Running head: EFFECTS OF RELEVANT CONCRETENESS ON TRANSFER

Individual Differences in the Effect of Relevant Concreteness on Learning and Transfer of a Mathematical Concept

Stephanie A. Siler

Kevin J. Willows

Carnegie Mellon University

Stephanie A. Siler, Psychology Department, Carnegie Mellon University.

Kevin J. Willows, Psychology Department, Carnegie Mellon University.

This research was supported in part by grants from the Institute of Education Sciences (R305A100404 and R305B090023), and in part by a grant from the National Science Foundation (SBE-0836012). Correspondence concerning this article should be addressed to Stephanie Siler, Psychology Department, Carnegie Mellon University, Pittsburgh, PA 15213.

E-mail: siler@cmu.edu

Accepted for publication in the Journal of Learning and Instruction.

Abstract

The concreteness of training materials influences learning and—perhaps more importantly—transfer. Building on prior research finding abstract representations best facilitated transfer to a game task, we conducted a similar study using training figures varying in concreteness but directly assessed transfer to modular arithmetic problems. Training figures: (a) were purely abstract, (b) were abstract but with features relevant to the transfer task, or (c) included additional concrete-relevant features. We hypothesized that concreteness—or number of relevant features—would be positively correlated with learning and transfer—especially among younger and/or lower-ability students. Although there was no overall difference in initial learning, the concrete-relevant and abstract-relevant features independently facilitated near-transfer, where concrete-relevant features supported lower-reasoning students. For far-transfer, eighth-graders benefited from the abstract-relevant features, whereas sixth-graders required additional concrete-relevant features. These findings suggest that concreteness interacts with learner and task characteristics to produce learning and transfer outcomes.

Keywords: abstract problem features; concrete problem features; transfer; mathematics; learning.

Individual Differences in the Effect of Relevant Concreteness on Learning and Transfer of a Mathematical Concept

1. Introduction

A key goal of education is enabling learners to transfer their knowledge beyond the initial learning context. One factor shown to play a role in learning and transfer is the degree of concreteness of instructional materials. The present study explored the effects of concreteness on learning and transfer of students in sixth through eighth grades. Prior work suggests that the effects of concreteness on learning and transfer may be moderated by individual differences, including differences in reasoning and metacognitive skills (e.g., Goldstone & Sakamoto, 2003; Goldstone & Son, 2005; McNeil, Uttal, Jarvin, & Sternberg, 2009; Sloutsky, Kaminski, & Heckler, 2005; Uttal, Scudder, & DeLoache, 1997). Although how the concreteness of instructional material and learner characteristics interact to predict learning and transfer would inform instructional design, there has been little empirical research directly addressing this issue.

The "concreteness" of a representation of an entity is defined in this article as it is widely used in the literature (e.g., Goldstone & Son, 2005; Koedinger, Alibali, & Nathan, 2008; McNeil et al., 2009), as its degree of similarity to an actual physical object. Concreteness may be represented along many dimensions, including visually (e.g., as figures) and textually. The concreteness of a representation is positively correlated with the amount of information it communicates (Kaminski, Sloutsky, & Heckler, 2013) and inversely related to the number of entities it could feasibly represent (Kaminski et al., 2011; 2013). As more concrete representations generally contain more unique features—and individual features themselves can vary in concreteness—they may convey more information that supports learning. However, abstract representations may better facilitate transfer (cf. Bassok & Holyoak, 1989)—especially when surface features differ between the initial and transfer contexts (cf. Bassok, 1990). If these are the primary trade-offs, concrete representations could potentially produce greater understanding of the transfer domain when transfer is facilitated.

Because typical experiments on the effects of concreteness use training conditions varying in concreteness then present transfer tasks in another domain, we first discuss how concreteness can support initial learning. Then we discuss transfer between training and transfer domains. Afterward, we discuss how the nature of concrete features can affect transfer. As the concreteness of various representations of an entity ranges along a continuum, a particular representation can more accurately be described as *more* or *less* concrete than another representation. That is, there is no fixed point delineating “abstract” from “concrete” representations. However, to be concise, in this article we refer to relatively more and less concrete entities as “concrete” and “abstract,” respectively.

1.1. Concreteness and Reasoning

The additional information conveyed by concrete representations may support inferencing or reasoning. For example, in the seminal research of Cheng and Holyoak (e.g., Cheng & Holyoak, 1985; 1989), participants were asked to verify an abstract rule of the form “If A, then B.” People correctly attempt to confirm the rule by selecting cases of “A” and verifying correspondences with “B.” However, they generally fail to check for “rule breakers,” or “not Bs,” to ensure they do not correspond with A. When reframed concretely—in terms of an intuitive “permission” schema (e.g., “If someone is drinking alcohol, they must be at least 21”)—performance dramatically improved. In the abstract form, a complete correct response requires participants to either know or infer that “If A, then B” implies “If not B, then not A.” However, permission schemas support (or even circumvent the need for) reasoning by providing knowledge associated with “checking for rule breakers.” Thus, concrete representations may support initial learning especially among students with weaker reasoning skills.

Concreteness may also affect performance on syllogistic reasoning tasks (Evans, Barston, & Pollard, 1983), such that believable conclusions of valid syllogisms are more likely to be accepted and unbelievable conclusions of invalid syllogisms rejected than belief-neutral conclusions. Concrete scenarios may also impair performance when the logical validity of conclusions contradicts beliefs, an effect termed “belief bias” (e.g., Evans et al., 1983; Klauer et al., 2000). Thus, concrete representations may support or impair performance, depending on whether belief-consistent responses are consistent or conflict with prescriptive performance.

1.2. Comparison Processes and Transfer

In typical studies on effects of concreteness, successful transfer requires that learners draw comparisons such as analogies between training (base) and transfer (target) domains. According to structure-mapping theory (Gentner, 1983, 2010), analogy formation is based on the process of establishing structural alignment between two domains. The resulting alignment produces an explicit set of correspondences between elements across domains. This alignment prompts inferences from the base to the target domain. Inferencing in analogy formation primarily involves mapping *relations* between elements. In literal similarity comparisons, attributes (or properties) of elements in the base domain are also mapped onto corresponding elements of the target domain.

Concrete training representations may support transfer by facilitating alignment between analogous elements of the base and target domains. Concrete elements may have attributes that facilitate understanding of inter-element relationships in the base that may be mapped onto the target domain. Attributes of concrete elements may also support understanding of the transfer domain via inter-domain mapping.

However, several recent studies have found advantages of abstract training domains for transfer (e.g., Goldstone & Sakamoto, 2003; Kaminski, Sloutsky, & Heckler, 2008; McNeil et al., 2009). In Goldstone and Son (2005), college undergraduates trained in one of two versions of a simulation of a complex system that behaved according to simple underlying rules. Although components in both versions were labeled “ants” and “food,” they differed visually. In the “concrete” version, elements were visually represented as ants and food patches (comprised of apples and pears) and in the “abstract” version, as dots and blobs. Although concrete training tended to produce better learning of the training simulation, *abstract* training tended to produce better transfer to an isomorphic simulation. Using the same training and transfer simulations, Goldstone and Sakamoto (2003) also found advantages of the concrete simulation for initial learning and the abstract simulation for transfer. However, *only* students who failed to learn the underlying rules during training benefited more from abstract training. Among students who learned these rules, concrete training continued to produce better transfer.

INSERT FIGURE 1 HERE

Similarly, Kaminski, Sloutsky, and Heckler (2006) found that sixth-grade children learned modular arithmetic rules better when trained with concrete visual figures (e.g., liquid in measuring cups) than with abstract symbols (i.e., shapes), though college undergraduates did not (Kaminski et al., 2008; McNeil & Fyfe, 2012). However, abstract training produced better performance on a “game” transfer task (see Figure 1) among children *and* adults (Kaminski et al., 2006; 2008; McNeil & Fyfe, 2012; De Bock et al., 2011). In De Bock et al. (2011), an extension of Kaminski et al. (2008), the abstract training again produced better transfer for undergraduates. However, participants' explanations suggested they did *not* learn the principles of modular arithmetic from the abstract training. Thus, they could not transfer them. However, *concrete* training better facilitated transfer to a novel task with concrete figures (i.e., sliced pizza; see Figure 1).

1.3. Feature Relevance

One explanation for such mixed findings on the effect of concreteness on transfer (e.g., De Bock et al., 2011; Goldstone & Sakamoto, 2003)—also suggested by reasoning research previously discussed—involves the nature of the features of the concrete representations (Goldstone & Son, 2005). Specifically, the additional features may "provide critical information about likely behaviors and relevant principles" of the training and transfer domains (and thus support reasoning) or may distract learners' attention from critical behaviors and principles. "Relevant features" elicit conceptions relevant to underlying principles common to the training and transfer tasks whereas "irrelevant features" elicit conceptions not relevant in understanding common underlying principles. Only mapping *relevant* features from the training to the transfer domain can facilitate understanding of the transfer domain.

