THE JOURNAL OF THE LEARNING SCIENCES, *14*(1), 69–110

The Transfer of Scientific Principles Using Concrete and Idealized Simulations

Robert L. Goldstone and Ji Y. Son
Department of Psychology
Indiana University

Participants in 2 experiments interacted with computer simulations designed to foster understanding of scientific principles governing complex adaptive systems. The quality of participants' transportable understanding was measured by the amount of transfer between 2 simulations governed by the same principle. The perceptual concreteness of the elements within the first simulation was manipulated. The elements either remained concrete throughout the simulation, remained idealized, or switched midway into the simulation from concrete to idealized or vice versa. Transfer was better when the appearance of the elements switched, consistent with theories predicting more general schemas when the schemas are multiply instantiated. The best transfer was observed when originally concrete elements became idealized. These results are interpreted in terms of tradeoffs between grounded, concrete construals of simulations and more abstract, transportable construals. Progressive idealization ("concreteness fading") allows originally grounded and interpretable principles to become less tied to specific contexts and hence more transferable.

Cognitive psychologists and educators have often debated the merits of concrete versus idealized materials for fostering scientific understanding. Should chemical molecules be represented by detailed, shaded, and realistically illuminated balls or by simple ball-and-stick figures? Should a medical illustration of a pancreas include a meticulous rendering of the islets of Langerhans or convey in a more stylized manner the organ's general form? Our informal interviews with mycologists at the Royal Kew Gardens (personal communication, Brian Spooner and David Pegler, May 1998) indicate a schism between authors of mushroom field guides.

Correspondence and requests for reprints should be sent to Robert Goldstone, Department of Psychology, Indiana University, Bloomington, IN 47408. E-mail: rgoldsto@indiana.edu

Some authors argue that the ideal book format for teaching readers to identify mushrooms is to present actual photographs of specimens. Other authors argue that line drawings are more effective than photographs despite their decreased realism because they can subtly emphasize diagnostic features and reduce variation unrelated to species identification. Our experiment explored the cognitive costs and benefits of concrete and idealized external representations.

The influence of concreteness is studied in the context of participants learning abstract scientific principles while interacting with computer simulations that instantiate those principles in particular domains. Participants' effective understanding of a principle is measured by their ability to comprehend and solve problems in a new domain that is governed by the same principle. The initial and transfer domains are superficially dissimilar, so effective transfer requires participants to be sensitive to their abstract commonality. We are interested in analogical transfer not as an end in itself, but as an indicator of a participant's understanding of the shared deep similarity between superficially dissimilar but abstractly related domains (Barnett & Ceci, 2002; Bransford, Brown, & Cocking, 1999; Carraher & Schliemann, 2002; Gentner, 1989; Holyoak, 1984; Novick & Holyoak, 1991; Rumelhart, 1980).

We first describe some of the competing costs and advantages associated with conveying scientific principles with concrete, contextualized materials. This analysis leads us to propose a possible method for improving transfer that we call "concreteness fading." Concreteness fading is the process of successively decreasing the concreteness of a simulation with the intent of eventually attaining a relatively idealized and decontextualized representation that is still clearly connected to the physical situation that it models. Then we present two experiments examining the transfer of knowledge from an initial simulation to an abstractly related simulation. Concreteness fading is compared to the converse method of "concreteness introduction," as well as to controls that maintain a constant level of concreteness or idealization throughout the simulation. These conditions allow us to explore the roles of external concreteness, graphical variability, and sequential order of graphical formats in the acquisition of abstract scientific understanding.

THE BENEFITS OF BEING CONCRETE

Well-established lines of philosophical and psychological argumentation have converged on the thesis that abstract understanding is most effectively achieved through experience with perceptually rich, concrete representations. Even if one's goal is to develop abstract scientific knowledge, there are several reasons to believe that these abstractions are not effectively learned through direct exposure to the abstractions in verbal or mathematical form. Instead, the abstractions are effectively learned by exposure to pictures, movies, interactive simulations, and real-world physical experi-

ences that embody the abstractions. By concrete external representations, we are including both perceptually detailed and rich materials as well as materials affording concrete, perceptual-motor activity. A brief description of some of these arguments is presented here and summarized in the left column of Table 1.

Decision makers are often more strongly affected by vivid, perceptual information than abstract statistical information (Nisbett & Ross, 1980) or descriptions lacking rich detail (Reyes, Thompson, & Bower, 1980). Deductive reasoning is facilitated when the domain is familiar and concrete rather than abstract (Wason & Shapiro, 1971). For example, people are better able to solve logically equivalent versions of a reasoning task when the cover story for the task concerns a familiar rather than an unfamiliar situation (Johnson-Laird, Legrenzi, & Legrenzi, 1972).

Concrete materials can support abstract reasoning because they can be explicitly designed to promote true inferences from perceptual representations to abstract principles (Bassok, 1996). Good external representations are crafted so that perceptually salient properties correspond to critical formal abstractions (Goldstone & Barsalou, 1998; Larkin & Simon, 1987). For example, the acquisition of mathematical expertise is facilitated if the perceptual chunks afforded by automated tutoring systems provide inductive support (Koedinger & Anderson, 1998) for important mathematical formalisms (Koedinger & Anderson, 1990). In general, we can engineer our environment and our perspective on this environment so that the perceptually salient cues properly support higher level cognitive processes.

TABLE 1
Advantages of Concrete and Idealized Representations

Advantages of Concreteness	*Advantages of Idealization*
Concrete information is easier to remember than abstract information.	Idealizations are potentially more transferable to dissimilar domains because knowledge is not as tied to a specific domain.
It is often easier to reason with concrete representations using mental models than abstract symbols.	The critical essence of a phenomenon is highlighted because distracting details are eliminated.
Visual processes used for concrete objects can be co-opted for abstract reasoning.	There may be an active competition between treating an entity as a symbol versus an object, and idealization makes symbolic interpretations more likely.
Concrete details are not always "superficial," but rather provide critical information about likely behavior and relevant principles.	Cognitive processing of less important but complex concrete elements is conserved.
Concrete materials are often more engaging and entertaining and less intimidating.	Idealizations facilitate interpretations of a situation in terms of abstract relations rather than specific attributes.
Concretely grounded representations are more obviously connected to real-world situations.	

The value of concrete representations has been frequently noted in education. One recommendation coming out of this research has been to use concrete materials that can be physically manipulated to teach abstract concepts (Moch, 2001). A well-known example of this strategy is to teach addition and subtraction of large numbers using Dienes blocks (Fuson & Briars, 1990). These blocks are arranged in units (individual cubes), longs (lines of 10 cubes), flats (sets of 10 lines), and boxes (stacks of 10 flats). The goal of using Dienes blocks is not only to solve specific numeric problems but also to provide students with a well-grounded, intuitive understanding for more formal arithmetic operations. Other well-known pedagogical manipulatives include geoboards (boards with a lattice of pegs and loose rubber bands to wrap around the pegs), Cuisenaire rods (colored wooden bars cut to integer lengths), and balance beams. The use of physical manipulatives has been advocated as providing scaffolding for abstract concepts (Bruner, 1966). Although proposals for manipulatives have outstripped carefully executed studies documenting their value, there is growing empirical evidence that manipulatives do offer educational benefits, by connecting abstract mathematical concepts to the real world (Kennedy & Tipps, 1994), overcoming students' anxiety about mathematics (Martinez, 1987), and teaching mathematics in an entertaining and engaging manner.

Manipulatives have traditionally been understood as concrete objects, but the seemingly oxymoronic term *virtual manipulative* has been coined to refer to computer-based, interactive simulations with graphical and dynamic elements that model real-world entities (Moyer, Bolyard, & Spikell, 2002). A computer simulation explicates scientific concepts by creating simplified, working models that are typically under parametric control by the simulation's user (Miller, Lehman, & Koedinger, 1999; Resnick, 1994; Schank & Farrel, 1988; Wilensky & Reisman, 1999). Computer simulations have been shown to confer a number of advantages over traditional educational practices. First, they provide a perceptual grounding for concepts that might otherwise be too abstract to readily comprehend. Second, they promote an active, hands-on, problem-solving stance that, in turn, often fosters a deep understanding of a phenomenon (National Research Council, 1999). Third, they provide effective exposure to experimentation skills that involve a cycle of hypothesis formation, testing, evaluation, and revision (White, 1993; White & Frederiksen, 1998). A noteworthy example of an apparently effective, grounded pedagogical simulation environment is Starlogo and its offshoots, StartlogoT and Netlogo (Resnick, 1994; Resnick & Wilensky, 1998; Wilensky, 1991, 1999, 2001). The Starlogo world consists of fixed patches, and moving agents that interact with the patches and each other. Students control dynamic parameters of the world by manipulating sliders, buttons, and the mouse. Interactive, graphical, and flexible simulations allow students to learn scientific principles such as the diffusion of beliefs, the spread of epidemics, self-organization in slime molds, the population dynamics of predator–prey systems, and the behavior of populations of gas molecules under pressure. Although formal laboratory studies of the pedagogical

efficacy of Starlogo have not yet been conducted, students' interactions with Starlogo systems indicate that these systems can foster intuitive understandings of complex systems more effectively than experience with formal equations or noninteractive movies (Resnick & Wilensky, 1998; Wilensky & Resnick, 1999).

An important design consideration when developing pedagogical computer simulations is determining how concrete to make the simulation elements. High levels of detail, produced by realistic rendering of objects within the simulation, may benefit learners by making a phenomenon concrete and by increasing the similarity between the simulation and real-world situations (DiFonzo, Hantula, & Bordia, 1998). In fact, most research in virtual reality has as an explicit goal the realistic mimicking of real-world phenomena (Grady, 1998; Heim, 2000). Using highly realistic computer simulations is often an intrinsically motivating experience because they are entertaining and evocative.

THE BENEFITS OF IDEALIZATION

Although this evidence suggests the pedagogical power of concrete, or at least "virtually concrete," learning materials, this conclusion has received a number of challenges, summarized in the right column of Table 1. Although the virtual reality community typically assumes "the more realistic, the better," other researchers have argued that simplified, relatively idealized representations are useful for distilling a situation to its essence (Gianutsos, 1994; Goldstone, Steyvers, & Rogosky, 2003; Smith, 2003). Idealization may assume various forms, from representations that eliminate detailed physical properties to symbolic, mathematical formalisms. In the following review, we focus on idealizations that are simplified physical-spatial representations because this is how we operationalize idealization in the reported experiments.

A strong empirical case for the value of less realistic representations comes from DeLoache's (1991, 1995; also DeLoache & Burns, 1994; DeLoache & Marzolf, 1992) research on children's use of physical models as representations. In her standard paradigm, a child around the age of 2.5 years is shown a model of a room; the child watches as a miniature toy is hidden behind or under a miniature item of furniture in the model, and the child is told that a larger version of the toy is hidden at the corresponding piece of furniture in the room. Children were better able to use the model to find the toy in the actual room when the model was a two-dimensional picture rather than a three-dimensional scale model (DeLoache, 1991; DeLoache & Marzolf, 1992). Furthermore, DeLoache (2000) decreased children's access to a model of a room by placing it behind a window, allowing the children to see but not manipulate the model. This manipulation improved children's performance. Conversely, in another experiment, children who were given

10 min of playtime with the model before being asked to use it as a symbol for the real room were less effective at finding the toy.

DeLoache (1995), Uttal, Liu, and DeLoache (1999), and Uttal, Scudder, and DeLoache (1997) explained these results in terms of the difficulty in understanding an object as both a concrete, physical thing and as a symbol standing for something else. As an object's concrete, physical properties become more salient, its ability to serve as a symbol decreases. A two-dimensional picture serves as a better symbol because its concrete properties are less salient. More generally, Uttal et al. (1999) argued that concreteness is not always beneficial. Although concreteness can help young children detect symbolic relations, it can make it more difficult for them to comprehend the abstract concepts represented by the concrete object. Another example of this is found in transfer among word problems in mathematics. Bassok and Holyoak (1989) examined transfer between isomorphic domains of arithmetic progression in algebra and constant acceleration in physics. They found greater positive transfer from the algebra problems to the physics problems than vice versa. They interpreted their results as showing that when abstract mathematics is easily isolated from the content-specific cover story in which it occurs, it transfers widely to different situations. Likewise, children (Rattermann & Gentner, 1998) and adults (Goldstone, Medin, & Gentner, 1991; Markman & Gentner, 1993) are less likely to respond on the basis of abstract relations among objects in a scene and are more likely to respond on the basis of superficial object attributes as the richness of the objects in the scene increases. As with the DeLoache studies, when the concrete manifestation of an abstraction is difficult to ignore, then it adversely impacts responding on the basis of abstractions.

