Transfer, and the Effects of Context Outside of the Training Task

Samuel B. Day (day9@indiana.edu) **Sarah Manlove (smanlove@indiana.edu)**

Robert L. Goldstone (rgoldsto@indiana.edu)

Dept. of Brain & Behavioral Sciences, 1001 E. 10th St.
Bloomington, IN 47405 USA

Abstract

While the use of concrete, contextualized and personally relevant examples can benefit learners in terms of comprehension and motivation, these types of examples can come with a cost. Examples may become too bound to their particular context, and individuals may have a difficult time recognizing when the underlying principles are relevant in new situations. In the current study, we provide evidence that contextualization may impair knowledge transfer even when that context occurs outside of the training example itself. Specifically, when students were taught about positive feedback systems in the context of polar ice-albedo effects, those individuals that had previously learned about the effects of global warming on polar bear populations showed reliably poorer transfer performance.

Introduction

In virtually all educational domains, the ultimate goal of learning is not simply to acquire some static body of specific facts. Rather, the objective is to gain more general kinds of knowledge structures that can be applied in new situations and under novel conditions. For example, computer programmers may learn very general algorithms that can be used in a variety of different tasks and instantiated in very different programming languages; students of literature or history learn about themes and patterns that can occur across a wide variety of situations that are superficially dissimilar; and those learning about science may discover principles that are relevant not only across different contexts, but even across disciplines. In short, the goal of learning is primarily to acquire generalizable knowledge that may be used productively.

Unfortunately, it is not immediately clear what the most effective means of conveying this kind of knowledge might be, and the data on this topic can seem counterintuitive or contradictory. This is a critical area to understand, however. Instructors have a great deal of latitude in their selection of teaching examples and methods, and research suggests that even subtle differences in these choices may have an important impact on students' learning. The current study extends our understanding of this issue. Specifically, we explore the ways in which the context surrounding a particular training example may influence what students learn, and their ability to apply this knowledge to new cases.

One straightforward approach of conveying generalizable knowledge is to present information in a highly abstracted way, removing any context-specific details and features. For example, consider the concept of *positive feedback systems*. Such systems are ubiquitous in science, and can be instantiated in an almost limitless number of ways. In order to maximize the set of situations where a learner's knowledge can be applied, this concept could be presented in a way that is not specific to any particular context, such as: "A system in which increases to a variable cause still further increases to that variable." Such a general definition would make the concept applicable to a wide range of domains, capturing positive feedback phenomena in biology, physics, chemistry and even interpersonal interactions.

However, it has long been argued that while such abstract presentations might capture the relevant information efficiently, they do so at the expense of comprehensibility (e.g., Bruner, 1966), making this approach ultimately counterproductive. Learners cannot apply information that they do not understand. Consistent with this, research has found that people tend to rely on more concrete examples when possible. For example, in one study (LeFevre & Dixon, 1986) researchers provided participants with explicit verbal instructions on how to perform a task, while also giving them a concrete example of the task being performed. However, for some individuals, those sources of information conflicted with each other, and actually reflected different tasks and goals. Under those conditions, participants were overwhelmingly more likely to act on the basis of the concrete example rather than the more abstract verbal instructions. Similarly, Ross (1987) found that even when students were given the appropriate mathematical formula to use in solving a story problem, their performance was influenced by the concrete examples they had previously seen (see also Anderson, Farrell, & Sauers, 1984).

One of the most striking examples of the merits of contextualization is the Wason selection task, which leads to uniformly poor performance (typically about 10% correct) when presented in its abstract symbolic form, but is often solved when instantiated in a familiar context (e.g., Johnson-Laird, Legrenzi, & Legrenzi, 1972; Wason & Shapiro, 1971). As we will discuss, however, there are issues with learning from contextualized training examples as well.

First, it can be challenging to define exactly what contextualization means. At one end of the spectrum, it could simply refer to the concrete perceptual features that are associated with a situation. For example, instructors could teach a principle with an animated simulation, using more or less realistic perceptual features. Another way of construing contextualization is in terms of the learner's existing knowledge structures. A case may be considered contextualized to the extent that it draws on familiar schemas, in which many of the relationships are already known. Alternatively, contextualization could be a function of personal relevance or perspective, with more engaging or interactive tasks possibly leading to deeper kinds of understanding. Of course, in most situations, these factors are probably highly interrelated.