Features of concrete entities may impair transfer by eliciting conceptions irrelevant or even contradictory to those inherent in the transfer task or may simply divert attention from the underlying principles (cf. DeLoache, 1991; McNeil et al., 2009). The concrete representations in the studies previously discussed may have been more likely to elicit irrelevant or contradictory conceptions than the abstracted representations. For example, in the ants/food simulation (Goldstone & Sakamoto, 2003; Goldstone & Son, 2005), the concrete and abstracted versions differed in *visual* characteristics specific to the ants and food. As the abstracted version conveyed all relevant features of the concrete version (e.g., the ants' movements and attraction to food), the additional visual features may not have added any relevant information. In fact, students given the concrete version of the simulation were more likely to express what Goldstone and Sakamoto called “domain-specific interpretations” that did not transfer very effectively (e.g., ants were “getting tired” or “happy with their share of food”). Thus, for some learners, the concrete visual attributes elicited irrelevant conceptions that apparently interfered with learning the underlying rules.

Furthermore, in Kaminski et al. (2006; 2008), sixth-graders learned rules analogous to those of modular arithmetic either with abstract figures (i.e., shapes) or concrete figures (e.g., liquid in measuring cups). In modular arithmetic, numbers “loop” or “reset” upon reaching a particular value, the modulo. For example, in modulo-3 arithmetic, numbers reset to zero upon reaching 3, so, 2 + 1 = 0; 2 + 2 = 1, etc. The conceptions elicited by features of the concrete figures (and by instruction to “report the remainder” in the Concrete condition), relevant in modular arithmetic, may have actually been irrelevant—even distracting—to the transfer task (see Figure 1). If students viewed the task as a literal similarity comparison, they may map concepts including numerical quantity and fullness (analogous to the identity element used by Kaminski et al., 2006; 2008) from the concrete to abstract figures. However, even if students did map, it is unclear how relevant concepts such as quantity could be represented in the transfer context and lead to understanding in the transfer domain. Thus, the concrete features may have been at best irrelevant to this transfer task.

1.4. The Present Study

1.4.1. Aims and materials

The present study built on prior work (De Bock et al., 2011; Kaminski et al., 2006; 2008; McNeil & Fyfe, 2012). However, the transfer tasks of these studies did not directly assess participants' understanding of modular arithmetic principles. In addition, their transfer figures may not have allowed alignment with concrete training features to produce understanding in the transfer domain. Using modular arithmetic as the transfer domain alleviates both issues.

In the present study, we compared abstract, abstract-relevant, and concrete-relevant training figures (Figure 2) to investigate whether this increasing support would produce better learning and transfer outcomes. As they included only relevant features, we used the concrete training figures of Kaminski et al. (2008) in the "Concrete-relevant" condition of the present study. These figures were represented as simple containers with spouts and handles with different proportions of the containers shaded. The shaded proportions were described as "different amounts of liquid." Although the type of container was not specified in training, we call them "measuring cups" in this article (as in prior studies). The type of liquid was not specified because doing so could elicit conceptions irrelevant to the transfer task (e.g., those related to its color, taste, or viscosity). As in Kaminski et al. (2006; 2008), the figures used in the "Abstract" condition of the present study were common shapes.

INSERT FIGURE 2 HERE

To separate the effects of concrete-relevant features from abstract-relevant features (i.e., three-part segmentation), we added an "Abstract-relevant" condition with abstracted versions of the liquid-in-measuring cups figures (see the second column of Figure 2). These figures—analogous to the "faded concrete" representations of McNeil and Fyfe (2012)—were represented as segmented rectangles and described generically as "symbols." They did not include concrete features conveying the idea of measuring cups (handles, spouts). As their segmentation was not contextualized as a "liquid," even this common feature was more abstract in the Abstract-relevant condition. The only difference between the Abstract and Abstract-relevant conditions was the segmentation; both involved abstract shapes and descriptors (i.e., “symbols”). Thus, although the Abstract-relevant figures’ segmentations make them more concrete than the abstract figures, this difference may be minimal.

INSERT TABLE 1 HERE

As outlined in Table 1, figures in both Abstract-relevant and Concrete-relevant conditions conveyed visual information—shaded segmentations—corresponding to numerical quantities most relevant to arithmetic modulo 3. As they likely elicit learners’ conceptions of these quantities, the segmented representations are likely relevant features. Figures in the Abstract condition (shapes) did not obviously convey numerical quantities. The Concrete-relevant figures also included handles and spouts, and the text described its content as a “liquid.” Together, these features may convey notions of pouring or adding liquid. As addition is one common underlying principle, these (concrete) features are potentially relevant. Furthermore, the idea that no more liquid can be added to a full container is analogous to the identity element of modular arithmetic of the present study. Because they are undefined, figures in the Abstract-relevant condition may represent a broader array of entities. Thus, it is less likely that notions of “fullness” or limit-reaching would be elicited in the Abstract-relevant condition (e.g., if students interpret the abstract-relevant figures as stacked solids, there is no physical reason for a limit).

Although the three-part segmentation is analogous to numbers relevant in arithmetic modulo 3, it is not directly relevant in the far-transfer task of the present study, arithmetic modulo 4. However, concrete-relevant features may elicit conceptions related to addition (i.e., “pouring”) and identity (i.e., a “full” measuring cup) that support far-transfer. However, concrete-relevant features may promote context-specific reasoning and hinder far-transfer performance (cf. Goldstone & Son, 2005), especially among younger and/or lower-reasoning students.

1.4.2. Task-related reasoning

We assessed initial rule-learning, near-transfer to modulo 3 problems then far-transfer to modulo 4 problems. Competent performance on assessments may involve drawing conclusions based on one or more given rules (see Appendices A-C). For example, a correct response to the first question on the training test—the result of symbols representing "zero" then "two"—requires applying Rule 1 (thus equating "zero" then "two" with "two" then "zero") then Rule 5, to conclude the result ("two"). The concrete-relevant and abstract-relevant features may support students' understanding that the training task involves quantities and—in concrete-relevant features in particular—evoke knowledge of addition. If so, applying rules to infer results is less necessary (e.g., students could infer that "zero" then "two" produces "two" using simple addition).

1.5. Follow-Up Study: Mapping Task

In a follow-up to the present study, we asked students to map figures and numbers relevant to modulo 3 arithmetic to determine whether the features support cross-task structural alignment. We expected the additional relevant features to support transfer in part by facilitating mappings of analogous elements across tasks. In particular, we expected mapping between figures and numbers to be facilitated in the two relevant-feature conditions, but mapping in the Concrete-relevant condition to be further facilitated by the relevant features that likely elicit conceptions of quantity and addition. We report results of this mapping task after the main results.

1.6. Hypotheses

Comparing the Abstract and Abstract-relevant conditions reveals effects of abstract-relevant features, and comparing the Concrete-relevant and Abstract-relevant conditions reveals the added effects of concrete-relevant features. We hypothesized that abstract-relevant features would facilitate initial learning (H1a), near-transfer (H1b) and far-transfer (H1c). Although we hypothesized general advantages of relevant features (H1), the better alignment of the segmentation with the near-transfer (Mod-3) than far-transfer (Mod-4) task—without additional support from concrete-relevant features—may better support near-transfer than far-transfer. That is, H1b may be more likely than H1c.

We also hypothesized that concrete-relevant features would facilitate initial learning (H2a), near-transfer (H2b) and far-transfer (H2c). We further hypothesized that the benefit of the concrete-relevant features in initial learning (H3a), near-transfer (H3b), and far-transfer (H3c) would be greater among lower-reasoning students. However, we must acknowledge the possibility that concrete-relevant features may promote context-specific reasoning and hinder far-transfer performance, especially among younger and/or lower-reasoning students.

To investigate possible age-related differences in the effect of material concreteness on learning and transfer not captured by the deductive reasoning assessment, students in grades 6-8 participated in this study. Within this age range, there are differences in specific cognitive functioning beyond deductive inferencing, including complex working memory functions such as retrieving information from long-term memory (Gathercole, 1999) and various measures of short-term recall, including digit, word, nonword, and backward digit recall (Gathercole, Pickering, Ambridge, & Wearing, 2004) that may influence the effect of concreteness on task performance. As they may need more support for learning and transfer, we hypothesized the relative advantages of concrete-relevant features would be greatest for the youngest students (H4), for initial learning (H4a), near-transfer (H4b) and far-transfer (H4c).

Finally, we predicted that mapping between figures and numbers would be independently facilitated by concrete-relevant and abstract-relevant features (H12). Hypotheses are summarized in Table 2.

INSERT TABLE 2 HERE

2. Method

2.1. Participants

One hundred thirty students (49 sixth-graders; 40 seventh-graders; 41 eighth-graders) in two sixth-grade (M_{age} = 11.59 years; SD = 0.39), two seventh-grade (M_{age} = 12.71 years; SD = 0.34), and two eighth-grade (M_{age} = 13.81 years; SD = 0.34) classes at a local science and technology magnet school within an urban school district completed all aspects of the procedure. The majority of students in this school (66%) were boys; 47.5% of the student population was African-American, 41% European-American, 6.5% multi-racial, 3% Asian-American, 1.6% Hispanic, and 0.4% American Indian.

2.2. Design

Participants were randomly assigned to one of three conditions within each class. In the Abstract condition, figures were shapes (see Appendix A). In the “Abstract-relevant” condition (Appendix B), figures were segmented shapes (rectangles). In the Concrete-relevant condition, figures were presented as liquid in measuring cups (see Appendix C).