Consistent with this tension between abstract and superficial construals, several researchers caution against the unbridled use of concrete manipulatives for education. One difficulty of teaching with concrete objects is that the connection between the objects and the desired abstraction is typically not transparent to a student, and unless students are explicitly instructed on the connection, use of manipulatives may not result in general arithmetic knowledge (Clements & McMillen, 1996; Uttal et al., 1997, 1999). Research has often failed to show a consistent advantage for manipulatives over more traditional instructional methods (Sowell, 1989). Resnick and Omanson (1987) found that mathematical knowledge gained with Dienes blocks did not generalize well to solving symbolic problems. Students did learn to effectively solve subtraction and addition problems with the blocks, but their performance did not positively correlate with their ability to solve symbolic problems such as "24 + 19 = ?"

Schwartz (1995), Schwartz and Black (1996a, 1996b), and Schwartz and Moore (1998) argued that realistic displays encourage people to use analog and dynamic imagery, whereas more schematic displays encourage analytic reasoning strategies. For example, Schwartz (1995; Schwartz & Black, 1996a) asked adults to determine whether marks on hinges and gears would meet if the mechanisms

were put in motion. When the display of the mechanisms was photo-realistic, people imagined the dynamics of the devices closing or rotating into position. In contrast, when the display of the mechanisms was a diagram, people instead extracted the metric properties of the display that they then compared or used in thumbnail derivations. Schwartz and Moore found greater use of mathematical reasoning when a proportional reasoning problem was displayed with a drawn diagram rather than with a photograph. Although these experiments did not reliably show better performance for idealized diagrams than more realistic representations, they did show that an advantage of idealization is that they promote more formal reasoning methods that are less influenced by physical constraints.

A major difficulty with manipulatives is that they bias students toward concrete construals of situations, and if a student is at risk to give a situation a concrete rather than abstract construal, then the manipulative may exacerbate this tendency (Ambrose, 2002). Consistent with this, Goldstone and Sakamoto (2003) explored factors that facilitate the transfer of an abstract scientific principle from one domain to another. Students interacted with two computer simulations that were governed by the same principle. For example, students learned about simulated annealing in the context either of balls falling on a hilly landscape or finding a pathway around obstacles. The initial simulation was presented using relatively realistic or idealized elements. Better transfer between simulations was found when the elements were idealized, particularly for students with relatively poor initial understanding. Following DeLoache's lead, they interpreted their results in terms of a competition between abstract and concrete construals of the simulations. Individuals prone to concrete construals tended to overlook abstractions when concrete properties or superficial similarities were salient. Thus, both concrete manipulatives and computer simulations potentially run the risk of tying students' understanding too tightly to the studied context, limiting the transportability of the underlying abstraction.

CONCRETENESS FADING AND INTRODUCTION

The previously discussed review indicates advantages and disadvantages to both concrete and idealized presentation strategies. On the one hand, concrete pedagogical materials provide useful and engaging perceptual scaffolding for abstract concepts that would be difficult to convey otherwise. On the other hand, concrete materials may encourage learners to develop internal representations that are overly literal and tied to specific contexts. A valuable pedagogical goal is to find instructional methods that combine the perceptual scaffolding provided by concrete materials with the potential for abstract transfer fostered by more idealized materials.

To this end, we put forth the method of concreteness fading as a candidate for partially satisfying both goals of grounded and generalizable representations. By concreteness fading, the graphical elements used in a simulation are originally

concrete but become more idealized over time. The intention is for early experience with concrete representations to promote intuitive and strong connections between the elements of the simulation and their corresponding real-world elements. Then, by later idealizing the elements, the psychological representations associated with the elements become less tied to their specific domain and hence more capable of being transferred to new domains that are superficially dissimilar.

One of the intellectual predecessors of concreteness fading is the work of Terrace (1963a) on errorless learning and his work on fading (Terrace, 1963b). Terrace (1963a) succeeded in training pigeons to learn a red–green discrimination without the pigeons making any errors. Pigeons pecked at a red key to receive food reinforcement. The green key started out as a dark key that appeared for only a short period of time and, as such, was unlikely to elicit responses. Pigeons continued to ignore the green key as it brightened and appeared for longer periods of time. Two advantages of errorless learning are that the learners do not experience unpleasant and distracting emotions associated with errors, and they do not confuse error responses with correct responses and thereby set down incorrect habits. Terrace (1963b) subsequently used fading to have pigeons who had already learned the red–green discrimination learn a horizontal–vertical discrimination without making errors. Vertical and horizontal lines were initially superimposed on top of red and green backgrounds, respectively. Pigeons would respond on the basis of their learned red–green discrimination. Gradually, the backgrounds were faded out so that only the lines remained. Both errorless learning and fading have been shown to dramatically improve discrimination learning. The analogy to concreteness fading is that, initially, participants can use a perceptually rich, concrete representation to understand a situation. Then, if this perceptual information is faded out, participants may continue to respond on the basis of other abstract relations that were associated with the perceptual information.

Another example of fading with direct relevance to developing abstract representations is the work of Kotovsky and Gentner (1996; see also Gentner & Wolff, 2000) on 4-year-olds' responding to relational similarities. Children were given a task requiring them to say whether a structure embodying symmetry (e.g., XOX) was more similar to another instance of symmetry (e.g., HIH) or to a second structure lacking symmetry (e.g., IHH). On some trials, there was perceptual support for placing X in correspondence with H and O in correspondence with I. For example, X and H were both dark and large, and O and I were both light and small. On other trials, this perceptual support was lacking. Children were better able to make their choices based on symmetry when these choices were supported by perceptual similarities. More interesting, children were better able to make symmetry-based choices that lacked perceptual support when these trials were preceded by trials with the perceptual support present. The researchers (see also Gentner & Medina, 1998) argued that the initial matches between relationally similar structures helped children grasp relational structures by creating an initial set of correspondences

that were consistent with a relational interpretation. Then, when these perceptual supports were faded out, relational responding continued.

There are precursors to concreteness fading that are directly relevant to pedagogy. The aim of Freudenthal's (1983) "progressive formalization" program is to teach students mathematics by gradually shifting from concrete and realistic situations to more formally based representations. Koedinger and Anderson (1998) provided experimental evidence for the pedagogical efficacy of bridging from concrete to abstract representations in contrast to the opposite sequence. Nathan (1998) described a computational simulation called ANIMATE that allows students to directly manipulate equations to see visible consequences in the simulation that represents entities in the equation. Novice word-problem solvers face two potential difficulties—overreliance on the concrete situation models and overreliance on the formal methods to the exclusion of information from the represented situations. By initially introducing students to the concrete situation, and subsequently encouraging students to reason explicitly about the situations, Nathan (see also Nathan, Kintsch, & Young, 1992) found that the opposing difficulties can be avoided and students can successfully generalize their situation-specific knowledge to more formal understanding.

A final rationale for concreteness fading is that this process of idealizing a physical display may complement a corresponding psychological process of fading. Schwartz and Black (1996b) proposed a conceptual process of fading by which internally represented attributes of the external referents are removed. In a gear-turning problem, a set of gears that are initially represented in terms of surfaces and teeth become successively faded to simple circles, then rotation directions, and finally numeric quantities (see also Dixon & Bangert, 2004). A physical concreteness fading operation that parallels conceptual fading may serve as a helpful external cognitive aid that is well coordinated with preferred internal representations.

In testing the effectiveness of decreasing the concreteness of a simulation, we compare this concreteness fading technique to the converse method of "concreteness introduction." Concreteness introduction presents initially idealized elements that become more concrete over time. Both concreteness introduction and fading involve elements with variable appearance. Research has shown that presenting an abstraction with variable concrete manifestations increases the likelihood of acquiring the underlying abstraction (Bruner, 1966; Gick & Holyoak, 1980, 1983). One possible reason for good extraction of an abstraction with multiple physical instantiations is that the common abstraction is the primary commonality across the different instantiations. With only one physical instantiation, concrete and abstract properties are equally plausible candidates for attention. With multiple physical instantiations, only the abstract properties are retained throughout and thus become highlighted as good candidates for attention.

To explore whether stimulus variation in itself fosters abstract transfer, we included experiment conditions with a constant appearance of elements. In one con-

dition elements were idealized throughout the initial training simulation. In another condition, the elements were consistently concrete. If appearance variability is helpful in promoting transportable abstractions, then we would expect both the concreteness fading and concreteness introduction conditions to yield better transfer performance than these two conditions with no element variability. Conversely, if element variability distracts learners from developing abstract representations (Uttal et al., 1999), then the opposite ordinal pattern should result.

In our experiments, we employ a rather modest manipulation of concreteness and idealization. In both concrete and idealized conditions, the elements are in fact graphically presented. They vary in their amount of detail and the extent to which the graphical element intrinsically contains sufficient information to identify the real-world, concrete entity that it represents. In the concrete condition, for example, ants appear as line drawings of ants with details that include heads, thorax, abdomen, and legs. In the idealized condition, ants appear as simple black dots. This modest manipulation of concreteness is beneficial for our initial experimentation because, if an influence of concreteness is found, it can be relatively unambiguously interpreted. However, we acknowledge that the size of our experimental effects may well be reduced because of the subtlety of the manipulation. Furthermore, exploring a greater range of concreteness variation would allow greater generalization.

EXPERIMENT 1

Experiment 1 explored the relative value of consistently concrete elements, consistently abstract elements, concreteness fading, and concreteness introduction in terms of transferring an abstract principle across simulations. Learners interacted with two simulations that were related by a common scientific principle and were given quizzes that measured their comprehension of the principle.

The principle that we proposed to teach was "competitive specialization." One of the most important notions of cognitive science is that the parts of a system can organize themselves without the help of leader or centralized plan (Resnick, 1994; Resnick & Wilensky, 1993). Parts that start out homogeneous and undifferentiated can each become specialized as a result of interactions between the parts (O'Reilly, 2001). A well worked out example of this is the development of neurons in the primary visual cortex that are specialized to respond to visually presented lines with specific spatial orientations (von der Malsburg, 1973). Another application of self-organization is the allocation of resources to cover a territory. It is often optimal for different agents to be specialized for different regions. In these situations, a good solution is found if every region has an agent reasonably close to it. For example, an oil company may desire to place oil drills such that they are well spaced and cover its territory. If the oil drills are too close, they will redundantly access the same oil deposit. If the oil drills do not cover the entire territory, then some oil reserves will not

be used. As another example, flies and dung beetles often distribute themselves to food sources so that they are well separated and have a plentiful supply of food to cover. It has been shown that in some situations animals will distribute themselves optimally, matching the distribution of food resources (Stephens & Krebs, 1987).

An elegant solution to problems of (close to) optimal covering such as these is to allocate agents (e.g., neurons, oil drills, flies) by executing the following three steps repeatedly: (a) randomly selecting a resource from among the entire set of resources to be covered, (b) determining the closest agent to this resource, and (c) adapting this closest agent toward the resource with a relatively fast rate and adapting all other agents toward the resource with a relatively slow rate.[1] This algorithm works by creating adapted and differentiated agents. If all agents adapt as quickly as the closest agent, then, ironically, they cover the territory less optimally as a group, because all of the agents occupy a similar position at the territory's center of mass. By quickly moving the agent closest to a selected resource, it becomes even more adapted to the resource for which it was already specialized. By slowly moving the other agents, they are still free to become specialized for other, less well covered resources.

