a within-subjects factor; training condition, pretest–posttest rotation, and training problem sequence as between-subjects factors; and pretest score as a covariate. As in Experiment 1, the covariate had a large and significant effect on transfer, $F(1, 207.22) = 161.67$, $p < .001$, $\eta_p^2 = .426$, with lower pretest scores leading to higher levels of transfer. There was also a significant main effect of transfer distance, $F(2, 144.80) = 51.22$, $p < .001$, $\eta_p^2 = .279$, such that transfer performance was higher for nearer transfer distances and lower for farther transfer distances. Pairwise contrasts between different transfer distances revealed significant differences between all pairs (*p*s $< .001$). Neither the main effect of training condition nor the interaction of training condition with transfer distance was significant: $F(1, 67.81) =$ .001, $p = .973$, $\eta_p^2 \approx .000$, for the main effect and $F(2, 141.30) =$ .555, $p = .575$, $\eta_p^2 = .004$, for the interaction. Thus, the two training conditions do not appear to have differed in their effects on transfer performance. Finally, the effects of rotation and sequence were not significant (*p*s $> .15$).

Comparison to Experiment 1. To see whether the changes to list fading and pure formalism instruction in Experiment 2 had any effect on transfer, we pooled and compared the data from Experiments 1 and 2. Data from the pure listing and formalism-first conditions in Experiment 1, and from the very far transfer problems in Experiment 2, were excluded, as they had no analogues in the other experiments. Pretest scores did not differ significantly between the two experiments. However, a significant interaction effect of distance and training on pretest score was found, reflecting the fact that pretest scores for near transfer problems were slightly higher in the pure formalism (.63) than in the list fading (.61) condition, whereas the reverse trend held for far transfer problems (pure formalism: .32; list fading: .42), $F(1, 124) = 4.54$, $p = .035$. This interaction must have been due to random variation, as the training conditions did not differ with respect to pretest. However, to account for the possibility that differences in pretest score may have influenced the effects of training on transfer, we included pretest score as a covariate in the subsequent analyses.

Average transfer scores across both experiments for each combination of distance and training condition are shown in Figure 6. These data were analyzed with the linear mixed model previously applied separately to Experiments 1 and 2, with experiment number added as a factor. Neither the main effect of experiment number nor its interactions with training condition, transfer distance, or both were significant: $F(1, 119.96) = .622$, $p = .432$, $\eta_p^2 = .003$, for the main effect; $F(1, 119.97) = .005$, $p = .005$, $\eta_p^2 \approx .000$, for the interaction with training condition; $F(1, 122.72) =$ .259, $p = .612$, $\eta_p^2 = .001$, for the interaction with transfer distance; and $F(1, 122.53) = .693$, $p = .407$, $\eta_p^2 = .002$, for the three-way interaction. Thus, despite the changes made to the implementations of pure formalism and list fading instruction in Experiment 2, their effects on transfer performance did not differ from those of the corresponding instruction in Experiment 1. As in the separate analyses of the two experiments, significant main effects were found for pretest score, $F(1, 201.22) = 206.66$, $p <$ .001, $\eta_p^2 = .497$, and transfer distance, $F(1, 145.99) = 85.59$, $p <$ .001, $\eta_p^2 = .230$. No other significant main effects or interactions were found (*p*s $> .45$).

In particular, the main effect of training condition was not significant, $F(1, 120.08) = .027$, $p = .870$, $\eta_p^2 \approx .000$. Also, the interaction of training condition with transfer distance did not achieve significance, $F(1, 123.21) = .484$, $p = .488$, $\eta_p^2 = .001$, despite the trend for higher transfer performance in the list fading condition for near transfer problems and in the pure formalism condition for far transfer problems. This trend appears to be an artifact caused in part by the above-mentioned interaction of condition and distance in pretest scores. Given its small size ($\eta_p^2 =$ .001), the effect would be unlikely to attain significance even with considerably larger experimental samples. Indeed, a power analysis conducted with resampling found that a sample of 4,050 or larger would be required for 80% power to detect the interaction.

[6] When pretest–posttest rotation and training problem sequence were added as factors in this model, neither had a significant effect nor interacted significantly with any other factor, and the other results reported here remained qualitatively the same.