2.3. Materials and Procedure

This study was conducted at an urban science and technology magnet school during students' regularly-scheduled math classes. In each participating class, each student worked individually at a computer for the entire study.

2.3.1. Deductive reasoning test

To account for possible individual differences in the effect of concreteness due to deductive reasoning skills, participants first completed a 9-item (9-point) multiple-choice test of deductive reasoning. Test items were adapted from the Verbal Reasoning section of the InView assessment for Grade 8-9 and ranged in difficulty. Each item consisted of two or more premises and four response choices. A sample item follows:

Ellen swims every Saturday. There is no place to swim in Ellen's town.

A Ellen's family will move.

B Ellen goes out of town to swim.

C Ellen looks forward to Saturday.

D A pool is being built in Ellen's town.

2.3.2. Training task

Students then viewed the condition-specific version of six rules associated with modular arithmetic, presented serially then summarized (Appendices A-C). In contrast to previous studies (Kaminski et al. 2006; 2008; De Bock et al., 2011; McNeil & Fyfe, 2012), there was no mention of a "remainder" as an outcome during training in the Concrete-relevant (or any other) condition. Problem outcomes were called "results" across conditions.

2.3.3. Rules test

All students then answered six questions on the Rules test, which required application of each rule at least once (see Appendix D for example items). The maximum possible score was 11 (some items had multiple correct answers). Again, to control across conditions, instruction that students should report the "remainder" as the outcome (done in previous studies) was omitted. Unlike in previous studies, students were not given feedback on their responses during the Rules test because this may reduce or magnify any initial between-condition differences in rule understanding. These revisions likely made the Rules test of the present study more difficult than in previous studies (De Bock et al., 2011; Kaminski et al. 2006; 2008; McNeil & Fyfe, 2012). Chance performance was 33.3%.

Students could view the "Rule Summary" page anytime during the Rules and transfer tests by clicking on then hovering over a button. When the Rule Summary page was open, students could also see the current test question.

2.3.4. Near-transfer assessment

Students were then introduced to modulo arithmetic with: "The rules you just learned are the same as rules in a particular type of mathematics called "Modulo 3" arithmetic and shown the correspondence between "Rule 3" in their respective condition and "1 + 2 = 0". They were asked to "try to figure out the results for the following 'Mod 3' addition problems using what you just learned…" for ten multiple-choice problems on the Mod-3 assessment (see Appendix E for a sample item). The maximum possible score was 10. As each Mod-3 question had four response choices, chance performance was 25%. Applying (non-modular) addition would result in 20%.

2.3.5. Far-transfer assessment

Students were then introduced to the far-transfer task, Modulo 4 arithmetic, with: "For the final task, can you figure out the solutions to the following 'Mod 4' addition problems? To answer the following questions, you will need to adapt the rules you have learned for Modulo 3 addition to Modulo 4 addition." Students typed their responses to nine Mod-4 addition problems into a text box (see Appendix F for a sample item). The maximum possible score was nine. As students could enter any number for their response, chance performance approached 0%; using (non-modular) arithmetic would result in one correct response (11%). Finally, to assess students' explicit knowledge, they were asked to explain the rules of Mod-3 and Mod-4 arithmetic.

2.4. Data Collection

Students' responses to the computer-presented questions and help elicitations were recorded in log files. The percent of logically correct responses on the deductive reasoning test served as the measure of student reasoning. The percentages of correct responses on the Rules test, Mod-3 assessment, and Mod-4 assessment served as measures of rule-learning, near-transfer, and far-transfer, respectively.

2.5. Explanation Coding

Students' open-ended explanations of modular arithmetic were classified as one of the following: (a) partial/full expression of modular arithmetic (e.g., expressing the idea that the sum cannot exceed a limit, the idea that numbers "loop" around after a limit), (b) vague expression, where the statement is not specific enough to determine whether the student understands any principles of modular arithmetic (e.g., "You look at the comparison in numbers then figure out where the zero is."), (c) matching-only, where the student only mentions aligning training and transfer figures (e.g., "… I tried to match up the numbers with the shapes and figure it out that way and still had no clue."), (d) wrong strategy (e.g., "… I think when zero is in the number sentence the following would also be zero."), (e) a "substance-less" response (e.g., "I don't know what I am doing"; no response). A second coder categorized a randomly-selected 15% of all responses. Inter-rater reliability was good (kappa = .87).

2.6. Analyses

For each assessment, we tested the hypothesized effects of abstract-relevant features (H1), concrete-relevant features (H2), and differential effects of condition by age (H4), in a “baseline” analysis with only condition and grade included in ANOVA. As viewing the Rules Summary page may have affected student performance, we then report any differences by condition in time this help was available. To investigate differential effects of condition due to reasoning (H3), we report analyses with reasoning terms included in statistical models.

3. Results

There were no differences in deductive reasoning performance across conditions, (p = .72), or grades (p = .43). Overall, students scored significantly above chance on all assessments; thus, non-significant results do not appear to be due to floor (or ceiling) effects. Mean percent correct on the deductive reasoning, training, and transfer assessments are reported in Table 3.

INSERT TABLE 3 HERE

3.1. Initial Learning

Inconsistent with H1a and H2a, there was no overall effect of abstract-relevant or concrete-relevant features for initial rule-learning. That is, there was no overall significant effect of condition on rule-learning, $F(2, 125) = 1.25$, $p = .29$, $\eta_p^2 = .02$. There was also no effect of condition on the total time the Rule Summary page was visible during the Rules test, $F(2, 127) = 0.74$, $p = .48$, $\eta_p^2 = .01$ (mean times are reported in Table 4).

INSERT TABLE 4 HERE

Counter to H3a, the condition-by-reasoning interaction was not significant, $F(2, 120) = 1.39$, $p = .25$, $\eta_p^2 = .02$. However, there was a significant grade-by-reasoning interaction, $F(2, 120) = 3.34$, $p = .04$, $\eta_p^2 = .05$. Among sixth- and seventh-graders, there was no effect of condition ($p = .66$; $p = .28$, respectively). However, for eighth-graders, there was a marginally significant condition-by-reasoning interaction, $F(2, 35) = 2.68$, $p = .08$, $\eta_p^2 = .13$ (see Figure 3). For higher-reasoning eighth-graders (i.e., those scoring above the mean of 59% on the deductive reasoning test), there was a significant effect of condition, $F(2, 19) = 3.91$, $p = .04$, $\eta_p^2 = .29$. Students in the Concrete-relevant condition out-performed students in the Abstract-relevant (p = .04) and Abstract (p = .01) conditions (M = 79.55, SD = 30.04; M = 53.54, SD = 20.55; M = 46.46, SD = 13.21, respectively). However, the Abstract-relevant and Abstract conditions did not differ (p = .46). Among *lower*-reasoning eighth-graders, there was no effect of condition, $F(2, 16) = 0.21$, $p = .81$, $\eta_p^2 = .03$. Thus—in partial support of H2a—concrete-relevant features produced better rule-learning only among higher-reasoning eighth-graders.

INSERT FIGURE 3 HERE

To investigate this result, we report correlates of rule-learning. In Grade 6, neither reasoning nor condition was related to rule-learning. In Grade 7, only reasoning was related to rule-learning, $r(38) = +.39$, $p = .01$, $\beta = +.34$. For Grade 8, reasoning was significantly correlated with rule-learning in the Concrete-relevant condition, $r(12) = +.64$, $p = .02$, $\beta = +.80$, but not in the Abstract-relevant (p = .94), or Abstract (p = .45) conditions. Thus, only those eighth-graders in the Concrete-relevant condition successfully applied their reasoning skills on the training task.

3.2. Near-Transfer Outcomes

We also hypothesized independent advantages of abstract-relevant features (H1b) and concrete-relevant features (H2b) for near-transfer performance. There was a significant main effect of condition, $F(2, 127) = 12.21$, $p < .001$, $\eta_p^2 = .16$. Consistent with H1b, students in the Abstract-relevant condition scored significantly higher than students in the Abstract condition (p = .002). Consistent with H2b, students in the Concrete-relevant condition scored marginally higher than those in the Abstract-relevant condition (p = .09). These same patterns were found when reasoning was statistically controlled.

There was also a significant difference in time the Rule Summary page was visible during the Mod-3 test, $F(2, 127) = 5.43$, $p = .005$, $\eta_p^2 = .08$. This time was significantly less in the Concrete-relevant than Abstract (p = .002) or Abstract-relevant (p = .01) conditions. The abstract conditions did not differ (p = .60). Thus, concrete-relevant training materials not only improved performance on the near-transfer assessment, but also appeared to make transfer easier. When controlling for time the Rule Summary page was visible during the Mod-3 test, consistent with H2b, near-transfer performance was significantly better in the Concrete-relevant than Abstract-relevant condition (p = .03). And, again consistent with H1b, performance was better in the Abstract-relevant than Abstract condition ($p < .001$).