Ants and Food

The first example of competitive specialization involved ants foraging food resources drawn by a user. The ants followed exactly the three rules described previously. At each time step, a piece of food was randomly selected, and the ant closest to the food moved with one rate, and all of the other ants moved with another rate. In interacting with the simulation, a learner can reset the ants' positions, clear the screen of food, place new ants, move ants, start or stop the ants' movements, and set a number of simulation parameters. The two most critical user-controlled parameters determine the movement speed for the ant that is closest to the selected food (called "closest rate" in Figure 1) and the movement speed for all other ants ("not closest rate"). Starting with the initial configuration of three ants and three food piles shown in Figure 1, several important types of final configuration are possible and are shown in Figure 2. If only the closest ant moves toward a selected piece of food, then this ant will be the closest ant to *every* patch of food. This ant will continually move to new locations on every time step as different patches are

[1]Although it might seem that determining the closest agent to a particular resource patch requires centralized leadership, Grossberg (1976) has shown how the closest agent can be identified on the basis of purely local interactions among the agents. Specifically, if the agents engage in several rounds of inhibiting each other from adapting based on their own adaptation rates, then no matter what their initial proximities are to a resource, only the closest agent will have a positive adaptation rate. Less extreme versions of this "winner-take-all" network can be achieved by parameterically varying the number of rounds of mutual inhibition.

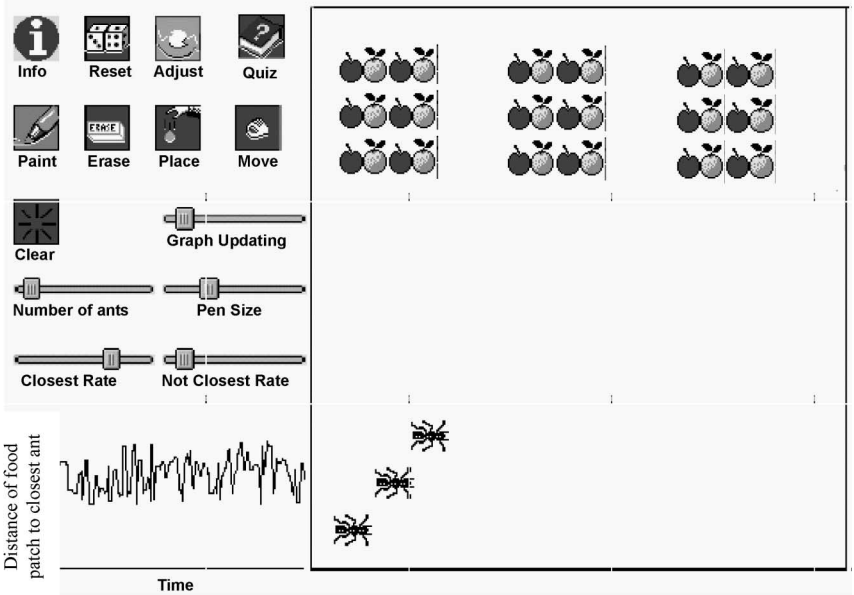


FIGURE 1 A screen-dump of an initial configuration for the "ants and food" simulation. At each time step, a patch of food is randomly selected, and the ant closest to the patch moves toward the patch with one speed (specified by the slider "closest rate") and the other ants move toward the patch with another speed ("not closest rate").

sampled but will tend to hover around the center of mass of the food patches. The other two ants will never move at all because they will never be the closest ant to a food patch. This configuration is suboptimal because the average distance between a food patch and the closest ant (a quantity that is continually graphed) is not as small as it would be if each of the ants specialized for a different food pile. If all of the ants move equally quickly, then they will quickly converge to the same screen location. This also results in a suboptimal solution because the ants do not cover the entire set of resources well. Finally, if the ant closest to a selected patch of food moves more quickly than the other ants, but the other ants move too, then a nearly optimal configuration is achieved. Although one ant will initially move more quickly toward all selected food patches than the other ants, eventually this ant will move toward a patch of food, thereby distancing itself from another patch of food that will then be controlled by another ant.

An important, subtle aspect of this simulation is that poor patterns of resource covering are self-correcting if good parameter values are used. If there are two ants and two resources, both ants will often initially be closer to one of the resources. However, when food from the other resource is selected, one of the two ants will still be the closest ant to it, and this ant will quickly move toward the uncovered resource.

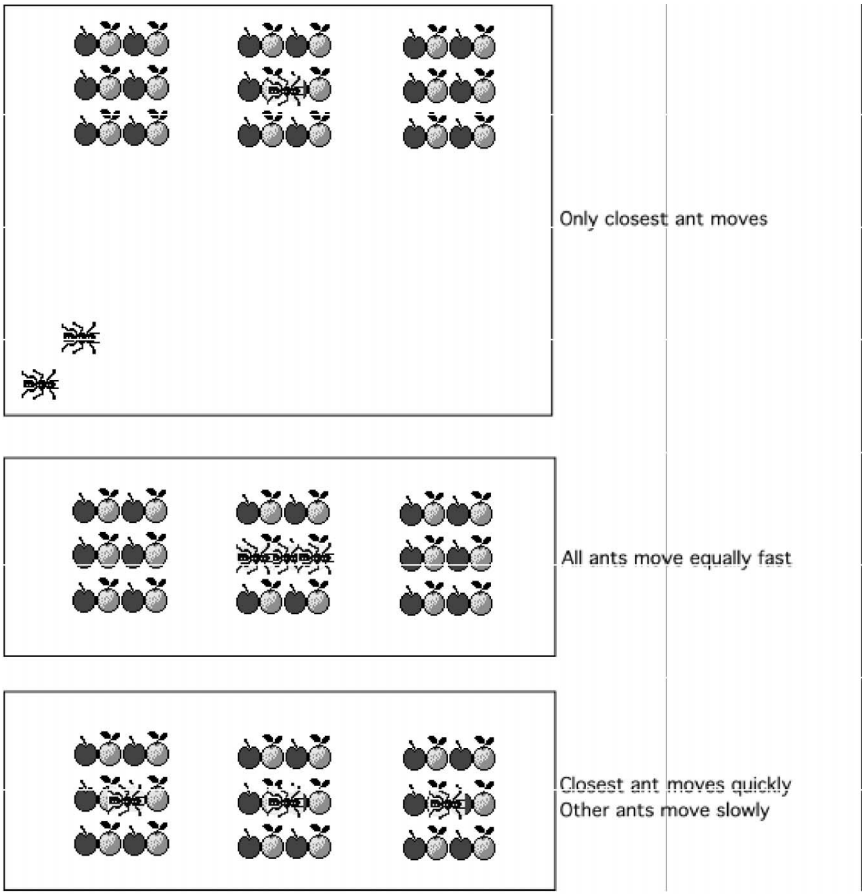


FIGURE 2 If only the ant closest to a selected food patch moves, and if all of the patches are fairly close, then often a single ant will move toward the average position of all of the food patches. If all ants move equally quickly, then all ants will move toward the average position. If the ant closest to a selected food patch moves much faster than the other ants but all ants move a bit, then each of the ants will become specialized for one food patch. This third possibility illustrates competitive specialization.

As it moves away from the crowded resource, the other ant will take control of territory formerly controlled by the departing ant. Rather quickly, the departing ant will no longer be the closest ant to any food in the formerly crowded resource and will then quickly move toward the less covered resource. The macroscopic impact of this interaction between ants and resources is almost always surprising to our participants. Ants will almost always self-organize themselves in a one-to-one relation to the resources regardless of the lopsidedness of their original arrangement.

Pattern Learning

The second example of competitive specialization is more abstract, dealing with the development of categories of visually presented patterns. The simulation was based on the unsupervised neural network learning algorithm of competitive learning (Rumelhart & Zipser, 1985). It is often desirable to have a system create categories that naturally capture the systematicities in a set of patterns. For example, if we present a system with examples of the letter "A" and examples of the letter "B," it is useful for the system to develop two categories—one for each type of letter. Once developed, these categories can be used as an efficient way of coding new inputs of the same types. One way of automatically creating appropriate categories is to randomly initialize the categories and then repeatedly select a picture, find the category that is most similar to the selected picture, adjust this category so that it even more closely resembles the selected picture, and adjust all of the other categories at a slower rate. This technique does not always produce optimal categories (for improvements to this algorithm, see Goldstone, 2003) but usually results in categories that resemble the major categories implicit among the pictures. Learners interacting with the simulation can control the rates by which the most similar category (the "winner") and the other categories (the "losers") adjust toward a selected picture, draw and edit pictures, set the number of pictures and categories, and start or stop the adjustment of categories. A continuously updating graph shows the similarity between the selected picture and the closest category to the picture and is computed by taking the inner product of the vectors representing the picture and category in question. Figure 3 shows an initial configuration of randomized categories and three pictures of letters that were drawn by a user. Figure 4 shows three potential category configurations resulting from different parameter settings. If the adaptation rate for the winning category is positive but the rate for losing categories is zero, then a single category will adapt toward all three letters (which are similar because of their common black backgrounds), leaving the other two categories unchanged. The single winning category will become a blend of all three letters, and consequently none of the letters will have a category that closely resembles it. The second panel shows the categories that emerged when all categories adapted equally quickly toward a presented picture. All categories will quickly become an identical blend of the same three pictures. The categories become progressively more similar to each other rather than differentiated over time because they are influenced by each presented picture in the same way. The third panel shows the category differentiation that occurs when the winning category adapts rapidly whereas the losing categories adapt much more slowly. Now each category becomes specialized for one and only one of the pictures. Although one category initially adapts more quickly toward all three pictures than the other categories, the other categories will eventually become similar to the pictures. At some point, one picture will attract the fast adapting category, and as the category adapts to become

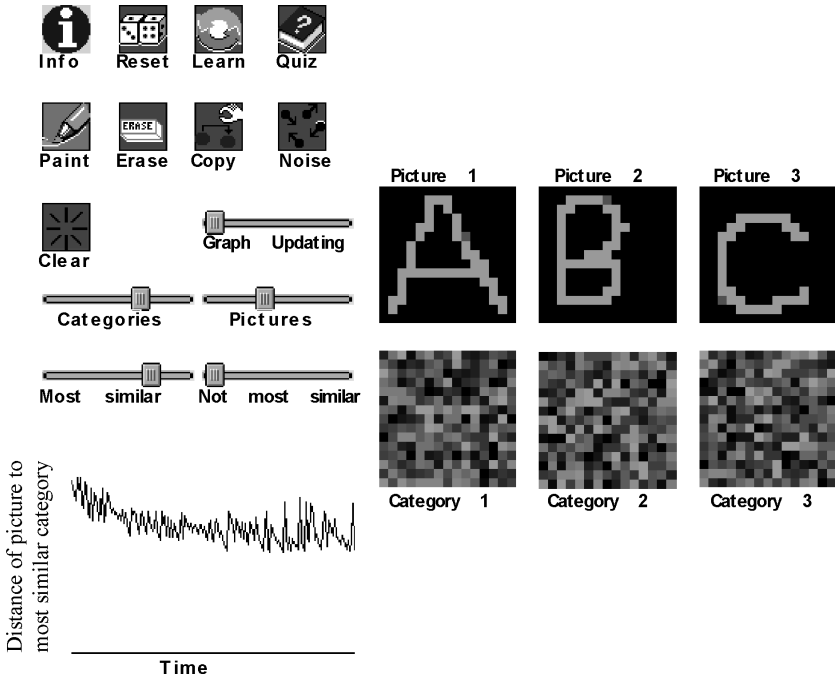


FIGURE 3 A screen-dump for the simulation "pattern learn." Users drew pictures, and prior to learning, a set of categories were given random appearances. During learning, a picture was selected at random, and the most similar category to the picture adapted its appearance toward the picture at one rate (specified by the slider "most similar") whereas the other categories adapted toward the picture at another rate ("not most similar").

more similar to the picture it will become less similar to the remaining pictures than another category. When this happens, one of the remaining categories will now adapt quickly to the remaining pictures. This process eventually leads to a one-to-one assignment of categories to pictures. The analogy between this situation and the ants foraging for food is hopefully clear. The three panels of Figure 2 are analogous to the respective panels in Figure 4. The two domains are deeply related because they are both instantiations of the principle of competitive specialization and are governed by the same mathematical formalism (Kohonen, 1995).