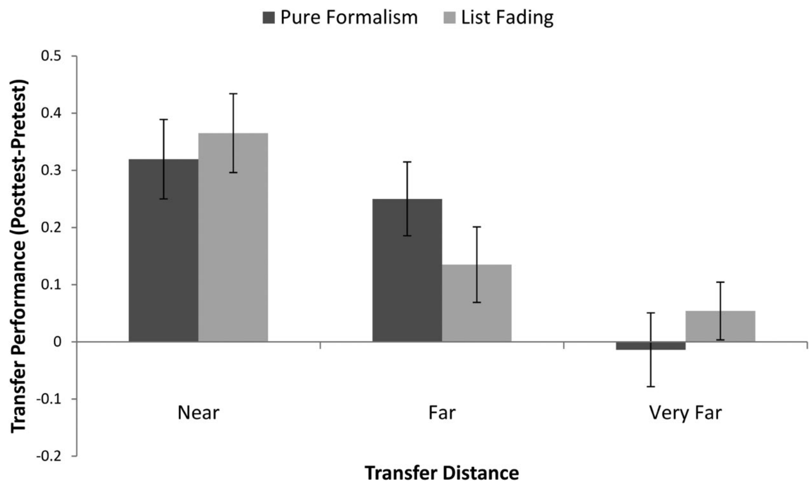

Figure 5. Transfer performance by transfer distance and training condition (Experiment 2). Error bars indicate standard errors.

Training compliance and individual differences. As in Experiment 1, participants were classified as having complied or not with training instructions based on whether they used the designated solution method for at least one of each pair of training problems. By this standard, compliance was 100% in both training conditions. Due to the lack of variation in compliance, it was not possible to analyze effects of pretest strategy use on compliance or effects of compliance on transfer performance, as was done for Experiment 1. Analyses of the effects of pretest strategy on transfer are reported in the supplemental materials (Section F).

Discussion

Experiment 2 tested modified versions of the two best-performing instructional approaches of Experiment 1: pure formalism and list fading. Its goals were, first, to test whether Experiment 1's failure to find a difference between list fading and pure formalism instruction in their effects on transfer performance could be replicated with a larger set of transfer problems, and second, to understand the effects of various modifications made to the instructional conditions on participants' transfer performance.

With respect to the first goal, the negative finding of Experiment 1 was replicated. Despite the efforts made to improve the list fading instruction, Experiment 2 still revealed no difference in transfer performance between the list fading and pure formalism conditions. Similarly, analysis of the combined data from both experiments revealed no differences between the two conditions. Thus, although instruction involving fading from outcome listing to numerical calculation appears to be a viable instructional approach, no evidence was found for its superiority to a simpler approach employing only formal representations throughout. Possible reasons for this result are deferred to the General Discussion.

With respect to the second goal, some of the modifications made in Experiment 2 were expected to benefit both training conditions, whereas others were expected differentially to benefit the list

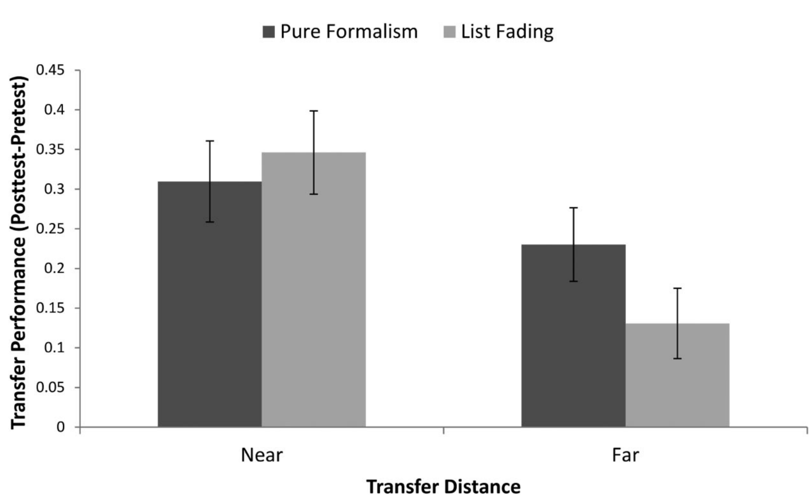

Figure 6. Transfer performance by transfer distance and training condition (pooled data from Experiments 1 and 2). Error bars indicate standard errors.

fading condition, insofar as they were intended to alleviate possible weaknesses of that condition identified in Experiment 1. However, these predictions were not confirmed. No improvement in transfer performance was observed from Experiment 1 to Experiment 2, suggesting that the modifications made in Experiment 2 did not produce any overall benefit for both conditions. Nor was any interaction of condition with experiment number observed, suggesting that the modifications did not differentially benefit either condition. We consider the implications of these null findings for each modification separately.