Although we hypothesized that advantages of the concrete-relevant features would be greater for lower-reasoning students (H3) and younger students (H4), neither the condition-by-reasoning nor condition-by-grade interactions were significant. As rules knowledge may have influenced near-transfer outcomes (and possibly obscured interactions with reasoning or age), we included Rules score in the model. There was a significant condition-by-reasoning interaction (Figure 4a), $F(2, 118) = 3.79$, $p = .03$, $\eta_p^2 = .06$, where—consistent with H3b—the benefit of the concrete-relevant features was greater among lower-reasoning students. Among lower reasoners, there was a significant effect of condition, $F(2, 67) = 12.49$, $p < .001$, $\eta_p^2 = .27$, where Mod-3 scores were higher in the Concrete-relevant than Abstract-relevant condition (p = .002). The Abstract-relevant and Abstract conditions did not differ (p = .16). Thus, only the concrete-relevant features supported lower reasoners' near-transfer performance. Among *higher* reasoners, there was a significant effect of condition, $F(2, 55) = 4.62$, $p = .01$, $\eta_p^2 = .14$. The

Concrete-relevant and Abstract-relevant conditions did not differ (p = .89). However, both the Concrete-relevant and Abstract-relevant conditions out-performed the Abstract condition (p = .01 for both). Thus, lower-reasoning students benefited from the concrete-relevant features, but higher-reasoning students benefited from the abstract-relevant features.

INSERT FIGURE 4 HERE

There was also an unexpected condition-by-rule-learning interaction for near-transfer performance, $F(2, 118) = 3.11$, $p = .048$, $\eta_p^2 = .05$ (Figure 4b). The independent benefits of the abstract-relevant and concrete-relevant features increased with increasing Rule test performance.

3.3. Far-Transfer Outcomes

We also hypothesized independent advantages of abstract-relevant features (H1c) and concrete-relevant features (H2c) for far-transfer. As with near-transfer, there was a significant main effect of condition, $F(2, 127) = 3.29$, $p = .04$, $\eta_p^2 = .05$. However, the only significant pair-wise result was an advantage of the Concrete-relevant condition over the Abstract condition (p = .01). There was no effect of condition on time the Rule Summary page was visible during the Mod-4 test, $F(2, 125) = 1.89$, $p = .16$, $\eta_p^2 = .03$. Thus, neither H1c nor H2c were supported in this baseline analysis.

INSERT FIGURE 5 HERE

Inconsistent with H3c and H4c, condition did not interact with either reasoning or grade. However, there was a significant grade-by-reasoning interaction, $F(2, 122) = 3.19$, $p = .045$, η_p^2 = .05. Among sixth-graders, there was a significant effect of condition, $F(2, 45) = 4.13$, $p = .02$, η_p^2 = .16 (Figure 5). Consistent with H2c, far-transfer scores of sixth-graders were significantly higher in the Concrete-relevant than Abstract-relevant condition (p = .03), but—inconsistent with H1c—the abstract conditions did not differ (p = .82).

Among seventh-graders, there was no effect of condition, $F(2, 36) = 0.13$, $p = .88$, $\eta_p^2 = .007$. Among eighth-graders, there was a marginal effect of condition, $F(2, 37) = 3.14$, $p = .06$, $\eta_p^2 = .15$. Consistent with H1c, eighth-graders in the Abstract-relevant condition scored marginally higher than those in the Abstract condition ($p = .09$). Mod-4 scores in the Concrete-relevant condition were significantly higher than in the Abstract condition ($p = .02$). However, inconsistent with H2c, the Concrete-relevant and Abstract-relevant conditions did not differ ($p = .56$). Although condition did not interact with reasoning or age, the pattern of Mod-4 results is generally consistent with H4c. Concrete-relevant features supported the youngest students (sixth-graders) and abstract-relevant features supported the oldest students (eighth-graders).

3.4. Summary of Key Findings

In summary (Table 5), concrete-relevant features supported initial rule-learning only among the oldest, higher-reasoning students. However, the concrete-relevant and abstract-relevant features independently facilitated near-transfer. In addition, the concrete-relevant features in particular facilitated near-transfer among lower reasoners. For far-transfer, the effect of concreteness differed with age: sixth-graders benefited only from concrete-relevant features and eighth-graders from abstract-relevant features. Results vis-à-vis each hypothesis are summarized in Table 6.

INSERT TABLES 5 AND 6 HERE

3.5. Articulation of Modular Arithmetic

Finally, there was an overall significant difference in response frequency between conditions, $\chi^2(8, 130) = 22.41$, $p = .004$, which was due to different rates of partial/full expressions of modular arithmetic and matching-only responses. In the Concrete-relevant condition, 38% of students expressed at least a partially-correct understanding of modular arithmetic, compared to 24% of students in the Abstract-relevant condition and 7% of students in the Abstract condition, $\chi^2(2, 130) = 12.71$, $p = .002$. Conversely, 17% of students in the Abstract condition expressed only that they tried matching training figures with numbers compared to 5% in the Abstract-relevant condition and 0% in the Concrete-relevant condition, $\chi^2(2, 130) = 10.10$, $p = .006$. This measure of explicit understanding generally parallels the results of the near- and far-transfer assessments: performance was better in the Concrete-relevant than Abstract-relevant condition and better in the Abstract-relevant than Abstract condition.

3.6. Follow-Up Study: Mapping task

To investigate the hypothesis that mappings are facilitated by relevant features (H12), the following year, 50 sixth-graders from the same school (who had not participated in the present study) completed one of four versions of a hand-out asking them to “Pair up each figure on the left with the number/figure on the right that you think is most like it (for whatever reason)…” (see Figure 7).

INSERT FIGURE 7 HERE

In each hand-out version, students were first asked to pair each abstract figure given in the present study (Figure 2) with numbers 0-3. The next pairing task was for either concrete-relevant or abstract-relevant figures (depending on version), and the final pairing task involved either the concrete or abstract figures and figures from the game transfer task of previous studies (e.g., De Bock et al., 2011; Kaminski et al., 2006; 2008; McNeil & Fyfe, 2012) (see Figure 1).

3.6.1. Mapping analogous figures

Almost all of the students (22 of 23, or 96%) correctly aligned all four concrete-relevant figures with their corresponding numbers; 20 of 26 students (77%) correctly aligned all abstract-relevant figures with their corresponding numbers. This difference was marginally significant,

$\chi^2(1, 50) = 3.50$, $p = .06$. None of the 50 students correctly aligned all abstract figures with their analogous numbers (two would be expected by chance); however, students mapped 40% of all abstract-symbolic training figures with their analogous transfer elements (significantly above chance, at 25%). Consistent with H12, these results suggest the additional concrete-relevant features and abstract-relevant features independently increased the salience of the deep feature, quantity.

3.6.2. Conceptual associations

Concrete-relevant figures appeared to elicit conceptions of physical size in particular. Among students who had previously mapped concrete-relevant figures onto numbers, 7 of 12 (58%) mapped the one-unit measuring cup with the ladybug, the two-unit cup with the ring, and the three-unit cup with the vase on the final task. In informal interviews with three students who response this way, two said they had mapped them in terms of object size (i.e., the ladybug was smallest and the vase largest) and one student could not explain her rationale. In contrast, only two of 11 students (18%) who had previously mapped the abstract-relevant figures with numbers mapped in this manner. This difference was marginally significant *p*(Fisher's) = .06.

3.6.3. Mapping analogous figures of prior studies

Finally, among students who paired the abstract shapes given as training figures in prior studies with figures of the transfer task used (i.e., ladybug, ring, and vase), 15 of 25 (60%) aligned all three analogous figures. In contrast, only 5 of 25 (20%) students who paired the concrete training figures with the game transfer task figures paired all analogous figures (similar to chance, at 17%). This difference was significant, $\chi^2 (1, 50) = 8.33$, $p = .004$. These results suggest that both perceptual similarities between abstract training and transfer figures and lack of alignment—either structural or surface/perceptual—between concrete training and transfer figures may have contributed to the benefits of the abstract training previously found.

4. Discussion

The present study investigated the effects of instructional concreteness on initial learning and transfer and how these effects were moderated by age and reasoning skill. Specifically, we investigated the effects of conceptual associations conveyed by concrete- and abstract-relevant features on transfer to a topic in mathematics. Previously, in this line of research (e.g., De Bock et al., 2011; Kaminski et al., 2006; 2008; McNeil & Fyfe, 2012), abstract symbolic training figures caused better transfer than concrete figures. However, this may have been due to the irrelevance of features of concrete training figures to the particular transfer task. This interpretation is suggested by results of the present study, in which abstract figures produced the poorest transfer to modular arithmetic. Consistent with prior findings (De Bock et al., 2011), based on the open-ended explanations given in the present study, abstract training rarely led to explicit understanding of modular addition.

Because of the increasing potential support provided by relevant features, we generally predicted that students would benefit least from abstract and most from concrete-relevant training figures. Specifically, we hypothesized independent benefits of abstract-relevant features (H1) and concrete-relevant features (H2). Based on prior findings that concrete contexts support reasoning (Cheng & Holyoak, 1985; 1989), we also hypothesized that the additional benefit of concrete-relevant features would be greater for lower-reasoning (H3) and younger (H4) students.

4.1. Effect of Relevant Features in Training Task

Consistent with prior findings (Kaminski et al., 2008; McNeil & Fyfe, 2012), but inconsistent with H1a and H2a, there was no general effect of condition for initial rule-learning.