Method

Participants. Eighty-four undergraduate students from Indiana University served as participants to fulfill a course requirement. The students were split evenly into four conditions: consistently idealized, consistently concrete, ideal-

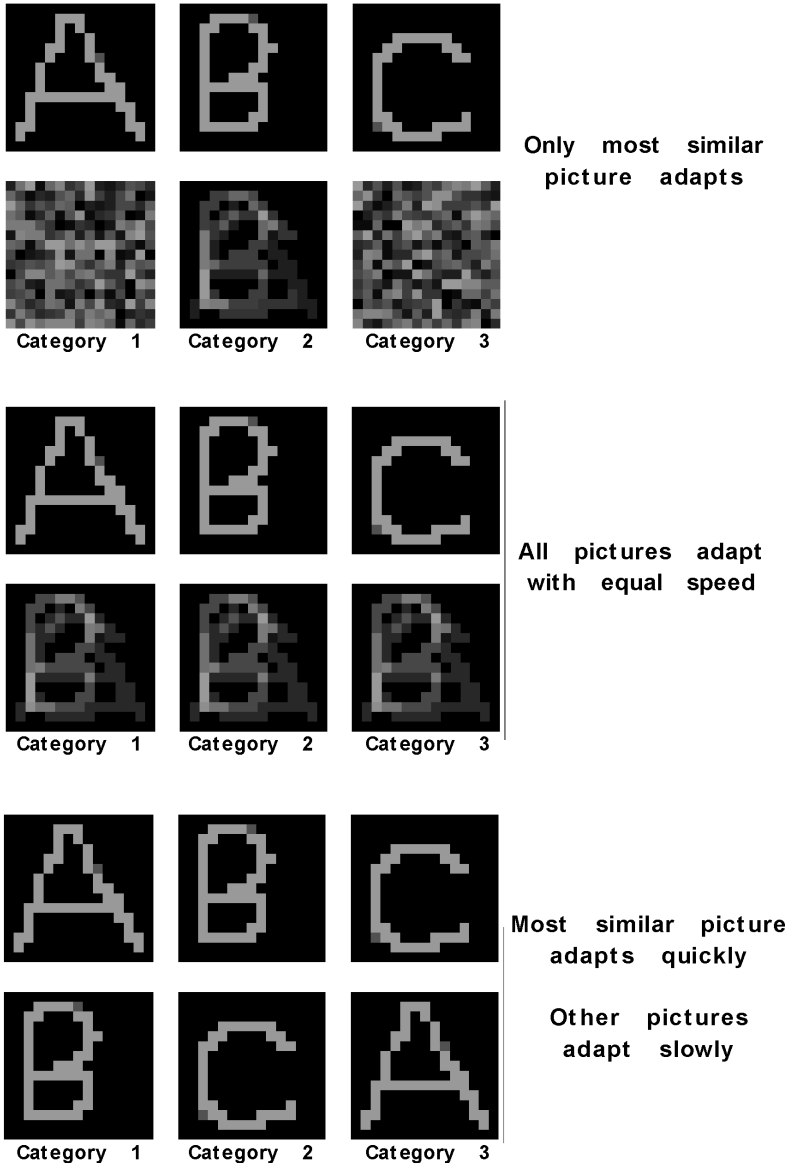


FIGURE 4 If only the category closest to a selected picture adapts, and the pictures are fairly similar, then often a single category will become a blend of all of the pictures. If all categories adapt equally quickly, then each category will become a blend of all of the pictures. If the category most similar to a selected picture adapts much more quickly than the other categories but all categories adapt a bit, then each category will become specialized for one picture. These three outcomes are analogous to the three outcomes shown in Figure 2.

ized-then-concrete (concreteness introduction), and concrete-then-idealized (concreteness fading).

Materials. Participants were allowed to freely explore the simulations shown in Figures 1 through 4. When presented with a simulation, participants were also given a one-page instruction sheet showing the rules by which the simulation operated and a general goal or line of inquiry to pursue while exploring the simulation. For the ants and food simulation, the goal was to have the ants distribute themselves to the food resources such that all of the resources had an ant that was reasonably close to it. For the pattern learning simulation, the goal was to develop category pictures that captured the natural groups present in the user-drawn pictures. Participants were given graphic examples of the goals. Participants were told that there were no hidden rules that described the behavior of the systems; everything that the agents did was governed by the three rules. Finally, participants were given descriptions of the important parameters of each simulation that were under their control.

Each simulation had a set of button and slider parameters on the left side of a 43-cm screen and a graphic window on the right side of the screen. Participants could draw and erase food resources and pictures in the graphics window. By moving the mouse and pressing buttons, participants could directly affect the graphics window by drawing, erasing, moving, and placing agents and cells. Buttons were used to reset the simulations, clear the screen, obtain help on using the simulations, turn on and off the simulations, and initiate the quiz. User-controlled sliders were used to control the continuously varying parameters of the simulations, such as the distance that the closest ant moved when a piece of food was randomly selected. See Figures 1 through 4 for the parameters used in both simulations and their arrangement. Each simulation also contained a continuously updating graph that plotted how a relevant measure of performance in the simulation varied over time (e.g., average closeness of ants to food patches). Free versions of the Macintosh software can be downloaded at http://cognitrn.psych.indiana.edu/rgoldsto/complex/.

Both simulations had dynamically changing displays inside the graphic window. These displays were updated every 17 msec and were instantaneously affected by user-controlled changes to parameter values. For example, as a participant reduced the movement amount of the closest ant to a piece of food, he or she would immediately see the ant move more slowly.

The simulations were all run on Macintosh iMac G4 computers. The participants were run in groups of 4 to 7, with each participant in an individual, sound-proofed cubicle with an overhead 25-watt light.

Procedure. Participants were told that they would be exploring two computer simulations but were not told that they were related in any way. For each of the simulations, they were first given an instruction page orienting them to the rules, parameters, and goal of the simulation. Then they were allowed to freely explore the simula-

tion for 20 min. During this time, a research assistant was available to answer questions that the students had about the rules governing the agents' behavior and the interface controls. During the exploration period, all key presses, mouse movements, and parameter changes were recorded by the computer, with time and duration information attached. See Appendix A for the instructions given to participants.

After the exploration period, participants were told to take the quiz by pressing the "Quiz" button. The computer then presented seven multiple-choice questions, samples of which are shown in Appendix B. Participants indicated their choice by clicking the mouse button while the cursor was inside a circle associated with the choice. After a choice was made, the computer automatically proceeded to the next question without giving any feedback on the correctness of a response. Participants were not allowed to go back to earlier questions.

Participants were first given one of four different versions of the ants and food simulation. After exploring this simulation and taking its associated quiz, participants were all transferred to the pattern learning simulation and after 20 min took the quiz associated with this simulation. In the consistently concrete condition, the ants and food had the appearances shown in Figures 1 and 2. The ants were depicted by relatively simple line drawings of black ants, and the food consisted of an orange peach and a red apple. In the consistently idealized version of the ants and food simulation, the ants were small black dots and the food sources were green patches, as shown in Figure 5. Participants "painted" food in the same manner in the concrete and idealized simulations, although in the concrete version, fractional portions of food were not allowed. The instructions for the two versions of the ants and food exercise were changed to reflect their different appearances, and extra

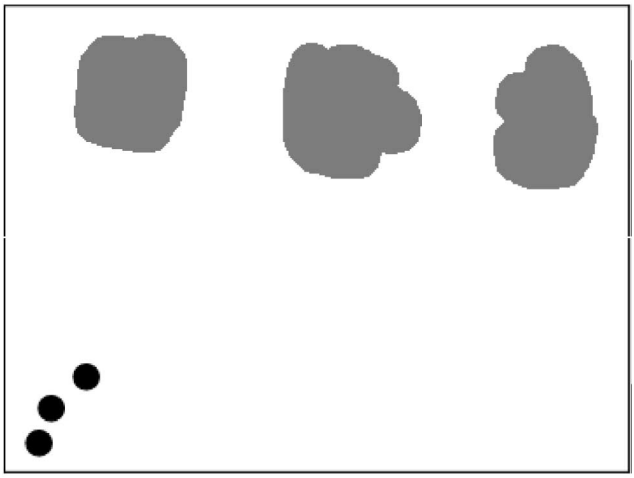


FIGURE 5 An example of the idealized version of the ants and food simulation.

measures were taken to assure that participants would interpret the small black dots as ants in the idealized condition. In particular, in a diagram from the simulation, green and black dots were labeled as food and ants, respectively. In the concreteness fading condition, concrete versions of ants and food were used for the first 10 min of the simulation. At the 10-min point, a message appeared on the screen that read "We are now changing the appearances of the food and ants, but they still behave just as they did before." From that point on, idealized versions ants and food were used. In the concreteness introduction condition, idealized versions of ants and food were replaced by concrete versions after 10 min.

After the second pattern learning simulation, participants were invited to describe what they felt they had learned about the simulations. Participants were invited to make drawings of scenarios from the simulations, describe the behavior, and then explain why the behavior occurred. Participants were also encouraged to explain why the simulations' critical parameters had their effect. These interviews were conducted one-on-one, so typically only 1 participant per experimental session was interviewed, and, in several sessions, time did not permit interviewing any of the participants. From the 84 participants, 12 participants were interviewed, with equal numbers of participants in each of the four conditions. Each interview lasted about 10 min.

Results

The four appearance conditions for the ants and food simulation differed in their quiz performances according to a between-subjects analysis of variance (ANOVA), $F(3, 80) = 9.1$, $p < .01$. Quiz scores on the initial training simulation are shown by the four left bars of Figure 6. Post hoc tests revealed that the idealized training condition produced worse performance than either concreteness fading or concreteness introduction, and the concrete training condition produced worse performance than concreteness fading, Fisher's PLSD (Protected Least Significant Difference) $p < .01$. To assess the benefits of appearance variability, the idealized and concrete conditions were combined together to form an average "consistent" quiz performance, and the concreteness fading and concreteness introduction conditions were averaged together to form a "variable" performance measure. The variable conditions performed better than the consistent conditions, with average quiz error rates of 56.8% and 61.3%, respectively, $F(1, 82) = 8.7$, $p < .01$.

The four training conditions also differed in their transfer to the pattern learning simulation, as shown by the four right bars of Figure 6, $F(3, 80) = 13.5$, $p < .01$. Performance on pattern learning was significantly better when the ants and food simulation was presented using concreteness fading than any of the three other presentation conditions, Fisher's PLSD $p < .01$. Concreteness introduction resulted in better transfer performance than either concrete or idealized conditions, Fisher's PLSD $p <$

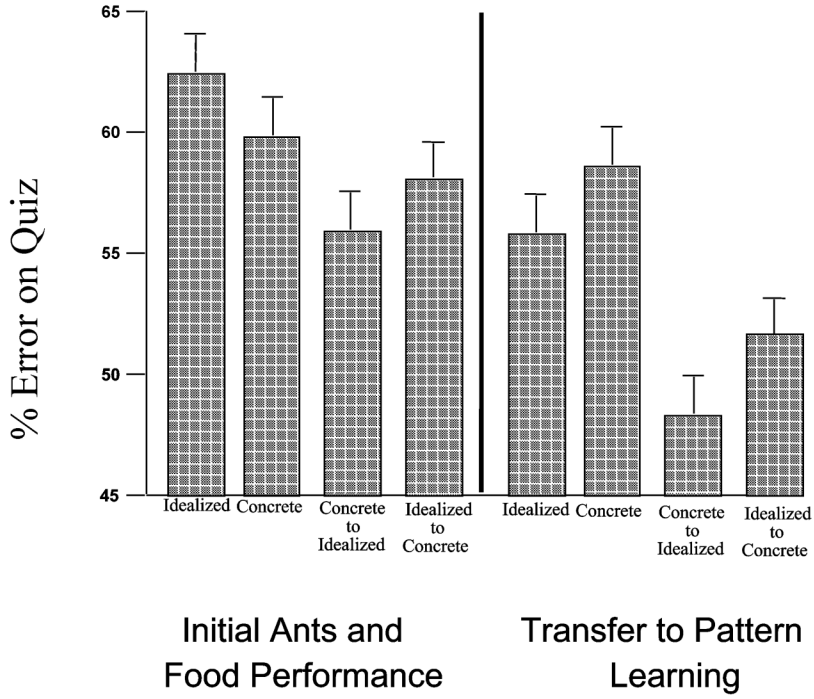


FIGURE 6 Quiz results from Experiment 1. The four bars on the left show performance on the initial quiz as a function of how elements from the initial simulation were presented. The four bars on the right show performance on the transfer quiz as a function of how elements from the initial simulation were presented.