Perhaps surprisingly, for each of these ways of construing contextualization, there is at least some empirical evidence suggesting that greater contextualization can impair people's learning. For instance, when Goldstone and Sakamoto (2003) examined students' ability to learn and transfer a scientific principle from a perceptually concrete computer simulation, they found superior performance when the simulations used relatively less detailed or less realistic entities (e.g., portraying ants as dots, as opposed to more realistic ant animations; also see Kaminski, Sloutsky & Heckler, 2008). Likewise, while the activation of schemas can sometimes support performance (as in some versions of the Wason selection task), schemas may also be detrimental if they suggest irrelevant or inappropriate relationships (e.g., Bassok, Wu & Olseth, 1995). And while personal interaction and personal relevance has been argued to support learning both cognitively (e.g., McCombs & Whistler, 1997) and in terms of motivation (e.g., Lepper, 1988), research has also called these potential benefits into question. For instance, Son and Goldstone (2009) found that when participants were taught the principles underlying signal detection theory through a concrete training example, their performance was impaired when the task was made more personally relevant, either by giving participants first-hand detection experience, or by framing the task in a first-person perspective (e.g., "Imagine that you are a doctor…" vs. "Imagine a doctor…"). Similarly, DeLoache (2000) found that young children's ability to use materials as symbolic representations was impaired after they were given the chance to directly manipulate and play with them.

Findings such as these raise some serious cognitive and pedagogical questions. For example, teachers are frequently told of the benefits of using concrete examples in their classes, and of making instruction engaging and relevant to the students (e.g., Rivet & Krajcik, 2008). The empirical research, however, suggests that the picture may be more complex than that advice would suggest. Injudicious use of contextualization and personalization in the classroom could actually hurt students' performance under some circumstances, particularly if performance is measured in terms of transfer to new situations. Furthermore, the fairly broad scope of the factors that could count as contextualization opens the possibility that seemingly subtle differences in the way that an example is described or introduced could have important consequences for learning. In the current research, we investigate the possible effects of these subtle kinds of contextualization.

Experiment

In the previous research discussed thus far, context has been manipulated by directly altering perceptual or conceptual aspects of the training task itself. In the current study, we examined the effects of manipulating context less directly. While the task itself, and even the introductory description of the task, were identical between conditions, the preceding introduction to the general content domain differed. Specifically, one condition described information that was expected both to be associated with more background knowledge and to be more personally relevant and engaging to the students.

Participants

144 students from a public middle school participated in this study, as part of their regular class time in a General Science course. The group included both 7th- and 8th-grade students (n = 70 and 74, respectively) from six class periods. A little more than a third of the students (n = 49) were part of the school's Accelerated Learning Program (ALPs), which is composed of students passing a science achievement test. The students were roughly evenly divided between males (n = 68) and females (n = 76).

Materials and Design

Our experiment was conducted during the course of regular class periods in a public middle school. Students first completed a pretest, in which they read several brief scenarios and decided whether each was an example of a positive feedback system. The instructions for this test included a brief definition of positive feedback, along with an example. Students then read a short introduction to the topic of polar melting. The wording of this introduction varied between participants in terms of its contextual richness (High Context vs. Low Context), and this variation was the only difference between the experimental conditions. All students then interacted with a computer simulation of the behavior of the polar ice caps. Next, students responded to an open-ended item, asking them to write a short paragraph describing positive feedback systems in general. Finally, the pretest feedback scenarios were administered again as a posttest.

Pretest and Posttest. The pretest and posttest materials were designed to assess students' understanding of positive feedback systems. The materials included eight brief scenarios (averaging 48 words apiece), each describing a real-world phenomenon. Half of these scenarios represented positive feedback systems and half did not. For example, one scenario was the following:

> Economic inflation involves a complex set of factors. Here is an example scenario. Minimum wage is increased; therefore the cost of producing goods increases; this causes a rise in the price of the goods; this in turn increases the cost of living,

leading to a call for an increase in the minimum wage.