Thus, neither concrete-relevant nor abstract-relevant features supported learning of the training task. However, when reasoning was included in analyses, counter to our hypotheses that concrete-relevant features would support younger students (H3a) and lower-reasoning students (H4a), only the oldest (eighth-grade) higher-reasoning students benefited from the concrete-relevant features.

4.2. Effect of Abstract-Relevant Features on Transfer

We expected that figure segmentation would support transfer by eliciting numbers relevant to mod-3 arithmetic. Thus, we hypothesized that abstract-relevant features would support near-transfer (H1b). This hypothesis was supported. Results of the mapping task suggest that students were able to recognize the deep feature, quantity, in the abstract-relevant training figures—at least when figures were presented with analogous numbers.

Additionally, that students in the Abstract-relevant condition showed larger near-transfer gains for corresponding gains in reasoning skill than students in the Abstract condition (see Figure 4a) suggests that the abstract-relevant features supported students' reasoning on the near-transfer task. However, for far-transfer performance, only the oldest students (eighth-graders) benefited from the abstract-relevant features, per se. Thus, H1c was supported only among the oldest students. As discussed next, the youngest students required additional support on the far-transfer task.

4.3. Effect of Concrete-Relevant Features on Transfer

We expected concrete-relevant features would further support transfer by eliciting conceptions that support understanding in the transfer domain (including notions of combining amounts of liquid via pouring and the significance of "full" measuring cups). Thus, we predicted that students in the Concrete-relevant condition would be most successful transferring their knowledge to modular arithmetic.

Consistent with H2b, concrete-relevant features supported near-transfer and—consistent with H3b—in particular among lower-reasoning students. Thus, as with previous findings that conceptual knowledge can support reasoning (e.g., Cheng & Holyoak, 1985; 1989; Cheng et al., 1986), the additional features of the concrete figures supported transfer. Results from the mapping task indicated that concrete-relevant features' facilitation of mapping onto corresponding numbers and elicitations of conceptions of quantity (and size) are both plausible mechanisms underlying the advantage of the concrete-relevant features. Evoking notions of size may have even facilitated cross-task quantity mappings.

We also expected concrete-relevant features (in particular, the idea of measuring cups being "full," analogous to the identity concept of modular arithmetic) to support far-transfer. Partially consistent with H2c, the additional concrete-relevant features facilitated far-transfer only for the youngest students (sixth-graders). Among the oldest students, although far-transfer was supported by abstract-relevant features, there was no additional benefit of concrete-relevant features. This pattern of results suggests that concrete-relevant features supported far-transfer only among the youngest students (H4c). Although there was a possibility that concrete-relevant features would promote context-specific reasoning and hinder far-transfer performance, especially among younger and/or lower-reasoning students, there was no evidence this occurred.

In summary, only the oldest, higher-reasoning students benefited from the concrete-relevant features during training. Lower-reasoners benefited from the concrete-relevant features on the near-transfer task, and the youngest students benefited from the concrete-relevant features on the far-transfer task. These results suggest that the benefit of concreteness for initial learning and transfer depends on deductive reasoning skills and age, as well as task characteristics (cf. Koedinger & Alibali, 2008).

4.4. Practical Implications

In spite of this apparent aptitude-by-treatment-by-task interaction, there was no evidence of any detrimental effects of either the concrete-relevant or abstract-relevant features on initial learning or transfer performance. That is, increasing the number of relevant features did not appear to impair any outcomes (though it did not always have a positive effect). And, with the exception of rule-learning, the benefits of concrete-relevant features were greater among lower-reasoning and younger students. These results suggest using concrete materials for younger and/or lower-ability students—provided their salient features are relevant to and support students' understanding within the learning task. Younger and/or lower-ability students may need the additional support provided by appropriate concrete features to most effectively (and efficiently) learn more difficult concepts. However, even older and/or higher-reasoning students may benefit from additional concrete-relevant features in some circumstances (addressed next).

4.5. Open Issues

Many questions remain open, including the reason for the apparently conflicting results in which older/higher-reasoning students benefited more from the concrete features during initial learning whereas the younger/lower-reasoning students benefited more from the concrete features during the transfer tasks. One possibility is that stronger cues of the additive nature of the task were available on the modular arithmetic transfer tasks than on the training task. Nothing regarding "adding," (either terminology or related symbols) was referenced during training, whereas the near-transfer and far-transfer tasks were explicitly described as "[modular] addition" and included familiar symbols of addition ("+" and "="). Students may not have realized the significance of the features until they saw the modular addition problems. Thus, more inferencing was likely required for students to realize that the “relevant” features were indeed relevant to the training task than to realize their relevance in the transfer tasks. Consequently, only the oldest, highest-reasoning students in the most supportive condition were able to realize and make use of the relevant features. Consistent with this interpretation, performance in the relevant-feature conditions was significantly higher on the Mod-3 than Rules test.

However, it is less clear why only the youngest students benefited from concrete-relevant features on far-transfer. As discussed previously, within this age range, there are differences in specific cognitive functioning beyond deductive inferencing, including complex working memory functions such as retrieving information from long-term memory (Gathercole, 1999) and various measures of short-term recall (Gathercole, Pickering, Ambridge, & Wearing, 2004) that may influence the effect of concreteness on task performance. Further research addressing the cognitive processes involved in near- and far-transfer performance on these tasks may shed light on this issue.

Finally, although the assumed benefits of relevant concrete features are from facilitating mapping of analogous elements across tasks and eliciting knowledge that supports understanding of modular arithmetic, neither mechanism was directly assessed in the present study. Although the results of the mapping task suggested that both mechanisms may have played a role, the extent to which each mechanism (and possibly others) played a role is unclear. Future research investigating mechanisms underlying the effect of concreteness on learning and transfer processes—perhaps by incorporating think-aloud protocols—may provide more insight and ultimately enable more precise tailoring of instruction to individual students along the concrete-abstract continuum.

4.6. Study Limitations

There are several limitations to the present study that need to be addressed in future research. First, because student performance on the transfer tasks was assessed immediately after training, the effect of concreteness on longer-term retention (of either the training or transfer tasks) was not assessed in the present study. It is possible that the current trends would differ when longer-term retention is assessed (as found in Johnson-Gentile, Clements, & Battista, 1994; McNeil & Fyfe, 2012). Thus, this is an important issue that must be addressed in future research and before educational recommendations can be made with confidence. Secondly, context such as the type of class may influence students' interpretations of problems and problem-solving performance, (Dewolf, Van Dooren, & Verschaffel, 2011). If students had worked in a context other than their mathematics class (e.g., during their regular science classes), their interpretations of the training materials and subsequent transfer outcomes may have differed.

In addition, although there was a "concrete" condition in this study, its figures were minimally concrete—only details deemed relevant to the transfer task were represented in these figures. Thus, this study did not compare a wide range of levels on the "concreteness continuum." Although previous findings indicate that the use of "perceptually rich" concrete objects may have negative impacts on learning and transfer if these features are not aligned with the instructional focus (e.g., DeLoache, 1991; 1995; Harp & Mayer, 1998; McNeil et al., 2009), the effect of using perceptually-richer figures was not assessed in the present study. Using perceptually-richer features may have produced better outcomes than the minimally-concrete figures of the present study or—as suggested by the current findings—the effect may depend on learner characteristics such as age or reasoning skills.

We must also point out the somewhat idiosyncratic interpretation of modular arithmetic used in this study. In addition modulo 3, 0 is the identity. However, in the present study, Rule 2 indicated 3 as the identity (consistent with this line of research De Bock et al., 2011; Kaminski et al., 2006; 2008; McNeil & Fyfe, 2012). The additional variable of 0 given in the Rules presented to students in the present study resulted in an under-defined system, where students must infer some results (e.g., they must infer that 2 + 2 = 1, perhaps by applying Rule 3 (1 + 2 = 0) and adding 1 to each side of the equation). The resulting increased difficulty may account to some extent for the benefits of relevant features found here.

References

Bassok, M. (1990). Transfer of domain-specific problem-solving procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 522-533. doi: 10.1037/0278-7393.16.3.522.

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.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391-416. http://dx.doi.org/10.1016/0010-0285(85)90014-3.

Cheng, P. W., & Holyoak, K. J. (1989). On the natural selection of reasoning theories. *Cognition, 33,* 285-313. doi: 10.1016/0010-0277(89)90031-0.

Cheng P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18*, 293-328. http://dx.doi.org/10.1016/0010-0285(86)90002-2.

De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., & Verschaffel, L. (2011). Abstract or concrete examples in learning mathematics? A replication and elaboration of Kaminski, Sloutsky, and Heckler's study. *Journal for Research in Mathematics Education, 42*(2), 109-126. http://www.jstor.org/stable/10.5951/jresematheduc.42.2.0109.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development, 62,* 736–752. doi: 10.1111/j.1467-8624.1991.tb01566.x.

DeLoache, J. S. (1995). Early understanding and use of symbols: The model model. *Current Directions in Psychological Science, 4*, 109–113. doi: 10.1111/1467-8721.ep10772408.

Dewolf, T., Van Dooren, W., & Verschaffel, L. (2011). Upper elementary school children's understanding and solution of a quantitative problem inside and outside the mathematics class. *Learning and Instruction, 21*(6), 770-780. http://dx.doi.org/10.1016/j.learninstruc.2011.05.003.