.01. Combining together the variable and consistent conditions as before yielded transfer quiz error rates of 50.1% and 57.5%, respectively, $F(1, 82) = 15.0$, $p < .01$.

Although the concrete and idealized conditions did not significantly differ from each other according to post hoc tests for either the ants and food or pattern learning simulations, the results in Figure 6 suggest an interaction between simulation and condition. A repeated-measures ANOVA conducted with simulation as a within-subject variable and condition as a between-subject variable revealed an interaction between these two variables with quiz score as a dependent measure, $F(3, 80) = 4.8$, $p < .01$. In the first simulation, the concrete condition yielded high quiz scores whereas the idealized condition showed higher scores in the transfer simulation. When we limited this ANOVA to only the concrete and idealized conditions, the interaction was still significant, $F(1, 80) = 5.3$, $p < .01$.

Participants who performed well on the first simulation quiz tended to also perform well on the second simulation. In particular, there was a Pearson correlation of $r = .60$ between the quizzes for the two simulations, Fisher's r to Z

transformation, $p < .01$. There were also significant correlations between quiz performance and participants' manipulations while exploring the simulation. For each participant, the number of times that they manipulated each slider and button was tallied and correlated with quiz performance. For the initial ants and food simulation, two sliders were significantly correlated with quiz performance: closest rate ($r = .4$, $p < .01$) and not closest rate ($r = .37$, $p < .01$). The analogous sliders were also the only two sliders or buttons that were significantly correlated with performance on pattern learning: most similar ($r = .44$, $p < .01$) and not most similar ($r = .42$, $p < .01$).

In assessing the observed condition-dependent transfer, it is helpful to determine whether the transfer differences could be completely accounted for by differences in initial simulation understanding. Initial simulation understanding cannot explain the transfer difference between idealized and concrete conditions because the idealized condition produced worse initial simulation quiz performance than the concrete condition but produced better transfer to the second simulation. However, to assess this account for all four conditions, we conducted a regression analysis in which we partialed out the influence of initial quiz performance on transfer quiz score. With the initial score partialed out and the condition dummy-coded along four variables, there was still a significant influence of condition on transfer, $F(3, 80) = 4.5$, $p < .01$, $R^2 = 0.19$. Thus, we concluded that the advantage of concreteness fading was not simply due to its producing the best understanding of the initial simulation.

Given the relatively modest effect sizes of our concreteness manipulations, a potential concern was that the effects may be caused by just one or two of the quiz items. We conducted an item analysis to investigate the generality of the influence of graphical concreteness. The percentage of error responses to the initial quiz questions ranged from 50% to 68% for the seven questions. An item by condition ANOVA revealed significant main effects of both item, $F(6, 480) = 6.5$, $p < .01$, and condition, $F(3, 80) = 7.9$, $p < .01$, but no significant interaction between these two variables. The lack of a significant interaction suggests that we do not have good cause to reject the null hypothesis that the influence of condition had equal effects over all of the items. A similar item analysis was conducted on the transfer simulation quiz with similar results indicating main effects of item, $F(6, 480) = 7.2$, $p < .01$, and condition, $F(3, 80) = 6.0$, $p < .01$, but not a significant interaction between these variables.

A preliminary analysis of the interviews indicated that the most distinctive differences between the four conditions occurred when participants described the behavior of the ants and food simulation. This is presumably because the four conditions manipulated the appearance of graphical elements in this simulation. Given the relatively small number of participants interviewed in this experiment, analyses of these results will be integrated with Experiment 2's interviews and reported in that experiment.

Discussion

The best performance on both the initial and transfer simulations was obtained when the initial simulation's elements were initially concrete but became idealized over the course of the initial simulation. Concreteness fading has the advantage of providing an initial, concrete grounding for simulation elements but also eventually idealizing these elements in a manner that allows them to be generalized to abstractly similar but superficially dissimilar situations. The benefits of concreteness fading are consistent with theories that stress the tradeoff between concrete representations that make model-world relations transparent and idealized representations that promote understandings that can be transported to new situations. A problem with exclusively idealized representations is that participants may require additional cognitive resources to keep track of the relation between the model and the represented world. Concrete representations make this relation clear, thereby freeing cognitive resources. However, a problem with exclusively concrete representations is that the resulting understanding may be too influenced by the training domain to produce wide generalization.

Both concreteness fading and concreteness introduction employ concrete and idealized representations. As such, the advantage of concreteness fading over concreteness introduction must be due to something more than employing both representations. One likely candidate is that the initial concrete presentation promotes correct correspondences between the world and model when they are most needed—at the beginning of the simulation when participants are most uncertain about the meaning of the simulation's elements. When originally concrete elements are idealized, learners are unlikely to forget the elements' original correspondences. However, if originally idealized elements are only subsequently made concrete, it may be difficult for learners to remember the interpretation of the simulation elements, and this would make learning inefficient during the initial simulation. Another way of putting the advantage of concreteness fading is that it allows learners to initially develop a grounded understanding of a specific domain using superficial appearances as a "crutch," and then it removes this crutch, thereby allowing the learner's representations to be more transportable to new situations.

A remaining question is "Why does concreteness fading help even the initial simulation performance?" Concreteness fading resulted in better performance on the ants and food simulation than either idealized or concrete conditions. Developing an abstract construal of a situation might be expected to result in better transfer to superficially dissimilar situations, but its value for comprehension of the situation itself is more surprising. Part of an explanation for this might be that the quiz questions for the initial simulation also required an abstract construal of the ants and food simulation. The questions were designed to test learners' appreciation of the principle of competitive specialization. Concreteness fading may offer

benefits over a continuously concrete presentation because it emphasizes the abstract competitive specialization principle.

Another part of the explanation for the benefit of concreteness fading is that it employs varied graphical elements. Both of the conditions that presented variable graphical elements produced better performance than the two conditions that used constant elements throughout. This advantage for presenting a simulation with variable elements was found for the simulation itself as well as for the transfer simulation. In our domains, abstractions are apparently fostered by presenting multiple physical instantiations of the same abstract roles. This is consistent with Bruner's (1966) arguments for presenting multiple kinds of manipulatives when teaching a novel abstract concept and Gick and Holyoak's (1983) finding that abstract schema induction is facilitated by presenting multiple instantiations of the schema.

On the question of whether simulation elements should be idealized or concrete, Experiment 1 yielded conflicting results. Overall, across both simulations, there was not a significant difference between consistently concrete and consistently idealized conditions. However, there was a significant interaction involving simulation. For the initial simulation, the trend was for a concreteness advantage, but for the transfer simulation the trend was reversed. This interaction replicates previous results using the same pair of simulations (Goldstone & Sakamoto, 2003) and indicates a dissociation between performance on a task and generalization of the performance to an analogous task. Such dissociations have been previously observed, for example, with the initial learning and transfer of dart trajectories traveling through water (Judd, 1908) and the alteration of matchstick shapes to create new shapes (Katona, 1940). The dissociation can be interpreted in terms of how tightly a learner's knowledge is tied to the particular training domain. In the concrete condition, the learners' knowledge is predicted to be tightly tied to a domain. This domain specificity could be modeled by overspecialized conditions in condition–action production rules (Anderson, 1993) or computational learning algorithms that overfit training data and subsequently produce insufficiently general rules. The consequence of overspecialization is that the learner may be able to perform well on the trained domain but shows little ability to transfer his or her knowledge to related domains. By contrast, if the learner's knowledge is more abstract, it will transfer well to analogous domains, but this increased capacity for transfer may be at the expense of a solid and concrete understanding of the original domain.

EXPERIMENT 2

Experiment 1 tested learners' understanding of an abstract principle by giving learners quizzes. Experiment 2 was an effort to replicate the results from Experiment 1 using an alternative, performance-based measure of transfer. This measure gauged whether participants can better solve problems in the transfer simulation

when the problems have been preceded by a related simulation (Schunn & Dunbar, 1996). Participants who have a hard time expressing in words the competitive specialization principle may still be able to apply knowledge gained from the initial simulation to solve posed problems during the transfer simulation. When learners were transferred to the second simulation, they were given problems to solve that required appreciation of competitive specialization. Performance was measured by the time required to supply solutions to these problems.

Method

Participants. Eighty-eight undergraduate students from Indiana University who did not participate in Experiment 1 served as participants to fulfill a course requirement. An equal number of participants were assigned to the four presentation conditions.

Materials. The same two simulations used in Experiment 1 were used again. The transfer simulation, pattern learning, was slightly modified so that some buttons and drawing functions were disabled. These were disabled so that participants could not find solutions for problems in the transfer simulation without using a competitive specialization strategy. In particular, the Reset, Paint, Erase, Copy, and Noise buttons were disabled, as were the No. of Categories and No. of Pictures sliders.

Procedure. As with Experiment 1, participants first explored the ants and food simulation for 20 min and answered its associated multiple-choice quiz. Then participants were transferred to the second simulation, pattern learning. Participants were given the rules for this simulation and descriptions of the controls. Rather than freely exploring the simulation as participants in Experiment 1 did, Experiment 2 participants were given three problems to solve of increasing difficulty. The goal for each of the problems was to create category units that perfectly covered the input patterns. In the goal state, every input pattern had a category unit that perfectly matched it. For all three problems, there were two input patterns and two categories. In the first problem, one of the inputs was a solid bar occupying the left eight cells of a 15 × 15 grid, and the other input was a solid bar occupying the right seven cells of a 15 × 15 grid. These inputs shared the same values on 0% of their cells. This is an easy problem within the simulation because, even though the categories start off as randomized and homogenous, the category that is closest to one of the input patterns is always different from the category that is closest to the other input. In the second problem, one of the inputs was a large letter "A" in the middle of the screen and the other input was a large letter "B." These letters had the same values on 70% of their cells.

In the third problem, one stimulus was a large horizontal bar 15 cells wide and 6 cells high appearing at the horizontal midline with a thin vertical bar with dimen-

sions of 3 cells wide and 5 cells high appearing below the horizontal bar, thus forming a "T" shape. The other stimulus had the same vertical bar appear above the same horizontal bar, forming an inverted T shape. These two stimuli have the same values on 87% of their cells. This is a difficult problem to solve because even if appropriate slider values are given for how quickly the most similar category and other categories adapt towards an input, participants must still wait about 10 sec for the categories become specialized. Initially, one category almost always becomes the most similar category to both input patterns and quickly assumes an appearance that is the blend of both inputs. Assuming the optimal strategy of adapting the most similar category quickly and the other category slowly, the other category eventually adapts toward both inputs as well, and, once it is sufficiently similar to the inputs, it is the closest category to one of the two patterns and each category can then become specialized to a single input.

The computer simulation automatically detected when a participant solved a problem. When the normalized inner product of each input picture to its most similar category was greater than 0.98, the simulation announced a successful solution had been reached and automatically presented the next problem to the learner. After a learner successfully solved all three pattern learning problems (or 20 min had transpired), they were given the transfer quiz used in Experiment 1. All but 4 of the participants were able to complete the three problems within 20 min. A solution time of 20 min was recorded for the participants who failed to solve all three problems.