First, in Experiment 2, use of designated solution methods during training was encouraged by means of interactive feedback scripts. Although this modification greatly increased compliance in the list fading condition, particularly compliance with instructions to use outcome listing, transfer performance in that condition showed no advantage over either the same condition in Experiment 1 or the pure formalism condition of Experiment 2. Apparently, forcing learners to use outcome listing produced no better learning than simply showing outcome listing and leaving learners free to use it or not. Thus, effort devoted to monitoring whether students of combinatorics comply with instructions to use outcome listing, or similar grounded representations, may be effort poorly spent. Further research would be needed to determine whether analogous conclusions apply to grounded instruction in other mathematical domains and for other groups of learners.

Why did increased use of outcome listing in the list fading condition not lead to better transfer performance? One possible explanation is that simply seeing, rather than using, outcome listing may have been enough to learn from it. Day and Goldstone (2011) found evidence consistent with this possibility for a different type of grounded representation. Participants who either directly manipulated or simply viewed manipulations of a physical system simulation subsequently showed equal levels of analogical transfer to a different domain. To the contrary, however, Martin and Schwartz (2005) found that children using concrete objects to learn about fractions showed superior learning outcomes when they physically manipulated the objects, rather than just viewing them. Further research will be needed to understand the specific conditions under which actually manipulating grounded representations can produce learning benefits, relative to merely perceiving them.

A second possible explanation is that those participants who could benefit from using outcome listing were precisely those who would have done so without being forced, whereas those who would not have done so without being forced may have been just the ones who did not stand to benefit much from the experience. For example, some participants may have had greater prior knowledge of formal methods in combinatorics and thus preferred to use formal representations during training. In line with this possibility, noncomplying participants in the list fading condition of Experiment 1 scored higher on pretest than complying participants (.67 vs. .39). This explanation, if correct, would suggest that learners are most likely to benefit from grounded representations if they have not yet learned formal ones. This conclusion dovetails with the poor performance of the formalism-first condition in Experiment 1.

Besides the addition of feedback scripts intended to increase compliance with instructions, another addition expected to benefit the list fading condition was the animations highlighting correspondences between the grounded and formal representations. The fact that no improvement was observed in that condition suggests that highlighting these correspondences did not help participants with respect to posttest performance. Participants may have gained some benefit from the animations that was not reflected in the posttest (e.g., better conceptual understanding). This possibility is consistent with the previously mentioned study of Berthold and Renkl (2009), which found benefits of a similar "relating aid" only on measures of conceptual knowledge, but not on procedural knowledge. On the other hand, the animations of Experiment 2 may not be the best method of highlighting correspondences between representations. Another possible method is to train learners in translating from one representation to the other. This approach has shown promise in other mathematical domains (Kellman, Massey, & Son, 2010; Silva & Kellman, 1999), but has yet to be tested in combinatorics.

A final modification in Experiment 2 was the separation of explanations from solution representations in time and by media modality. This change was intended to reduce cognitive load incurred by attending to both explanations and representations through the same sensory channel at the same time. The absence of an increase in transfer relative to Experiment 1 suggests that this modification did not benefit participants. Cognitive load may not have been an issue to begin with, or alternatively, the change may not have reduced cognitive load enough to produce a significant benefit for learners.

Finally, participants' transfer performance on very far transfer problems deserves mention. Despite performing as well on these problems as on the far transfer problems at pretest, participants showed no significant posttest improvement on the very far transfer problems. Although participants were specifically told at posttest that these problems could not be solved by direct application of the methods learned during training, the need to employ a formal method different from the one learned during training may have been an obstacle nevertheless. Learners seem to naturally limit how general and extended is the impact of their training. One likely factor that constrained participants' effort to extend their training is motivational. Extending formalism or listing procedures beyond their literal forms requires hard cognitive work. In this context, recent proposals (Belenky & Nokes-Malach, 2012; Engle, Lam, Meyer, & Nix, 2012; Perkins & Salomon, 2012; Schwartz et al., 2012) for increasing students' motivation to extend their knowledge may offer keys to achieving very far transfer.

