Instruction abstracted from specific and concrete examples is frequently criticized for ignoring the context-dependent and perspectival nature of learning (e.g., Bruno, 1962, 1966; Greeno, 1997). Yet, in the effort to create personally interesting learning contexts, cognitive consequences have often been ignored. To examine what kinds of personalized contexts foster or hinder learning and transfer, three manipulations of perspective and context were employed to teach participants Signal Detection Theory (SDT). In all cases, application of SDT principles was negatively impacted by manipulations that encouraged participants to consider the perspective of the signal detector (the decision maker in SDT situations): by giving participants active detection experience (Experiment 1), biasing them to adopt a first-person rather than third-person perspective (Experiment 2), or framing the task in terms of a well-known celebrity (Experiment 3). These contexts run the risk of introducing goals and information that are specific to the detector’s point of view, resulting in sub-optimal understanding of SDT.

Cognition can be painted as both context-dependent and context-independent. On the context-dependent side, problem solving is often easiest when framed by supportive, concrete contexts (i.e., Baranes, Perry, & Stigler, 1989; Nisbett & Ross, 1980; Wason & Shapiro, 1971). Unfortunately, there have also been many documented failures for such contextualized understandings to generalize flexibly (e.g., Lave, 1988; Lave & Wenger, 1991; Nunes, Schliemann, & Carraher, 1993). Having wide experience with a range of contexts can help students recognize relevant information for generalization (for a thorough treatment of this idea, see coordination classes, diSessa & Wagner, 2005). But when only a limited exposure to contexts is available, strategic decontextualizations, if successfully employed, can allow human learners to function across contexts and extend prior learning to solve new problems. When symbolic representations such as graphs, equations, or rules are acquired through time-consuming and effortful study, such achievement has been shown to lead to flexible transfer to new situations and novel problems (e.g., Bassok & Holyoak, 1989; Judd, 1908; Novick & Hmelo, 1994). Transfer of learning is often critically important because the specific training domain is just one example of a much
deeper principle (Morris, Bransford, & Franks, 1977; Bransford & Schwartz, 1999; Thorndike & Woodworth, 1901). For example, a teacher presents a case study of evolved coloration of Manchester’s peppered moths not only for students to learn about these moths specifically, but so that they will understand and apply natural selection, variation, and phenotype concepts when they arise subsequently in scenarios involving alligators or bacteria.

Successful pedagogy recognizes the potentials and pitfalls of decontextualized abstractions in the attempts to impart knowledge of general forms, stripped of situational details, actions, and perspectives. The philosophy of decontextualization has historically been challenged because abstract formalisms often seem unnaturally difficult for novices to grasp and learn (Bruner, 1962, 1966). Learning decontextualized representations can result in “inert” knowledge because they are unconnected to particular contexts (Bransford, Franks, Vye, & Sherwood, 1989). For example, statistical tests are useful abstract formulations for a wide variety of situations, but students are often at a loss as to when to actually implement them beyond their classroom examinations (Franks, Bransford, Brailey, & Purden, 1990; Schwartz & Martin, 2004; Schwartz, Sears, & Chang, 2007). Abstract representations that are too separate from the rest of a student’s knowledge will not promote the discovery of productive commonalities across diverse situations. In a recent study exploring the tradeoffs between contextually grounded versus abstract (equation-based) representations, Koedinger, Alibali, and Nathan (2008) found that for simple problems, grounded word problems were solved better, but for complex problems, equations were solved more accurately.

In the effort to bring contexts back into learning, the learning sciences have seen a proliferation of what can be considered “contextualization,” which can broadly be construed as how much learning is embedded in a specific domain or situation. Many types of contextualizations are aimed toward the goal of student-centered learning (McCombs & Whistler, 1997), to make what is learned meaningful to the student. Personalizing learning through direct experience, perspective, and interests offers straightforward means of couching concepts in meaningful experiences. In the richest form, this could mean full immersion in real-world problems in communities of practitioners (i.e., Brown, Collins, & Duguid, 1989; specific instantiations include Gorman, Plucker, & Callahan, 1998; Barab, Squire, & Deuber, 2000). A caveat is that these rich programs are complex and difficult to incorporate immediately into traditional instructional contexts. Considering the potential gains in motivation and knowledge that may come through personalization, it is worthwhile to consider how to personalize learning in traditional instruction as well. Micro-manipulations of contextualization can be implemented easily, allowing them to be replicated many times in the course of a teaching unit.