As with Experiment 1, a subset of participants were given an interview after the second simulation in which they made drawings of scenarios from the simulations, described the resulting behavior, and explained why the behavior occurred. Twelve of the 88 participants were so interviewed, 3 from each of the four conditions.

Results

The four appearance conditions for the ants and food simulation differed in their quiz performance according to a between-subjects ANOVA, $F(3, 84) = 7.6$, $p < .01$. Quiz scores on the initial training simulation are shown by the four left bars of Figure 7. Post hoc tests revealed that the idealized training condition produced worse performance than all of the other conditions, Fisher's PLSD $p < .01$. As with Experiment 1's results, the idealized and concrete conditions were combined together to form an average consistent quiz performance, and the concreteness fading and concreteness introduction conditions were averaged together to form a variable performance measure. The variable conditions performed better than the consistent conditions, with average quiz error rates of 57.9% and 60.7%, respectively, $F(1, 86) = 7.3$, $p < .01$.

The four training conditions also differed in their explicit quiz-based transfer to the pattern learning simulation, as shown by the four gray bars on the right side of Figure 7, $F(3, 84) = 11.7$, $p < .01$. Performance on pattern learning was signifi-

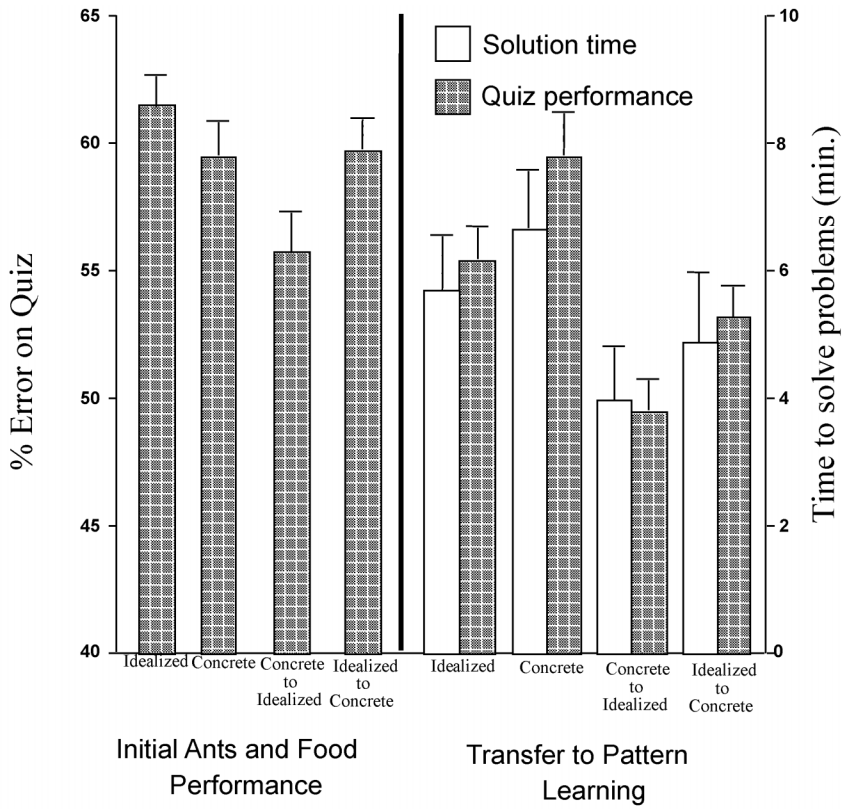


FIGURE 7 Explicit and implicit performance results from Experiment 2. Quiz result performance (gray bars) is read off of the left vertical axis, and time to solve problems (white bars) is read off of the right vertical axis. The four presentation conditions all refer to the presentation format during the initial ants and food simulation.

cantly better when the ants and food simulation was presented using concreteness fading than any of the three other presentation conditions, Fisher's PLSD $p < .01$. Concreteness introduction resulted in better transfer performance than the concrete condition, Fisher's PLSD $p < .01$. Combining together the variable and consistent conditions yielded transfer quiz error rates of 51% and 57.6%, respectively, $F(1, 86) = 9.5$, $p < .01$.

The results from the transfer measure of performance closely mirrored the explicit transfer measure. Conditions that produced relatively good quiz performances also produced fast solution times to the three problems. The solution times to all three problems were summed together to create the right axis of Figure 7. Overall, there was a highly significant effect of condition on cumulative solution

times, $F(3, 84) = 6.5$, $p < .01$. Transfer solutions were found significantly faster when preceded by the concreteness fading version of the initial simulation than either consistently idealized or concrete versions, and the concreteness introduction version resulted in significantly faster solution times than the concrete version, Fisher's PLSD $p < .01$.

To address the question of differences between the two measures of transfer, a multivariate analysis was conducted with the performance measure variable having two levels—time to solve problems and percentage of errors on transfer quiz—and condition as a between-subject variable. The resulting ANOVA revealed the same main effect of condition that the previous analyses revealed but did not reveal an interaction between performance measure and condition, $F(3, 84) = 1.5$, $p > .2$. This result, then, did not suggest any reliable differences between our measures of transfer.

Although the concrete and idealized conditions did not significantly differ from each other according to post hoc tests for either the ants and food or pattern learning simulations, the results in Figure 7 suggest the same interaction between simulation and condition that was found in Experiment 1. Using quiz performance as a dependent measure, a repeated-measures ANOVA conducted with simulation as a within-subject variable and condition as a between-subject variable revealed an interaction between these two variables, $F(3, 84) = 5.2$, $p < .01$. When we limited this ANOVA to only the concrete and idealized conditions, the interaction was still significant, $F(1, 86) = 5.9$, $p < .01$.

As with Experiment 1, we conducted item analyses to gauge whether the influence of concreteness was found for all of the initial quiz questions. The percentage of error responses to the initial quiz questions ranged from 47% to 65% for the seven questions. An item by condition ANOVA revealed significant main effects of both item, $F(6, 504) = 7.3$, $p < .01$, and condition, $F(3, 84) = 7.2$, $p < .01$, but no significant interaction between these two variables. A similar item analysis was conducted on the transfer simulation quiz with similar results indicating main effects of item, $F(6, 504) = 5.7$, $p < .01$, and condition, $F(3, 84) = 6.8$, $p < .01$, but no significant interaction between these variables.

Interviews. The interviews from Experiments 1 and 2 were combined because the procedure involving the initial simulation was identical for the two experiments, and analyses indicated that the largest difference between the conditions was in participants' descriptions of the first ants and food simulation. This is expected because the conditions only varied in the graphical representations used for the first, not the second, simulation. Given the greater diagnosticity of the ants and food simulation descriptions, only these descriptions were further analyzed. To analyze the descriptions of the 24 participants, unique descriptions were identified for each of the four conditions. A unique description was defined as a phrase given by a participant in a condition that was not given by a participant in any other condition. The large majority of the descriptions (76%) were not unique, but rather were given by several partic-

ipants across different conditions. Table 2 provides a list of the unique descriptions. To obtain this table, equivalence matching was performed to combine phrases from different participants that conveyed nearly the same meaning. For example, 3 participants mentioned ants scaring each other: "One ant scared the others away," "This ant is scared off by this other ant," and "because they are frightened off by the ant who's

TABLE 2
Unique Descriptions by Participants in Experiments 1 and 2 to the Ants and Food Simulation[a]

Idealized	*Concrete*	*Concrete to Idealized*	*Idealized to Concrete*
"Food in this clump pulls this ant away even though it starts off far from it."	"One ant scares the others away." (3)	"The ants move at different speeds, depending on what food they're moving to."	"When other ants move too quickly, then they leave their old, good locations."
"Animals move quickly to food that they are close to."	"The ants get tired after awhile." (2)	"The closest ant to this food is attracted to it." (2)	"The ants that are far away from a piece of food aren't attracted to it."
"Each ant will move to one of the piles." (2)	"This ant decides to focus on one food pile."	"When all ants move the same speed, they end up in the same spot." (2)	"The fruit is so close together that the ants try to cover both groups."
"It helps to make an ant move quickly at first, and then more slowly."	"The ants are tempted by both food piles." (2)	"Because this ant is the only one that's moving, it's trying to cover too many blobs."	"The ants end up getting caught in the middle between the two piles because they move toward each pile."
"Food is randomly selected and then ants move to it."	"The ants decide not to go to food that is being eaten by another ant."	"If this food pile is picked, then this ant will move to it the fastest."	
"The dots should be spacing themselves out more, but are clumping up instead."	"The ants don't like to be crowded." (3)	"The ants are like magnets that repel each other."	
	"This ant sees that the other ant is already at this food pile, and so it stays at its own food pile."		

[a]Each description was provided by a single participant, except for the phrases followed by numbers indicating the frequency of the phrase within the condition.

already there." These three descriptions were combined together in Table 2 and labeled by one of the participants' actual phrase.

The clearest trend is that participants in the concrete conditions were more likely to give domain-specific, arthropocentric interpretations of the ants' behavior. Only in the concrete condition did participants describe ants as scaring each other away, avoiding crowds, being tempted by food, or being tired. A judge tallied the number of domain-specific interpretations over all of the descriptions, both unique and shared. For an interpretation to count as domain-specific it needed to be (a) applicable to sentient agents like ants but not nonsentient agents like the categories of the second simulation, (b) described in terms of "ants" or "food" rather than more abstract language such as "forager" or "groups," and (c) not simply a specific instantiation using ants and food of the abstract rules that in fact governed the ants' behavior. With these criteria, for the concrete, idealized, concreteness fading, and concreteness introduction conditions, the number of domain-specific descriptions were 18, 4, 6, and 8, respectively.

A second textual analysis was conducted on the abstractness of the terms used to describe the ants and food simulation. For example, the food resources were variously described by different participants as "resources," "blobs," "clumps," "splotches," "supply," "edible stuff," "food," "fruit," "apples," and "oranges." The first five terms were classified as abstract because they did not directly invoke edibility, and the last five terms were classified as concrete because they do presume edibility. Similarly, ants were most often called "ants" by participants but were occasionally described using more abstract language that was not tied to the ants context: "thing," "mover," "dot," "spot," and "forager." Restricting our attention to only verbal descriptions of ants and food, the percentage of abstract descriptions for concrete, idealized, concreteness fading, and concreteness introduction conditions were 28%, 59%, 42%, and 34%, respectively. These values indicate greater use of abstract language for the idealized than concrete condition, with the remaining two conditions falling between these extremes but closer to the concrete condition.

Discussion

The results of Experiment 2 closely matched those from Experiment 1 in three important respects. First, the concreteness fading condition resulted in the best performance on both the initial and transfer simulations. Second, there were initial and transfer simulation advantages for conditions with variable, rather than consistent, graphical elements. Third, there was an interaction between the two consistent appearance conditions (idealized vs. concrete) and simulation (initial vs. transfer), such that concrete elements in a simulation tended to produce better performance on the simulation itself compared to idealized elements, but the converse ordinal relation was found for the transfer simulation. Neither of these ordinal relations was significant by itself, but the interaction was significant. This interaction

indicates that simulation appearance differentially affects comprehension of the simulation and transfer from the simulation. One likely account for the interaction is that a concrete simulation helps learners develop a strong representation of the simulation itself, but the representation may be too tied to the specific simulation to foster robust transfer.

The results from the problem-based measure of transfer were consistent with the quiz measure. The manipulations of concreteness affected not only verbal knowledge that was probed through written quizzes, but also practical knowledge of how to solve problems.

GENERAL DISCUSSION

Experiments 1 and 2 provide converging evidence that the choice of graphical elements in a simulation influences transfer from the simulation to another abstractly related but superficially dissimilar simulation. Taken together, the experiments indicated four important results. First, both initial simulation and transfer performance was better when a simulation was presented with elements of variable, rather than uniform, appearance. Second, performance was particularly good when originally concrete elements were switched to more idealized representations halfway into a simulation. This concreteness fading method resulted in reliably better performance than the converse method of concreteness introduction on both initial and transfer simulations. Third, concrete displays were superior to idealized displays for the trained simulation itself, but the opposite trend was found for transfer to an abstractly related simulation. Fourth, these trends were found for both a quiz-based measure of comprehension and a performance-based measure. These results have a bearing on the relation between concrete and abstract thought and also on practical issues for the design of pedagogically motivated simulations.