Participants were then asked whether or not the relationship between the relevant factors represented a positive feedback system. Responses were given by selecting one of the following options from a five-point rating scale: *Definitely not, Probably not, Don't know, Probably yes,* and *Definitely yes*.

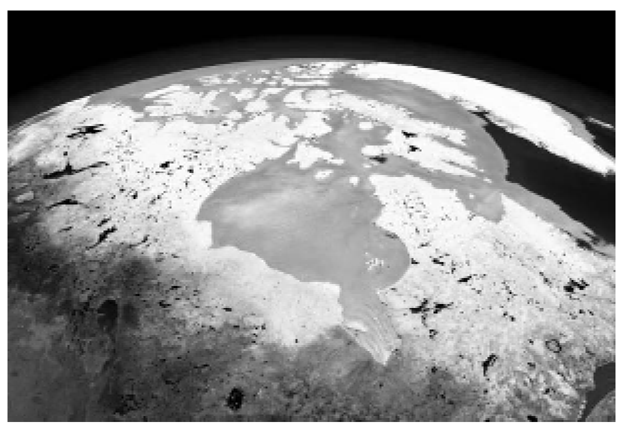
Identical items were given at pre-test and post-test. However, in order to minimize any explicit memorization and reference to previous answers, students were not informed about the post-test until later in the experimental session.

Computer simulation. All students interacted with a computer simulation demonstrating the effects of ice-albedo feedback (a kind of positive feedback system) on the Earth's polar ice caps. Prior to interacting with the simulation, each participant read a one-paragraph introduction to the topic of polar melting resulting from global warming (see Box 1). For roughly half of the students (n = 75), this introductory paragraph described recent patterns of polar melting in the Hudson Bay area (the Low Context condition). While this introduction is directly relevant to the topic of the simulation, it was expected to provide little in the way of subjective context because of students' limited background knowledge of, and personal relationship to, this specific content. The remainder of the students (n = 69) read an introduction describing the negative impact of global warming on polar bear populations (the High Context condition). We expected students to have fairly rich existing knowledge about polar bears, and to have a greater sense of personal relevance and identification with the plight of the bears. The full texts of these introductions, as well as the photographs accompanying each, are given in Box 1.

Next, all students read a more specific introduction to ice-albedo feedback effects:

> This is a simulation of the polar ice caps. Heat from the sun warms the earth, and can melt the ice.
>
> One interesting thing about the polar caps is that because ice is white, it REFLECTS much of the sunlight, so it isn't absorbed. Because of this, when some ice melts, less sunlight is reflected, so the earth gets warmer, which makes even MORE ice melt. On the other hand, when some water freezes, more sunlight is reflected, making the earth cooler, which can make even ***more*** water freeze.
>
> This kind of system is called "positive feedback." Any change in the system (like water freezing) tends to cause even more of that change (like more water freezing).

Polar bears are in danger. Climate research now shows that because of global warming, Canada's Hudson Bay sea-ice forms later in the winter, and is breaking up earlier in the spring than in the past. This shortage of sea ice leaves the population of polar bears there with more time waiting for ice to form, and less time on the ice. The less time they have on the ice, the less food they have. Currently the polar bears have been off the ice since July 15 and must rely on fat reserves which they lose quickly. The longer they wait the more at-risk they are of not being strong enough to hunt when they can. If polar bears cannot hunt, they will not survive.



Hudson Bay's sea-ice in Canada takes a long time to form during the colder months. It usually starts forming in October, and has full ice-cover by December. The bay's sea ice-extent and thickness is studied to determine the effects of global warming. "Ice-extent" is a measurement of the area of the ocean where there is at least some ice. Scientists say Hudson Bay's ice extent in the winter is much less than it used to be. One reason for the lack of ice in the bay is the warmer temperatures in the past twenty years. Given its size, history, and impact on global climate patterns, Hudson Bay's sea ice processes will continue to be very important to study as we struggle to understand global warming.

Box 1: Pre-simulation context materials. The photo and text on the left were given to those in the high context condition; those on the right were given to the low context group.

Students then interacted with the ice-albedo simulation itself (This was implemented in NetLogo, a software package for developing agent-based simulations; Wilensky, 1999). Students were guided through the simulation with very specific instructions, which were designed to highlight the feedback effects. Additionally, these instructions explicitly reiterated at multiple points that these effects were a demonstration of *positive feedback* behavior, and why. Box 2 provides a thorough description of the simulation and instructions.