Evans, J., Barston, J & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition, 11*, 295–306. doi: 10.3758/BF03196976.

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences, 3*(11), 410-419. http://dx.doi.org/10.1016/S1364-6613(99)01388-1

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177-190. doi: 10.1037/0012-1649.40.2.177.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752-775. doi: 10.1111/j.1551-6709.2010.01114.x.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170. doi: 10.1207/s15516709cog0702_3.

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

Goldstone, R. & Son, J. (2005). The transfer of scientific principles using concrete and idealized simulations. *The Journal of the Learning Sciences, 14*(1), 69-110. doi: 10.1207/s15327809jls1401_4.

Harp, S. & Mayer, R. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*(3), 414-434. doi: 10.1037/0022-0663.90.3.414.

Johnson-Gentile, K., Clements, D. & Battista, M. (1994). The effects of computer and noncomputer environments on students' conceptualizations of geometric motions. *Journal of Educational Computing Research, 11*, 121-140. doi: 10.2190/49EE-8PXL-YY8C-A923.

Kaminski, J. A., Sloutsky, V. M. & Heckler, A. F. (2006). Do children need concrete instantiations to learn an abstract concept? *Proceedings of the XXVIII Annual Conference of the Cognitive Science Society*, 1167-1172. Mahwah, NJ: Erlbaum.

Kaminski, J. A., Sloutsky, V. M. & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 25*, 454-455. doi: 10.1126/science.1154659.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2013). The cost of concreteness: The effect of nonessential information on analogical transfer. *Journal of Experimental Psychology: Applied*, *19*(1), 14. doi: 10.1037/a0031931.

Klauer, K.C., Musch, J., & Naumer B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107*(4), 852–884. doi: 10.1037/0033-295X.107.4.852.

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

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

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

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

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

Table 1

Features expected to support transfer by condition

Condition	Modular arithmetic principle		
	Numbers 0 – 3 (Mod-3)	Addition	Identity
Abstract	(n/a)	(n/a)	(n/a)
Abstract-relevant	Segmentation and shading	(n/a)	(n/a)
Concrete-relevant	Segmentation and shading[a]	Liquid→Pouring	Liquid→“Full” measuring cup

[a] Segmentation and shading is contextualized as “amounts of liquid.”

Table 2

Summary of hypotheses

Hypothesis 1 (H1)	Abstract-relevant > Abstract
Hypothesis 2 (H2)	Concrete-relevant > Abstract-relevant
Hypothesis 12 (H12)	Mapping facilitated most by concrete-relevant figures, then abstract-relevant figures, and least by abstract figures.
Hypothesis 3 (H3)	Lower reasoners will benefit more in Concrete-relevant vis-à-vis Abstract-relevant condition than higher reasoners.
Hypothesis 4 (H4)	Younger students will benefit more in the Concrete-relevant vis-à-vis Abstract-relevant condition than older students.
For H1, H2, H3, H4:	Hxa—rule-learning; Hxb—near-transfer; Hxc—far-transfer

Table 3

Test performance summary: Mean percent correct (standard deviation) by condition and assessment

	Assessment			
Condition	Reasoning	Rules	Mod-3	Mod-4
Abstract	55.80 (24.84)	45.06 (16.92)	39.35 (13.06)[b,c]	20.77 (15.29)[d]
Abstract-relevant	59.43 (18.13)	47.99 (18.11)	52.56 (20.48)[a,b]	25.32 (17.20)
Concrete-relevant	58.27 (21.34)	52.11 (24.06)	60.00 (24.80)[a,c]	31.71 (26.12)[d]

Scores differ at: [a] $p < .10$. [b] $p < .005$. [c] $p < .001$. [d] $p = .01$.

Table 4

Mean duration (seconds) the rule summary page was visible (and standard deviation) by condition and assessment

	Assessment		
Condition	Rules	Mod-3	Mod-4
Abstract	45.05 (46.99)	64.01 (76.93) [a]	40.92 (63.85)
Abstract-relevant	43.65 (50.85)	57.29 (62.68) [b]	24.49 (32.01)
Concrete-relevant	33.33 (47.18)	23.72 (27.72) [a, b]	22.06 (45.57)

Scores differ at: [a] $p < .005$. [b] $p = .01$.

Table 5

Summary of key outcomes by assessment, grade, and analysis

Assessment	Grade	Baseline analysis	Reasoning in model
Rules	6	(n.s.)	(n.s.)
	7		(n.s.)
	8		Low-reasoning: (n.s.) Higher-reasoning: C-r > A-r, A-s
Mod-3	6-8	C-r > A-r > A-s	C-r > A-r > A-s (H1a & H1b) Condition x reasoning (H2)[a] Low-reasoning: C-r > A-r; A-s Higher-reasoning: C-r, A-r > A-s
Mod-4	6	C-r > A-s	(H3) C-r > A-r, A-s (H1a)
	7		(H3) (n.s.)
	8		(H3) C-r, A-r > A-s (H1b)

[a] This interaction was significant when rules test performance was included in the model.

Table 6

Summary of results per hypothesis

H1: Benefit of abstract-relevant features: Abstract-relevant > Abstract.

- H1a: Rule learning: No.
- H1b: Mod-3 near-transfer: Yes.
- H1c: Mod-4 far-transfer: Only among eighth-graders.

H2: Benefit of concrete-relevant features: Concrete-relevant > Abstract-relevant.

- H2a: Rule learning: Only among higher-reasoning eighth-graders.
- H2b: Mod-3 near-transfer: Yes.
- H2c: Mod-4 far-transfer: Only among sixth-graders.

H3: Lower reasoners benefit more from concrete-relevant features than higher reasoners: reasoning-by-condition (C-r vs. A-r) interaction.

- H3a: Rule learning: No: higher-reasoning (oldest) students benefited.
- H3b: Mod-3 near-transfer: Yes.
- H3c: Mod-4 far-transfer: No.

H4: Youngest students benefit more from concrete-relevant features than older students: grade-by-condition (C-r vs. A-r) interaction.

- H4a: Rule learning: No: older (and higher-reasoning) students benefited.
- H4b: Mod-3 near transfer: No.
- H4c: Mod-4 far transfer: Only indirectly.

H12: Mapping facilitated most by concrete-relevant figures, then abstract-relevant figures, and least by abstract figures: Yes.

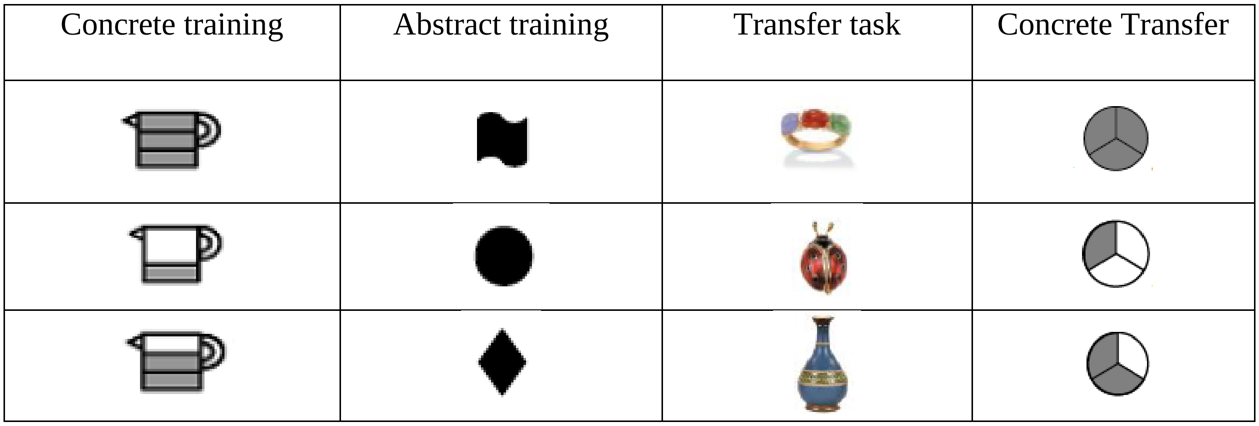


Figure 1. Figures in Kaminski et al. (2008) and De Bock et al. (2011) concrete training, abstract training, and transfer task. The last column shows the figures used in the concrete transfer task in De Bock et al. (2011). Corresponding figures across tasks are shown across each row.

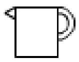
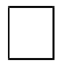
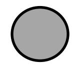
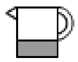
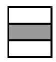
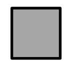
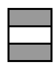
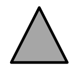
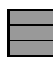
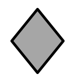
Training			
Concrete-relevant	Abstract-relevant	Abstract	Transfer
			0
			1
			2
			3

Figure 2. Analogous training figures by condition.