Concrete and Abstract Representations

The effect of graphical appearance in both experiments suggests competing advantages for concrete and abstract representations. Concrete displays are advantageous because they provide a strong and intuitive link between the elements of the modeling world and the elements of the modeled world. Even though all participants were explicitly told what the elements in a simulation represented, this knowledge is likely to be stronger when the superficial appearance of an element reflects its meaning. In describing cognitive advantages conferred by graphical representations, Scaife and Rogers (1996) noted that some representations allow "computational offloading"—they reduce cognitive effort by offering external information sources. A concrete representation of an ant that looks like an ant affords computational offloading because it does not require learners to keep in their mind

the correspondences between modeling and modeled worlds. Likewise, Larkin and Simon (1987) argued that good external representations of a problem make explicit aspects that would otherwise need to be internally computed.

The converse advantage of idealized displays is that they promote internal representations that are not tied to a single domain. Considerable research has documented the concrete nature of human thought (Johnson-Laird et al., 1972; Nisbett & Ross, 1980). With respect to learning specifically, the "situated cognition" movement emphasizes that knowledge is highly dependent on the particular context (Lave, 1988). Oftentimes, concreteness blinds people to the deep commonalities between situations (Gentner & Toupin, 1986; Gick & Holyoak, 1980; Holyoak & Koh, 1987; Ross, 1987). People frequently are reminded of situations on the basis of superficial rather than abstract similarities (Gick & Holyoak, 1983; Reed, Ernst, & Banerji, 1974; Reeves & Weisberg, 1994; Ross, 1989). Finding methods that foster more abstract remindings is challenging, but important and possible (Barnett & Ceci, 2002). One such method is to present relatively idealized, decontextualized elements. Bassok and Holyoak (1989) found that algebraic formulae were learned in more generalizable forms when taught using a relatively abstract algebra context rather than a more concrete physics context.

Consistent with the risk of concreteness producing overly context-tied representations, participants exploring earlier versions of these simulations were more likely to think about domain-general principles of competition between agents trying to collect resources in idealized than concrete graphics conditions. For example, learners in the concrete condition were more likely than learners in the idealized condition to describe the ants as "getting tired," "being happy with their share of food," "seeing another ant already occupying a food patch," and "scaring other ants away." Several of these descriptions were partially correlated with the true rule-governed behavior of the ants, and so participants may have been able to answer some of the ants and food quiz questions correctly using these arthropocentric interpretations. However, these descriptions are all specific to ants and consequently would not transfer very effectively to new competitive specialization situations.

The good performance obtained with concreteness fading makes sense from the perspective of these competing advantages of concrete and idealized representations. Concreteness fading was proposed as a promising instructional method because it allows simulation elements to be both intuitively connected to their intended interpretations but also idealized in a manner that promotes transfer. Concreteness introduction might be expected to have both of these properties as well, and so it is useful to assess why concreteness fading was superior to concreteness introduction. Part of the answer is likely that providing an initial concrete grounding is more helpful in promoting understanding than providing this grounding later. During the interviews, a few participants in the concreteness fading condition spontaneously reported that having the ants switch to dots was not disruptive because they could easily continue to interpret the dots properly. One participant

said, "I just treated the dots and ants as the same thing, and this was easy because they were both black." If participants are successful in naturally interpreting black dots as ants when they are preceded by ants, then one of the primary disadvantages of decontextualized representations is significantly mitigated. People's natural tendency to interpret ambiguous objects so as to be consistent with prior, unambiguous objects (Leeper, 1935; Medin, Goldstone, & Gentner, 1993) may allow idealized objects to continue to be interpreted as they were when they were concrete.

Schwartz (1995; Schwartz & Black, 1996a, 1996b) presented a related theoretical treatment of the complementary advantages of concrete and abstract representations. By this account, idealized, low-fidelity materials encourage people to think about the materials as referents. Concrete, high-fidelity materials encourage people to reason about the referent itself and consequently cause people to employ solution processes that have real-world analogs. These analog processes, like rotating and pulling, are intuitive and perceptually scaffolded, but more abstract construals are likely to be more general and efficient for complex problems. Schwartz speculated that

> One might use an actual image of a person sliding down an inclined plane to help the student think in terms of forces as they are perceptually experienced. After these perceptual notions are brought to mind and reflected upon, one may want to switch to more abstract drawings of two dimensional blocks on oblique lines. (p. 721)

This possibility is supported by our concreteness fading results.

The positive impact of concreteness fading is reminiscent of Kotovsky and Gentner's (1996) work on fostering abstract relational responses in children. Perceptual supports helped children grasp the relational similarities, and, once grasped, these supports can be removed without disrupting relational responses (see also Mix, 1999, for an example in which the abstraction involves numeric quantity). Similarly, concreteness fading allowed learners to first acquire interpretations of simulation elements that were supported by both concrete appearance and relational role. Then, when the concrete cues were removed, learners could continue to respond to the abstract roles. Together with Terrace's (1963b) original working on fading, our results indicate that an effective learning technique is to provide perceptual scaffolding that leads to responding that is consistent with an important cue. When the perceptual scaffolding is removed, responding to the important cue may continue without the learners making costly interpretational errors.

CONCLUSIONS

Caution is needed to avoid overinterpreting these experimental results, because a purposefully minimal manipulation of concreteness was employed. Still, some po-

tential recommendations can be offered. A first recommendation is to present simulation materials with varied rather than consistent appearances (see also Bruner, 1966; Gick & Holyoak, 1980, 1983; Paas & Van Merrienboer, 1994). A second pedagogical suggestion based on our results is that two educational uses of a simulation should be distinguished. Some scientific simulations are designed to teach students about the nature of a particular domain or situation such as heat flow, evolution, or infection spread. Other simulations are designed to teach students principles that are applicable in many domains other than the specific one presented (White, 1993; Wilensky & Resnick, 1999). These two uses must be distinguished, because different design choices are prescribed depending on which is the intended use of a simulation. In this respect, our results are consistent with Bransford and Schwartz's (1999) observation that instruction that leads to the best immediate performance is not always the same as instruction that prepares learners ideally for future learning opportunities. In our experiments, performance on the simulation itself was better with concrete than idealized graphics, but the opposite trend was found for transfer to an abstractly related simulation. The third, most obvious recommendation is to incorporate variants of concreteness fading in pedagogical simulations. Concreteness fading not only improved performance on the simulation that used this method, but it also promoted the most robust transfer to a new simulation governed by the same principle.

In recent years, the pedagogical use of computer simulations has grown dramatically and shows little sign of abating. They are powerful teaching tools because they provide a concrete grounding for concepts that might otherwise be too abstract to be readily comprehended. They promote an active, hands-on, problem-solving stance by learners (National Research Council, 1999). Although there is a considerable body of research that explores and tests design choices that could impact the usefulness of simulations (Jackson, Stratford, Krajcik, & Soloway, 1996; Klahr & Carver, 1988; Miller et al., 1999; Miller & Stigler, 1991; White & Frederiksen, 1998), there is still a deep need for research that explores systematic differences in simulations to try to optimize their educational impact and to assess what elements of the simulation are critical for imparting the educational benefit. Our research attempts to address one component of this long-term pursuit.

Experiments revealed that a combination of both concrete and idealized formats was valuable and that the particular sequence that was most valuable was to have initially concrete representations become more idealized over time. Consistent with the advantages of concrete representations described in Table 1, we believe, along with many others, that computer simulations are effective pedagogical devices precisely because of their concreteness and perceptual grounding. However, we are also interested in students applying what they have learned to domains that are superficially unrelated to the simulation's domain. Our results give us optimism that these motivations are not necessarily mutually exclusive and that both perceptually grounded and abstract understandings can be simultaneously achieved.

ACKNOWLEDGMENTS

This research was funded by National Institutes of Health Grant MH56871 and National Science Foundation Grant 0125287.

Many useful comments and suggestions were provided by Dedre Gentner, Ken Koedinger, Douglas Medin, Kelly Mix, Zack Patton, Daniel Schwartz, Vladimir Sloutsky, and Linda Smith. We thank Melissa Adkins, Adam Fritz, and Brent Myers for assistance in conducting the experiments.

REFERENCES

Ambrose, R. C. (2002). Are we overemphasizing manipulatives in the primary grades to the detriment of girls? *Teaching Children Mathematics, 9,* 16–21.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Journal of Experimental Psychology: General, 128,* 612–637.

Bassok, M. (1996). Using content to interpret structure: Effects on analogical transfer. *Current Directions in Psychological Science, 5,* 54–58.

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

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999*). How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (pp. 61–101).

Bruner, J. S. (1966). *Toward a theory of instruction.* Cambridge, MA: Belknap.

Carraher, D., & Schliemann, A. D. (2002). The transfer dilemma. *The Journal of the Learning Sciences, 11,* 1–24.

Clements, D. H., & McMillen, S. (1996). Rethinking "concrete" manipulatives. *Teaching Children Mathematics, 2,* 270–279.

DeLoache, J. S. (1991). Symbolic functioning in very young children: Understanding of pictures and models. *Child Development, 62,* 736–752.

DeLoache, J. S. (1995). Early understanding and use of symbols: The model. *Current Directions in Psychological Science, 4,* 109–113.

DeLoache, J. S. (2000). Dual representation and young children'*s* use of scale models. *Child Development, 71,* 329–338.

DeLoache, J. S., & Burns, N. M. (1994). Symbolic functioning in preschool children. *Journal of Applied Developmental Psychology, 15,* 513–527.

DeLoache, J. S., & Marzolf, D. P. (1992). When a picture is not worth a thousand words: Young children's understanding of pictures and models. *Cognitive Development, 7,* 317–329.

DiFonzo, N., Hantula, D. A., & Bordia, P. (1998). Microworlds for experimental research: Having your (control and collections) cake, and realism too. *Behavioral Research Methods, Instruments, and Computers, 30,* 278–286.

Dixon, J. A., & Bangert, A. S. (2004). On the spontaneous discovery of a mathematical relation during problem solving. *Cognitive Science, 28,* 433–449.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structure.* Dordrecht, The Netherlands: Kluwer Academic.

Fuson, K. C., & Briars, D. J. (1990). Using a base-ten blocks learning/teaching approach for first and second grade place-value and multidigit addition and subtraction. *Journal for Research in Mathematics Education, 21,* 180–206.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity, analogy, and thought* (pp. 199–241). New York: Cambridge University Press.

Gentner, D., & Medina, J. (1998). Similarity and development of rules. *Cognition, 65,* 263–297.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10,* 277–300.

Gentner, D., & Wolff, P. (2000). Metaphor and knowledge change. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 294–342). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Gianutsos, R. (1994). Driving advisement with the Elemental Driving Simulator (EDS): When less suffices. *Behavioral Research Methods*, *Instruments, and Computers, 26,* 183–186.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–39.

Goldstone, R. L. (2003). Learning to perceive while perceiving to learn. In R. Kimchi, M. Behrmann, & C. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 233–278). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Goldstone, R. L., & Barsalou, L. (1998). Reuniting perception and conception. *Cognition, 65,* 231–262.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology, 23,* 222–264.

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46,* 414–466.

Goldstone, R. L., Steyvers, M., & Rogosky, B. J. (2003). Conceptual interrelatedness and caricatures. *Memory & Cognition, 31,* 169–180.

Grady, S. M. (1998). *Virtual reality: Computers mimic the physical world.* New York: Facts on File.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23,* 121–134.

Heim, M. (2000). *Virtual realism.* New York: Oxford University Press.

Holyoak, K. J. (1984). Analogical thinking and human intelligence. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2, pp. 199–230). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332–340.

Jackson, S. , Stratford, J., Krajcik, S., & Soloway, E. (1996). Making system dynamics modeling accessible to pre-college science students. *Interactive Learning Environments, 4,* 233–257.

Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and a sense of reality. *British Journal of Psychology, 63,* 395–400.

Judd, C. H. (1908). The relation of special training to general intelligence. *Educational Review, 36,* 28–42.

Katona, G. (1940). *Organizing and memorizing.* New York: Columbia University Press.

Kennedy, L. M., & Tipps, S. (1994). *Guiding children's learning of mathematics* (7th ed.). Belmont, CA: Wadsworth.

Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a Logo debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology, 20,* 362–404.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science, 14,* 511–550.

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments, 5,* 161–180.

Kohonen, T. (1995). *Self-organizing maps.* Berlin: Springer-Verlag.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67,* 2797–2822.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11,* 65–100.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life.* New York: Cambridge University Press.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431–467.

Leeper, R. W. (1935). A study of a neglected portion of the field of learning—The development of sensory organization. *Journal of Genetic Psychology, 46,* 41–75.

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431–467.

Martinez, J. G. R. (1987). Preventing math anxiety: A prescription. *Academic Therapy, 23,* 117–125.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278.

Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science, 23,* 305–336.

Miller, K. F., & Stigler, J. W. (1991). Meanings of skill: Effects of abacus expertise on number representation. *Cognition and Instruction, 8,* 29–67.

Mix, K. S. (1999). Similarity and numerical equivalence: Appearances counts. *Cognitive Development, 14,* 269–297.

Moch, P. L. (2001). Manipulatives work! *The Educational Forum, 66,* 81–87.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching Children Mathematics, 8,* 372–377.

Nathan, M. J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments, 5,* 135–159.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra word problem comprehension and its implications for the design of computer learning environments. *Cognition and Instruction, 9,* 329–389.

National Research Council. (1999*). How people learn: Brain, mind, experience, and school.* Washington, DC: National Academy Press.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment.* Englewood Cliffs, NJ: Prentice-Hall.

Novick, L. R., & Holyoak, K. J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 398–415.

O'Reilly, R. C. (2001). Generalization in interactive networks: The benefits of inhibitory competition and Hebbian learning. *Neural Computation, 13,* 1199–1242.

Paas, F., & Van Merrienboer, J. (1994). Variability of worked examples and transfer of geometry problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86,* 122–133.

Rattermann, M. J., & Gentner, D. (1998). The effect of language on similarity: The use of relational labels improves young children's performance in a mapping task. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences* (pp. 274–282). Sophia: New Bulgarian University.

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology, 6,* 436–450.

Reeves, L. M., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin, 115,* 381–400.

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

Resnick, M. R. (1994). *Turtles, termites, and traffic jams.* Cambridge, MA: MIT Press.

Resnick, M. R., & Wilensky, U. (1993, April). *Beyond the deterministic, centralized mindsets: New thinking for new sciences.* Paper presented at the meeting of the American Educational Research Association, Atlanta.

Resnick, M. R., & Wilensky, U. (1998). Diving into complexity: Developing probabilistic decentralized thinking through role-playing activities. *Journal of Learning Sciences, 7,* 153–172.

Reyes, R. M., Thompson, W. C., & Bower, G. H. (1980). Judgmental biases resulting from differing availability of arguments. *Journal of Personality and Social Psychology, 39,* 2–12.

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,* 629–639.

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,* 456–468.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rumelhart, D. E., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science, 9,* 75–112.

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human–Computer Studies, 45,* 185–213.

Schank, R. C., & Farrel, R. (1988). Creativity in education: A standard for computer-based teaching. *Machine-Mediated Learning, 2,* 175–194.

Schunn, C. D., & Dunbar, K. (1996). Primary, analogy, and awareness in complex reasoning. *Memory & Cognition, 24,* 271–284.

Schwartz, D. L. (1995). Reasoning about the referent of a picture versus reasoning about the picture as the referent: An effect of visual realism. *Memory & Cognition, 23,* 709–722.

Schwartz, D. L., & Black, J. B. (1996a). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology, 30,* 154–219.

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

Schwartz, D. L., & Moore, J. L. (1998). The role of mathematics in explaining the material world: Mental models for proportional reasoning. *Cognitive Science, 22,* 471–516.

Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education, 20,* 498–505.

Smith, L. B. (2003). Learning to recognize objects. *Psychological Science, 14,* 244–250.

Stephens, D. W., & Krebs, J. R. (1987). *Foraging theory.* New Haven, CT: Princeton University Press.

Terrace, H. S. (1963a). Discrimination learning with and without "errors." *Journal of the Experimental Analysis of Behavior, 6,* 1–27.

Terrace, H. S. (1963b). Errorless transfer of a discrimination across two continua. *Journal of the Experimental Analysis of Behavior, 6,* 223–232.

Uttal, D. H., Liu, L. L., & DeLoache, J. S. (1999). Taking a hard look at concreteness: Do concrete objects help young children learn symbolic relations? In C. S. Tamis-LeMonda (Ed.), *Child psychology: A handbook of contemporary issues* (pp. 177–192). Philadelphia, PA: Psychology Press.

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.

von der Malsburg, C. (1973). Self-organization or orientation-sensitive cells in the striate cortex. *Kybernetik, 15,* 85–100.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Psychology, 23,* 63–71.

White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction, 10,* 1–100.

White, B. Y., & Fredericksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16,* 3–118.

Wilensky, U. (1991). Abstract meditations on the concrete and concrete implications for mathematics education. In I. Harel, & S. Papert (Eds.), *Constructionism*. (pp. 193–203). Westport, CT: Ablex.

Wilensky, U. (1999). *NetLogo.* Evanston, IL. Center for Connected Learning and Computer Based Modeling, Northwestern University.

Wilensky, U. (2001). Modeling nature's emergent patterns with multi-agent languages. *Proceedings of EuroLogo 2001.* Linz, Austria.

Wilensky, U., & Reisman, K. (1999). ConnectedScience: Learning biology through constructing and testing computational theories—An embodied modeling approach. *InterJournal of Complex Systems, 234,* 1–12.

Wilensky, U. & Resnick, M. (1999). Thinking in levels: A dynamic systems perspective to making sense of the world. *Journal of Science Education and Technology, 8,* 3–18.

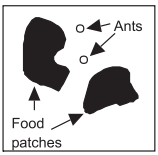
APPENDIX A
Instructions to Participants

The instructions given to participants in the ants and food simulation were:

To explore the simulation, try the following:

1. Draw different food patches that will attract ants.
2. Change the parameters of the simulation: how many ants there are, how fast the closest ant to a food patch (the "winner") moves toward the food patch, and how fast the other ("loser") ants move toward the food patch.
3. As the ants move around, you can pick them up, move them elsewhere, add new ants, and change any parameters.

In trying to understand this simulation, the question that you should keep in the back of your mind at all times is: How can you change the parameters so that each piece of food has an ant that is close to it? That is, how can you make it so that the food is optimally covered by ants? For example, think about how a situation that starts off with a display like



could eventually look like

When thinking about the simulation, remember that there are no hidden complexities to the ants' behavior. Everything the ants do can be explained by the following three rules:

1. One at a time, a piece of food (a green dot) is randomly selected from all the food present.
2. The ant that is closest to the selected food (the "winner") moves toward the food with a speed that you specify.
3. All of the other ants (the "losers") move toward the food with another speed."

The instructions for the pattern learning simulation were:

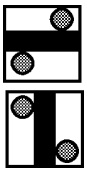
To explore the simulation, try the following:

1. Draw different pictures for the computer to use in forming its categories.
2. Change the parameters of the simulation: how many categories and pictures there are, how quickly a picture changes its most similar category, and how quickly the picture changes the other categories.
3. As the categories are being learned, you can change the pictures or categories by drawing on them, or by copying one picture to another.

In trying to understand this simulation, the question that you should keep in the back of your mind at all times is: How can you change the parameters so that the computer learns good categories for the pictures? That is, if there are several pictures that look similar, they should all be grouped in the same category, and pictures that look different should be grouped in different categories. Also, each of the categories should become specialized for some picture(s).

For example, think about how a set of four pictures that look like:

could create 2 categories like:



When thinking about the simulation, remember that there are no hidden complexities to the learning. Everything about how the categories move around can be explained by the following three rules:

1. The user tells the computer how many pictures and categories to allow.
2. One at a time, a picture is randomly selected.

3. The category that is closest to the picture changes itself by an amount that you specify so that it looks more like the selected picture. All of the other categories adapt themselves to the picture by a different amount.

APPENDIX B
Sample Multiple-Choice Quiz Questions for Simulations

Correct answers are indicated with asterisks. Rather than showing all seven questions for each simulation, two questions are shown with their analogs and two questions are shown without their analogs.

Ants and food

To make the ants as a population cover the food well, which strategy is the best:

1. Have the ants move as quickly as possible.
*2. Make the ant that is closest to a piece of food move more quickly than all the other ants.
3. Make the ant that is closest to a piece of food move more slowly than all the other ants.
4. Early on, make the closest ant move more quickly than the others, but later on, make the closest ant move more slowly.

Why don't the ants cover the food well if only the ant closest to a piece of food moves?

*1. If only the closest ant moves, then this ant may be responsible for many pieces of food, and the other ants may not cover any food.
2. If only the closest ant moves, then it may eventually get tired and stop moving at all.
3. The ant closest to a piece of food shouldn't move much because it is already close to the food. It is the other ants that need to move.
4. If only one ant moves, then as a population, the ants are not moving very much when a piece of food is selected.

Why don't the ants cover the food well if the closest ant and all of the other ants all move with the same speed?

1. The closest ant doesn't need to move as quickly as the other ants because it is already close to the food.

*2. If other ants move as fast as the closest ant, then when a new piece of food is selected, they will move away from food that they were previously covering well.
3. If all of the ants move with the same speed, then they will all get to the food at the same time, and they won't be able to share it efficiently if there are a lot of ants.
4. If the closest ant moves as fast as the other ants, then it will get to the food first and will prevent the other ants from benefiting from it.

To have the ants cover the food well, it is necessary to have the ants become specialized for particular food patches. Which action most directly allows for this specialization?

1. Make sure that there are not very many ants on the field. That way, each ant can be far away from other ants.
2. Make sure that there are many ants on the field. That way, each ant can become specialized for a tiny patch.
*3. Make the ants that are not closest to a piece of food move slowly to the food.
4. Make the ant that is closest to a piece of food move slowly to the food.

Pattern Learn

To make categories best represent the natural groups in a set of pictures, you should:

1. Have the categories adapt as quickly as possible.
*2. Make the category that is closest to a selected picture adapt more quickly than all the other categories.
3. Make the category that is closest to a selected picture adapt more slowly than all the other categories.
4. Early on, make the closest category adapt itself more quickly than the others, but later on, make the closest category adapt more slowly.

Why aren't good categories formed if only the picture closest to a selected picture adapts?

*1. If only the closest category adapts, then this category will become responsible for many pictures, and the other categories may not be adapted for any picture.
2. If only the closest category adapts, then its learning may eventually become exhausted, and it may stop learning at all.

3. The picture closest to a picture shouldn't adapt much because it is already close to the picture. It is the other pictures that need to adapt.
4. If only one category adapts, then in general the categories are not changing very much when a picture is selected.

If there are four pictures and two categories, the categories will emphasize the parts that the pictures in a category share. How does this occur?

1. Categories adapt most toward parts shared by category members because shared parts provide links between pictures belonging to different categories.
2. Categories adapt most toward parts shared by category members because the rate of adaptation will be faster for categories that are not closest to the selected picture.

*3. Categories adapt most toward parts shared by category members because these parts are always present in the members, and so there will be more opportunities for learning.

4. Categories adapt most toward parts shared by category members because these parts are at the very essence of the category, defining what it means to be part of the category.

If there are two pictures and only one category, what usually happens?

1. The category will alternate between the pictures, but only if it adapts very slowly.

*2. The category will be a blend of the two pictures, highlighting parts shared by the pictures.

3. The category will become specialized for one of the pictures only.
4. The category will not become adapted to either picture, unless it is highly similar to them in the first place.