General Discussion

The two experiments reported above tested the effectiveness of instruction based on fading from grounded to formal representations in the domain of combinatorics. Although there is general theoretical and empirical support for the pedagogical effectiveness of combining grounded and formal representations through such fading (Freudenthal, 1991; Goldstone & Son, 2005; McNeil & Fyfe, 2012), the present study represents, to our knowledge, the first controlled experimental test of that approach in the domain of combinatorics, and one of only a few studies involving fading from graphical to purely formal representations. A specific instantiation of that approach, termed list fading, was compared to several other approaches that differed in the types of representations used and their order of presentation, but were controlled as much as possible to be equivalent to list fading in all other respects.

List fading was found to promote transfer better than formalism-first, and marginally better than pure listing (Experiment 1). Its advantage over formalism-first is striking because the two conditions employed the same set of instructional materials and differed only in order of presentation. This result is consistent with that of Goldstone and Son (2005), who also found better learning outcomes following instruction in which a more grounded representation was presented before, rather than after, a less grounded one. The present study extends their conclusions from the case in which both types of representation are depictive, differing mainly in the degree of idealization, to the case in which the less grounded representation is a nondepictive formalism. Goldstone and Son's account may also apply the present study: When more grounded representations are encountered before less grounded ones, the former may serve as scaffolding that assists in understanding the latter, but this function cannot be served when the same representations are encountered in the opposite order.

An alternative explanation, not exclusive with that just mentioned, is that exposure to formal methods may decrease receptivity to grounded ones (Schwartz et al., 2012, 2011). Consistent with this view, in Experiment 1, compliance with instructions to use outcome listing was lower in the formalism-first condition, in which outcome listing was presented after numerical calculation, than in the list fading condition, and within the list fading condition, compliance was lower among participants who had used factorials on pretest than among those who had not. Apparently, exposure to formal representations either before or during instruction may create resistance to subsequently presented grounded representations. Of course, as college students, our participants all had had extensive prior exposure to numerical calculation. Their preference for formal over grounded representations may reflect the formalisms-first approach employed in many mathematics textbooks (Nathan, 2012; Nathan et al., 2002). Younger learners with less prior exposure to such materials might demonstrate greater willingness to employ outcome listing (Figueiras & Cañadas, 2010). However, the relatively poor transfer performance found in the formalism-first condition of Experiment 1 speaks against the appropriateness of similar approaches in the domain of combinatorics, even for college students who might prefer them. The present study thus joins a body of evidence suggesting that if a curriculum involves both formal and grounded representations, it may be preferable to start with the latter (Nathan, 2012).

A related implication of our findings is that simply requiring learners to use grounded representations may not solve the issue of resistance to them. In Experiment 2, participants were essentially forced to use grounded representations but showed no consequent improvement relative to Experiment 1. Clearly, for some participants, the mere experience of creating outcome lists was not beneficial, or was no more beneficial than whatever they would have done if left to their own devices. Understanding why some learners show resistance to grounded representations, whether such learners can benefit from such representations at all, and precisely how this goal may be achieved are all important subjects for future research.

Although list fading instruction led to better learning outcomes than formalism-first instruction, list fading was not more effective than pure formalism instruction, contrary to our predictions. This negative finding was replicated in Experiment 2. Thus, although formalisms-first may be suboptimal in the domain of combinatorics, formalisms-only appears to be acceptable. Other studies of combinatorics and probability learning have consistently found that instruction based on formal representations alone is as effective as or more effective than grounded instruction (Belenky & Nokes, 2009; Kolloffel et al., 2009). The present study extends these findings by demonstrating that instruction based on the specific method of concreteness fading/progressive formalization also shows no advantage over purely formal instruction. It is striking that despite several attempts to identify advantages to employing alternative representational formats in combinatorics instruction, the use of formal representations alone still appears viable.

Another prediction tested by the present study was that list fading would promote conceptual understanding and thus show an especially strong advantage for far transfer. To the contrary, list fading showed some advantage for near but not for far transfer. It is possible that list fading promoted better conceptual understanding, but that this advantage did not translate into better far transfer performance. Improvements in conceptual understanding may not be immediately evident at test, but instead manifest in subsequent learning (Berthold & Renkl, 2009; Schwartz & Martin, 2004). However, although other studies have found advantages of grounded representations for conceptual understanding in combinatorics and other mathematical domains (Berthold & Renkl, 2009; Corter & Zahner, 2007; Figueiras & Cañadas, 2010; Schwartz & Black, 1996; Schwartz & Martin, 2004; Zahner & Corter, 2010), the present findings offer no further evidence of such advantages.