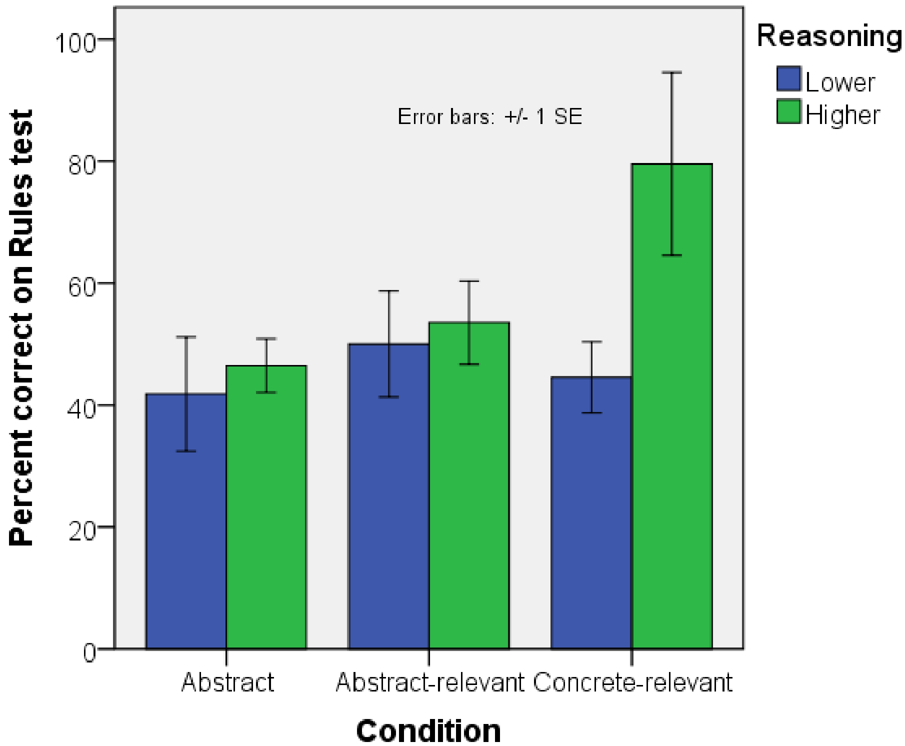


Figure 3. Condition by reasoning interaction for Rules test performance (eighth-graders).

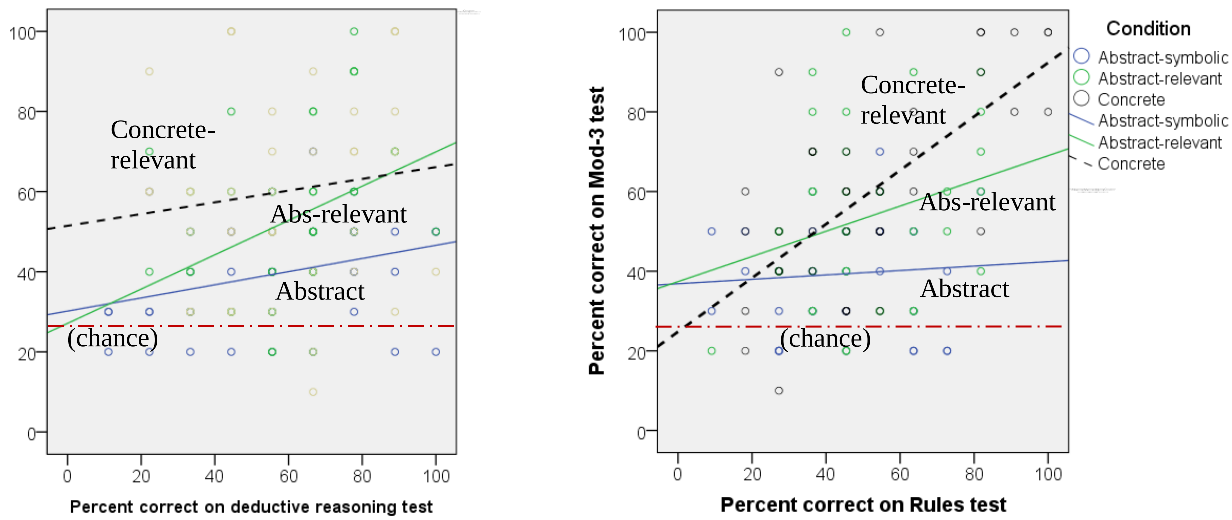


Figure 4. Interactions between condition and (a) reasoning, and (b) initial learning for near-transfer performance.

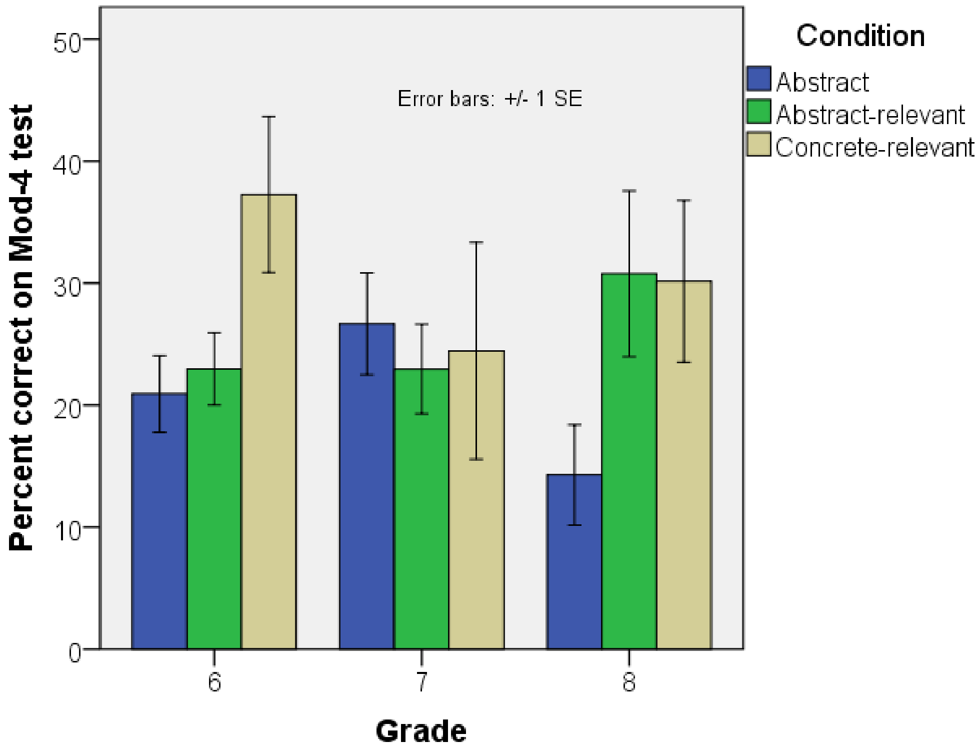


Figure 5 Percent correct on far-transfer assessment by condition and grade.