Empirical results of even minor contextualizations incorporating direct action, perspective, and personal interest have been fruitful for both a basic understanding of cognition and pedagogical inquiry. For example, when children manipulated objects and performed actions referred to in text using physical objects, their reading comprehension improved (Glenberg, Guiterrez, Levin, Japuntich, & Kaschak, 2004). Contextualizing abstract problems in personally relevant or interesting situations (such as familiar schemas: Nisbett & Ross, 1980; Wason & Shapiro, 1971; or fantasy situations: Parker & Lepper, 1992) has also produced learning benefits. These manipulations may have their effect by enhancing intrinsic motivation (Lepper, 1988), but the implications may be much broader. Consider that human reasoning may be inherently grounded in modality-specific, action-specific, and perspective-specific interactions between thinkers and their environments, as embodied psychology (e.g., Barsalou, 1999; Glenberg, 1997) and situated
education (Greeno, 1997) suggest. In light of such views, perhaps personalization has more of an effect than simply increasing interest and personal relevance. Personalized contexts may have an effect on the content that is actually learned.

POTENTIAL CONSEQUENCES OF PERSONALIZATION

Although there may be motivational reasons to personalize, the research reported here explores the possibility of counteracting cognitive implications. One important implication of personalized contexts has always been this extreme possibility: if all learning is tied to specific contexts, the possibility of transfer across domains and phenomena comes into question (e.g., Detterman & Sternberg, 1993; Lave, 1988). After all, if we define transfer as thinking and reasoning across contexts (e.g., Barnett & Ceci, 2002), knowledge must be decontextualized and abstracted (from the Latin, Abstrahere, to pull away) from particulars in order to be transferred (see Reeves & Weisberg, 1994).

The potential for transfer and generalization over a variety of situations provides compelling reasons to understand how individuals might benefit from decontextualization. Discovering, understanding, and using deep principles across domains seem critical for students (see Anderson, Greeno, Reder, & Simon, 1996 for a defense of this assumption) and have historically been common goals for educators (Klausmeier, 1961; Resnick, 1987). The benefits of decontextualization may be most apparent on tests of transfer, but more generally, transfer has also been proposed as a more sensitive indication of learning than other measures such as memory retrieval (Michael, Klee, Bransford, & Warren, 1993; Schwartz & Bransford, 1998).

However, there are reasons to be skeptical of transfer as a pedagogical goal and decontextualization as a strategy toward that goal. Evidence of life-to-school transfer failures such as the mathematical successes of housewives in supermarkets (Lave, 1988) and Brazilian fishermen at the fish market (Nunes, Schliemann, & Carraher, 1993; Schliemann & Nunes, 1990) coupled with their inability to exhibit their mathematical knowledge in school settings could be used to argue that authentic knowledge is based in concrete, real-world situations (Lave, 1988). Decontextualization is called into question considering the failures in the opposite direction, school-to-life transfer, such as the failure of American children who successfully represent negative numbers on a school-learned number line to relate that knowledge to money transactions (Mukhopadhyay, Peled, & Resnick, 1989). One reaction in education is to abandon efforts to foster transfer through abstract instruction and instead focus on training students in situations that are directly pertinent to important and probable future applications. This effort suggests that education should be contextualized in concrete domains as much as possible.

Also, decontextualization may not be the only way to foster transfer. Proper contextualization can also result in robust transfer. A situated learning perspective asserts that generalization occurs because of contextual interactions and commonalities (e.g., Beach, 1997; Lemke, 1997). In other words, transfer situations that are appropriately contextualized can reveal the influence of past learning. For example, a student who brings tools from school (e.g., calculators, software) into the workplace effectively changes their work context to be more like school and consequently implements school learning in the workplace (Beach, 1995). Situated perspectives also endorse learning based on problems that are continuous with everyday knowledge (Lampert, 1986). Changing the type of contextualizations in learning and transfer situations is a non-mutually
SON AND GOLDSTONE

exclusive alternative to decontextualizing learning in hopes of generalizing to dissimilar transfer situations.