In conclusion, the results of this study support a few general statements regarding mathematics education and also suggest some general questions. Combining grounded and formal representations in mathematics instruction is preferable to focusing on grounded representations alone and appears to be a viable, though not necessarily superior, alternative to focusing on formal representations alone. However, the effectiveness of such a combination may be sensitive to details of implementation. Order matters: Grounded representations may be more useful when encountered before rather than after formal ones. Learners' attitudes and prior knowledge matter: Grounded representations may not be perceived as relevant or useful, especially by learners with prior exposure to formal representations, and simply forcing them to use grounded representations may produce no benefit. How to deal with this potential challenge remains an open question. Finally, not only the representations themselves, be they grounded or formal, but also the connections between them matter. Learners may need active encouragement to notice and understand such connections in order to reap any rewards offered by multiple representations. How best to bring this about is another open question that, it is hoped, will receive more attention in the future.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*(2–3),131–152. doi:10.1016/S0360-1315(99)00029-9

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction, 16*(3), 183–198. doi:10.1016/j.learninstruc.2006.03.001

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637. doi:10.1037/0033-2909.128.4.612

Bassok, M., & Holyoak, K. J. (1989). Interdomain transfer between isomorphic topics in algebra and physics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(1), 153–166. doi:10.1037/0278-7393.15.1.153

Belenky, D. M., & Nokes, T. J. (2009). Examining the role of manipulatives and metacognition on engagement, learning, and transfer. *Journal of Problem Solving, 2*(2), 102–129.

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences, 21*(3), 399–432. doi:10.1080/10508406.2011.651232

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology, 101*(1), 70–87. doi:10.1037/a0013247

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durán, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal, 34*(4), 663–689. doi:10.3102/00028312034004663

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. doi:10.1111/j.1750-8606.2009.00098.x

Corter, J. E., & Zahner, D. C. (2007). Use of external visual representations in probability problem solving. *Statistics Education Research Journal, 6*(1), 22–50.

Day, S. B., & Goldstone, R. L. (2011). Analogical transfer from a simulated physical system. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(3), 551–567. doi:10.1037/a0022333

Engle, R. A., Lam, D. P., Meyer, X. S., & Nix, S. E. (2012). How does expansive framing promote transfer? Several proposed explanations and a research agenda for investigating them. *Educational Psychologist, 47*(3), 215–231. doi:10.1080/00461520.2012.695678

Figueiras, L., & Cañadas, M. C. (2010). Reasoning on transition from manipulative strategies to general procedures in solving counting problems. In M. Jouvert & P. Andrews (Eds.), *Proceedings of the 7th British Congress of Mathematics Education* (Vol. 30, pp. 89–96). Retrieved from http://www.bsrlm.org.uk/IPs/ip30-1/BSRLM-IP-30-1-Full.pdf

Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht, the Netherlands: Kluwer Academic.

Glenberg, A. M., Willford, J., Gibson, B., Goldberg, A., & Zhu, X. (2012). Improving reading to improve math. *Scientific Studies of Reading, 16*(4), 316–340. doi:10.1080/10888438.2011.564245

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46*(4), 414–466. doi: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*(1), 69–110. doi:10.1207/s15327809jls1401_4

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. W. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York, NY: Freeman.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*(4), 684–689. doi:10.1037/0022-0663.91.4.684

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence* (A. Parsons & S. Milgram, Trans.). New York, NY: Basic Books. doi:10.1037/10034-000

Kaminski, J. A. (2006). *The effects of concreteness on learning, transfer, and representation of mathematical concepts* (Unpublished doctoral dissertation). Ohio State University, Columbus.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*(5875), 454–455. doi:10.1126/science.1154659

Kellman, P. J., Massey, C. M., & Son, J. Y. (2010). Perceptual learning modules in mathematics: Enhancing students' pattern recognition, structure extraction, and fluency. *Topics in Cognitive Science, 2*(2), 285–305. doi:10.1111/j.1756-8765.2009.01053.x

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32*(2), 366–397. doi:10.1080/03640210701863933

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences, 13*(2), 129–164. doi:10.1207/s15327809jls1302_1

Kolloffel, B., Eysink, T. H. S., de Jong, T., & Wilhelm, P. (2009). The effects of representational format on learning combinatorics from an interactive computer simulation. *Instructional Science, 37*(6), 503–517. doi:10.1007/s11251-008-9056-7

Landy, D. H., & Goldstone, R. L. (2007). How space guides interpretation of a novel mathematical system. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 431–436). Austin, TX: Cognitive Science Society.

Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 33–40). Hillsdale, NJ: Erlbaum.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science, 29*(4), 587–625. doi:10.1207/s15516709cog0000_15

Mayer, R. E. (2002). Multimedia learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 41, pp. 85–139). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(02)80005-6

Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction. *Journal of Experimental Psychology: Applied, 11*(4), 256–65. doi:10.1037/1076-898X.11.4.256

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52. doi:10.1207/S15326985EP3801_6

McNeil, N. M., & Fyfe, E. R. (2012). "Concreteness fading" promotes transfer of mathematical knowledge. *Learning and Instruction, 22*(6), 440–448. doi:10.1016/j.learninstruc.2012.05.001

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*(2), 171–184. doi:10.1016/j.learninstruc.2008.03.005

Nathan, M. J. (2012). Rethinking formalisms in formal education. *Educational Psychologist, 47*(2), 125–148. doi:10.1080/00461520.2012.667063

Nathan, M. J., & Koedinger, K. R. (2000). An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction, 18*(2), 209–237. doi:10.1207/S1532690XCI1802_03

Nathan, M. J., Long, S., & Alibali, M. W. (2002). The symbol precedence view of mathematical development: A corpus analysis of the rhetorical structure of textbooks. *Discourse Processes, 33*(1), 1–21. doi:10.1207/S15326950DP3301_01

Nuñes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics*. Cambridge, England: Cambridge University Press.

Pape, S. J., & Tchoshanov, M. A. (2001). The role of representation(s) in developing mathematical understanding. *Theory Into Practice, 40*(2), 118–127. doi:10.1207/s15430421tip4002_6

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist, 47*(3), 248–258. doi:10.1080/00461520.2012.693354

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology, 91*(1), 175–189. doi:10.1037/0022-0663.91.1.175

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346–362. doi:10.1037/0022-0663.93.2.346

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(4), 629–639. doi:10.1037/0278-7393.13.4.629

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different Effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 456–468. doi:10.1037/0278-7393.15.3.456

Scheiter, K., Gerjets, P., & Catrambone, R. (2006). Making the abstract concrete: Visualizing mathematical solution procedures. *Computers in Human Behavior, 22*(1), 9–25. doi:10.1016/j.chb.2005.01.009

Schnotz, W. (2005). *An integrated model of text and picture comprehension*. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–60). Cambridge, England: Cambridge University Press.

Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*(2), 141–156. doi:10.1016/S0959-4752(02)00017-8

Schwartz, D. L., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science, 20*(4), 457–497. doi:10.1207/s15516709cog2004_1

Schwartz, D. L., Chase, C. C., & Bransford, J. D. (2012). Resisting overzealous transfer: Coordinating previously successful routines with needs for new learning. *Educational Psychologist, 47*(3), 204–214. doi:10.1080/00461520.2012.696317

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759–775. doi:10.1037/a0025140

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*(2), 129–184. doi:10.1207/s1532690xci2202_1

Shneiderman, B., Mayer, R. E., McKay, D., & Heller, P. (1977). Experimental investigations of the utility of detailed flowcharts in programming. *Communications of the ACM, 20*(6), 373–381. doi:10.1145/359605.359610

Silva, A. B., & Kellman, P. J. (1999). Perceptual learning in mathematics: The algebra–geometry connection. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 683–688). Mahwah, NJ: Erlbaum.

Sloutsky, V. M., Kaminski, J. A., & Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review, 12*(3), 508–513. doi:10.3758/BF03193796

Son, J. Y., & Goldstone, R. L. (2009). Fostering general transfer with specific simulations. *Pragmatics & Cognition, 17*(1), 1–42. doi:10.1075/pc.17.1.01son

Stenning, K. (2002). *Seeing reason: Image and language in learning to think*. Oxford, England: Oxford University Press. doi:10.1093/acprof:oso/9780198507741. 001. 0001

van Reeuwijk, M. (2001). From informal to formal, progressive formalization: An example on "solving systems of equations." In H. Chick, K. Stacey, J. Vincent, & J. Vincent (Eds.), *The future of the teaching and learning of algebra*: *Proceedings of the 12th ICMI Study Conference* (Vol. 2, pp. 613–620). Retrieved from http://repository.unimelb.edu.au/10187/2812

Zahner, D. C., & Corter, J. E. (2010). The process of probability problem solving: Use of external visual representations. *Mathematical Thinking and Learning, 12*(2), 177–204. doi:10.1080/10986061003654240

Received November 7, 2011
Revision received January 30, 2013
Accepted January 30, 2013 ■