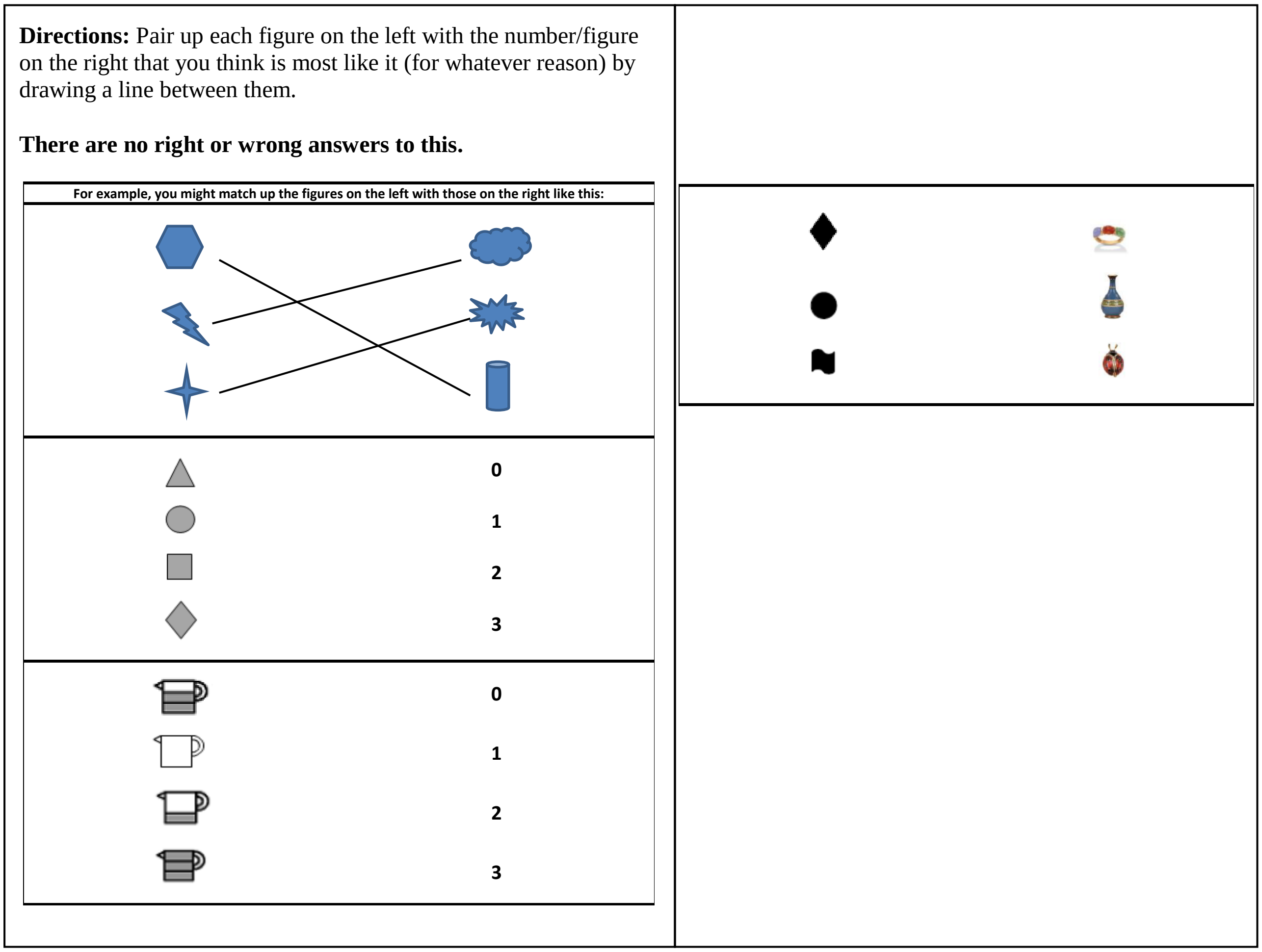
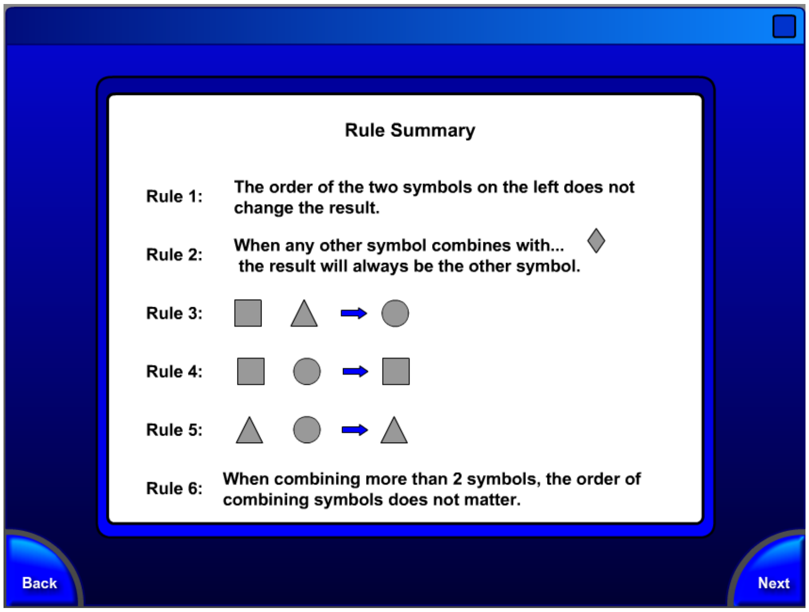
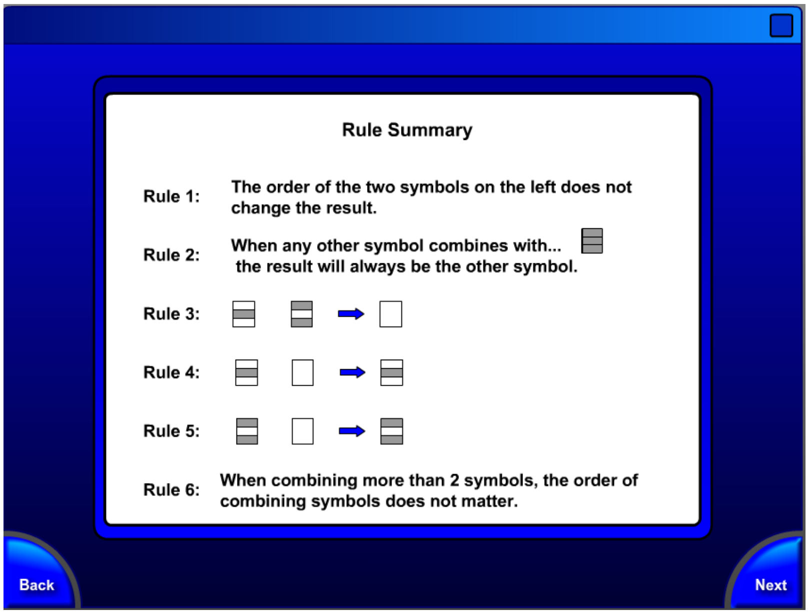


Figure 6. One version of mapping task given to sixth-grade students.

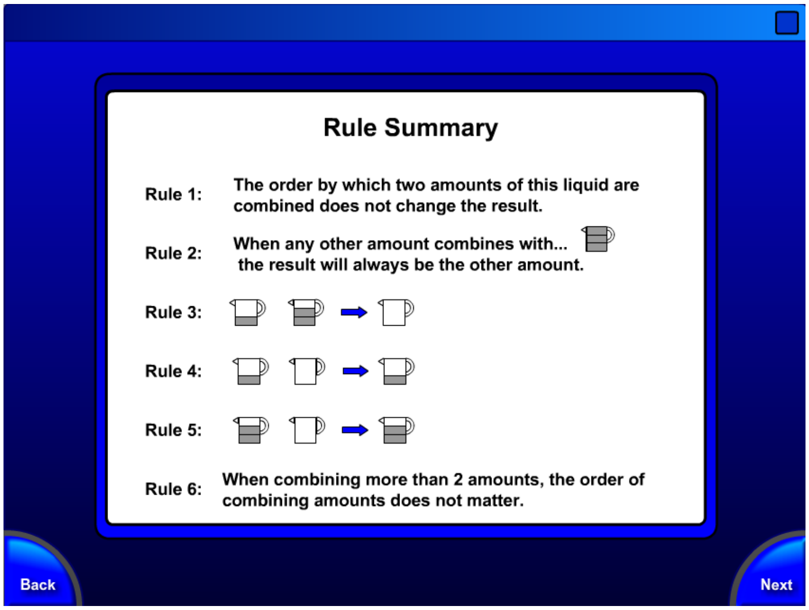
Appendix A. Screenshot of the rule summary page (Abstract condition).



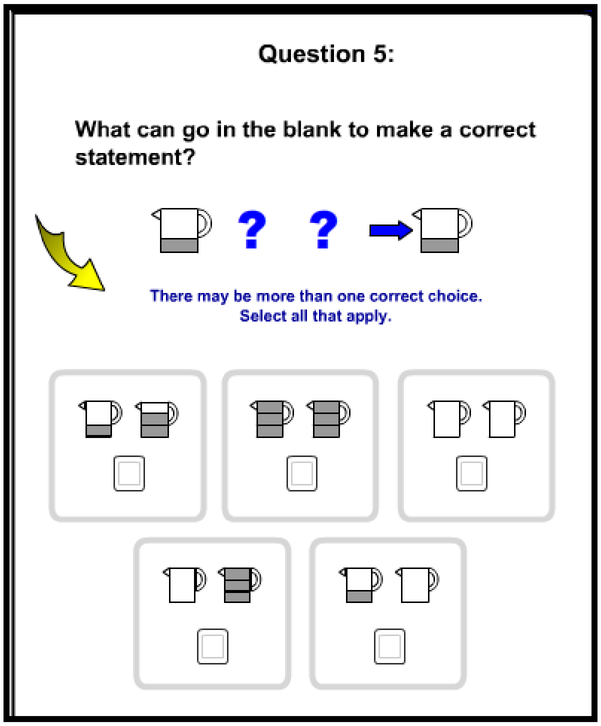
Appendix B. Screenshot of the rule summary page (Abstract-relevant condition).



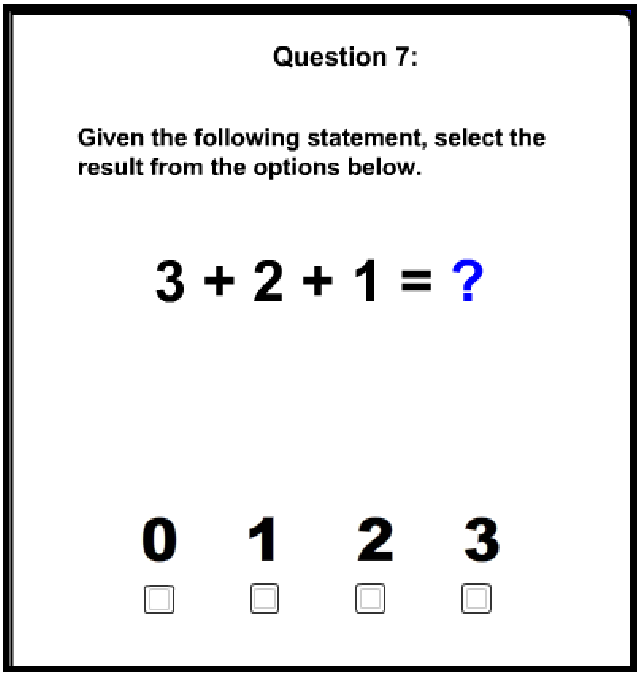
Appendix C. Screenshot of the rule summary page (Concrete-relevant condition).



Appendix D. Rules test items (Concrete-relevant condition).



Appendix E. Mod-3 assessment sample item.



Appendix F. Mod-4 sample item.