After the simulation, students were asked to define positive feedback systems in their own words: "We would like you to tell us what a positive feedback system is. Just do your best to describe it in your own words. Please *don't* just write about the simulation you just saw. Instead, try to write about positive feedback systems in general." Students had a full page to provide their answers, but there were no instructions regarding the required length for this response.

Finally, all students completed the scenario classification task again as a posttest.

Results

Our analyses found reliable differences between the context conditions in terms of ability to recognize positive feedback systems in new situations. Responses on the pretest and posttest were coded according to their proximity to the correct end of the rating scale. For instance, if a scenario actually reflected positive feedback behavior, a response of *Definitely yes* would be coded as a 5, a response of *Definitely not* would be coded as a 1, and a response of *Don't know* would be coded as a 3. These codings would be reversed for scenarios that did *not* reflect positive feedback (e.g., a response of *Definitely yes* would be coded as a 1). For each student, we calculated an improvement score, which was simply the sum of the posttest scores minus the sum of the pretest scores. 11 of the 144 students were dropped from the analysis because of items left blank during one of the two tests.

We found reliable differences between the improvement scores of the low and high context conditions ($t(132) = 2.42$, $p = .017$). Specifically, those in the low context condition improved reliably at posttest ($M = 1.58$, $t(68) = 3.30$, $p = .002$), while those in the high context condition showed a non-significant decrease in posttest performance ($M = -.36$, $t(63) = 0.55$, *n.s.*). Because of the poor performance by those in the high context group, there was no reliable improvement when collapsing across all participants ($M = .65$, $t(132) = 1.59$, $p = .114$). The superiority of the low context group also held in a separate analysis of the eight test items ($t(7) = 2.47$, $p = .043$).

This poor transfer performance did not appear to be the result of less effective learning of the training example itself. Students' responses to the open-ended definition item were coded on a scale from 0 to 5, based on a rubric assessing their understanding of positive feedback. These scores did not differ between the two groups ($M = 1.99$ and 2.11 for the low and high context conditions, respectively ($t(132) = 0.48$, $p = .63$).

The effects of context condition also did not appear to vary as a function of student ability. While the accelerated (ALPS) students outperformed those in regular classes at both pretest ($t(132) = 6.99$, $t < .001$) and posttest ($t(132) = 6.43$, $t < .001$), there were no differences between the groups in terms of overall improvement or differences in improvement between context conditions. Similarly, there were no differences in the effects of condition on 7th vs. 8th graders.

We also found evidence for a small but reliable bias in students' posttest responses, such that items at posttest were more likely to be classified as examples of positive feedback. To assess this bias, we coded each response based on its proximity to the end of the rating scale labeled *Definitely yes*, regardless of what the correct response for that item should be (e.g., responses of *Definitely yes* were coded as 5, responses of *Definitely not* were coded as 1). Bias for each student was simply calculated as the sum these scores at posttest minus the sum at pretest. Across all participants, this value was reliably greater than zero ($M = .30$, $t(132) = 3.47$, $p < .001$). The level of this bias did not differ between the low ($M = .22$) and high ($M = .39$) context conditions ($t(132) = 0.99$, $p = .32$).

Discussion

Individuals in the current study were adversely affected by rich contextualization, even when that contextualization occurred outside of the training example itself. All of the students in our experiment interacted with identical computer simulations, and the descriptions of both ice-albedo effects and positive feedback systems more generally were the same across conditions. However, those individuals who had previously read a contextually rich general introduction to the issue of polar melting (involving polar bears) showed reliably poorer transfer performance. Specifically, while students who had read a less personally relevant and engaging introduction had reliable posttest gains in their ability to classify new cases as examples of positive feedback, those in the high context condition showed no gains at all.

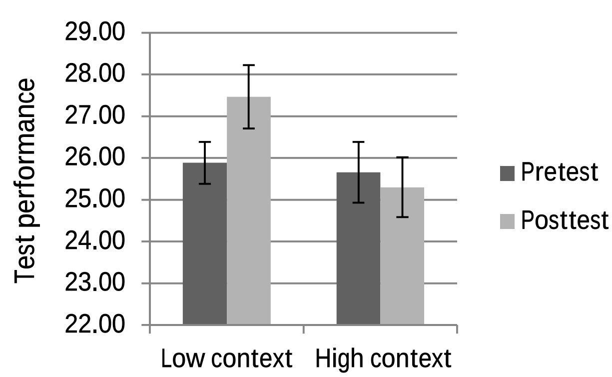
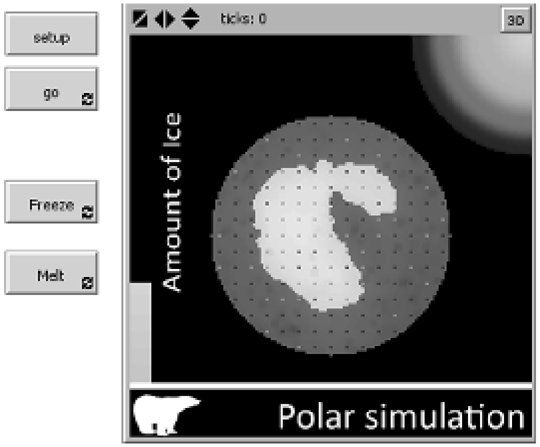


Figure 1: Pretest and posttest classification results.



The computer simulation displays a top-down view of a polar ice cap, surrounded by water. The primary observable dynamics of the system involve the size and shape of the ice surface, which is constantly changing. These changes are a function of the temperature at each location on the earth, and this temperature is affected by four different factors: cooling, diffusion, sunlight, and albedo. First, there is a slow but constant net cooling of the earth, reflecting the dissipation of heat into space. Similarly, there is a constant diffusion of heat between adjacent areas of the earth, which serves to "average out" the temperature in a given region. The most relevant factors for the students, however, are sunlight and albedo. The earth is receiving a steady flow of energy in the form of sunlight, which can be absorbed and can increase an area's temperature. However, not all of this energy is absorbed: much of the light is reflected back into space. Furthermore, the amount of light that is reflected depends on a given area's albedo or reflectance. Critically, ice has a much higher albedo than land or water, because of its white color. In our simulation, ice only absorbs one quarter of the energy that is absorbed by the surrounding water (20% vs. 80%, respectively). Because of this, greater ice coverage results in lower overall warming. It is this factor that produces the system's feedback behavior. A decrease in ice coverage results in more heat being absorbed, leading to even more melting, and so on. Conversely, an increase in ice coverage causes the reflection of more light, reducing the temperature and potentially causing even more water to freeze.

The simulation uses a grid of colored points to indicate each region's overall light reflectance and absorption (the "reflection grid"). Red dots (darker in the image above) indicate absorbed energy, while blue dots (lighter) indicate reflected energy. The actual location of these dots is irrelevant for the operation of the simulation itself, but they provide a way for students to directly perceive the relative balance of reflected and absorbed energy in different locations. Specifically, 80% of the dots on frozen areas show reflectance (blue dots), compared with 20% of the dots on the water. When active, this grid flashes on and off in one second increments.

Students were guided through the simulation via specific instructions, given through popup messages. Initially, students were familiarized with the operation of the system, first without the reflection grid, and then with. Messages appeared at brief intervals reminding them of the relevant principles of the system, such as:

> *Right now, the system of ice, water and heat is pretty balanced, and doesn't change much. One interesting thing about the polar caps is that because ice is white, it REFLECTS much of the sunlight, so that light isn't absorbed. Next, we will show you how much light is being reflected by the ice and by the water. Red dots show light and heat that are being ABSORBED by the earth. Blue dots show light and heat that are being REFLECTED away from the earth.*

Next, students were instructed to interact with the system in various ways. For example, students were asked to select the button labeled "Melt" and to click and drag a few lines through the ice. As they did so, the ice under the cursor changed to water (with a temperature of 36° F). After a sufficient amount of the ice had been melted in this way, the simulation resumed. At this point, the reduced albedo led to a positive feedback loop in which additional ice melted at an accelerating rate, until eventually all of the ice had melted. At this point, students were told:

> *Notice how this is a POSITIVE FEEDBACK system. Melting some of the ice tends to make MORE ice melt. This is because less sunlight is being reflected, so more heat is absorbed.*

Next, students used the simulation's controls to observe the complementary feedback effect, with greater ice coverage causing additional freezing. Again, they were explicitly reminded afterward the way in which this reflected positive feedback behavior.

Finally, students were able to freely interact with simulation for up to three minutes. Additionally, at this point we added sliders that allowed students to directly control the reflectance of ice and water in the simulation.

The complete simulation may be viewed and completed at:

http://cognitrn.psych.indiana.edu/albedo/albedo_F10_grid2.html

Box 2: Ice-albedo computer simulation

What accounts for these differences? One possible explanation is that students in the high context condition were simply distracted by the salient and emotionally engaging introductory content. If so, this distraction may have impaired their ability to attend to the subsequent task, and therefore inhibited their learning about the simulation itself. However, our data suggest that this is not the case. When participants were asked to describe positive feedback systems in their own words after the simulation, those in the high context condition did just as well (and numerically slightly better) than those in the low context group.

Instead, we would argue that the rich context provided by the general introduction served to tie students' knowledge more tightly to this particular content area. Rather than being able to construe the concepts underlying positive feedback behavior independently, those students perceived the ideas entirely in terms of this specific physical system: abstract ideas of variable values, causation and mutual influence were explicitly bound to the more particular notions of heat and reflectance. They were therefore less able to integrate those concepts with the content of new examples involving, for example, economics or biology.

Our findings are somewhat counterintuitive. It seems like a common sense truism that engaging examples are superior for instruction. However, previous research has confirmed that there can be a considerable disconnect between which factors learners (and instructors) *believe* best support learning and those factors that actually do (e.g., Kornell & Bjork, 2008). Just because students are more attentive to engaging examples and enjoy them more does not in itself indicate that they will lead to generalizable knowledge.

Of course, our study examined the use of supporting context that was extraneous to the principle to be learned. While the plight of polar bears is very much related to polar melting, it plays no causal role in the underlying feedback system itself (see Greeno, 2009). We would argue that such extraneous content is very much a relevant issue for classroom instruction, however. In the interest of making materials accessible and holding students' attention, teachers are likely to couch examples in whatever salient context may be available—for example, grounding a discussion of probabilities in terms of LeBron James shooting free throws. More subtle cases of contextualization—such as those in the current experiment, which were not directly involved or mentioned in the example itself—are likely to be even more common.

The current research adds to our understanding of the role of context and specific content in learning from instructional examples, and provides a striking example of just how broadly the definition of "context" may extend.

Acknowledgments

This work was supported by National Science Foundation REESE grant 0910218. We would like to thank Nancy Martin of Jackson Creek Middle School and Akshat Gupta for their help with our research.

References

Anderson, J. R., Farrell, R., & Sauers. R. (1984). Learning to program In LISP. *Cognitive Science. 8*, 87–129.

Bassok, M., Wu, L.L., & Olseth, K.L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition, 23,* 354–367.

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

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

Greeno, J. (2009). A theory Bite on Contextualizing, Framing, and Positioning: A Companion to Son and Goldstone. *Cognition and Instruction, 27,* 269–275.

Goldstone, R. L., & Sakamoto, Y. (2003). The Transfer of Abstract Principles Governing Complex Adaptive Systems. *Cognitive Psychology, 46,* 414–466.

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

LeFevre, J. A. & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction, 3,* l–30.

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

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction?" *Psychological Science, 19,* 585–592.

Lepper, M. R. (1988). Motivational considerations in the study of instruction. *Cognition and Instruction, 5,* 289–309.

McCombs, B. L., & Whistler, J. S. (1997). *The learner-centered classroom.* San Francisco: Jossey-Bass.

Rivet, A.E., & Krajcik, J.S. (2008). Contextualizing instruction: Leveraging students' prior knowledge experiences to foster understanding of middle school science. *Journal of Research in Science Teaching, 45,* 79-100.

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.

Son, J. Y., & Goldstone, R. L. (2009). Contextualization in perspective. *Cognition and Instruction, 27*, 51-89.

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

Wilensky, U. (1999). *NetLogo (and NetLogo User Manual).* http://ccl.northwestern.edu/netlogo/