The research reported here is motivated by this broad interest in generalization but specifically examines the role of decontextualization in learning and transfer. We have focused our efforts on variations of personalizing contexts (over other types of contextualizations) that often seem thoroughly beneficial (or at worst, benign). How could making something more "learner-centered" or "learner-relevant" be a bad thing? We suspect that any contextualization, implemented without considering cognitive implications, could have unexpected consequences. Personalization should be considered in light of the mounting evidence for inherently perspectival, action-specific, modality-specific representations (see Barsalou, 2003 for a review; Kosslyn & Thompson, 2000; Pulvermüller, 1999). If cognition is inherently perspectival, personalization may induce a particular perspective on the learning content. If that perspective is aligned with the abstract principles of the system, then such a perspective may be helpful. However, if the induced perspective encourages incorrect inferences, counter or orthogonal to the principles to be taught, then we may see a detrimental effect on learning.

We have focused on manipulations of personalization that are closely controllable to facilitate the laboratory study of learning and generalization. However, these are also pertinent to the simple kinds of personalization decisions that arise in everyday lesson planning. Three experiments explore different types of personalization: (a) action-involving experience, (b) conversational narration that places the reader in the story, and (c) familiar case studies with popular actors/characters. Despite the modest nature of our manipulations of personalization, they are nonetheless important because these manipulations reflect the types of contextualizations that are pervasively implemented by teachers, textbook writers, and educational media developers. Whenever pedagogical texts or materials are developed, design decisions related to contextualization are made. Common design decisions include how much background experience to give the learner with the domain, what voice to use in positioning the reader in the text, and whether to take advantage of well-known cases and situations. First we will describe the overall organization and predictions of the three experiments. Then we will review the potential effects of each type of personalization in turn as we introduce the corresponding experiment.

EMPIRICAL STUDIES

To examine the extent to which contextualizations affect the subsequent application of learned principles, we designed a computer-based tutorial about Signal Detection Theory (SDT), a useful structural description of decision-making based on uncertain evidence. A signal detector makes decisions regarding whether a signal is present or absent and each decision can have two outcomes—the signal is actually present or absent. There are four types of events: hits (detector decided “present” and actually present), misses (decided “absent” and actually present), false alarms (decided “present” and actually absent), and correct rejections (decided “absent” and actually absent). The contingent relationships between these four categories (e.g., deciding “present” more often results in more hits but also more false alarms) provide structure that can be used to effectively describe and predict outcomes in many different situations such as doctors detecting sick patients, meteorologists predicting storms, analysts picking out market trends, and farmers classifying ripe fruit.
In many classrooms, books (i.e. Thomas Wickens, *Elementary Signal Detection Theory*), handouts, (i.e., David Heeger, http://www.cns.nyu.edu/~david/handouts/sdt/sdt.html) as well as currently available computer-based tutorials (Claremont Graduate University, http://wise.cgu.edu/sdtmod/index.asp; California State University, Long Beach, http://www.csulb.edu/org/college/hfes/sdt.htm/), SDT is introduced through some example of an individual who must decide whether some signal is present or absent. A common example used in SDT is a doctor trying to determine whether a person’s medical test results indicate sickness or not. Most SDT tutorials describe such cases and then apply a more general framework with more abstract descriptions that could be applied to a variety of SDT situations.

In all of the experiments that follow, participants were taught SDT through a click-through tutorial, which implemented visuospatial representations accompanied by explanatory text since such simultaneous presentation (of graphics and text) has been empirically shown to be effective (Mayer, 2003). To contextualize SDT, the tutorial expanded the typical description of a case study. Instead of describing the principles of SDT in a general abstracted form, they were embedded in the context of a doctor trying to diagnose patients. By using a common example of a real-life SDT situation in a more narrative form, we hoped students would gain an advantage of using a familiar and more compelling problem context. A more detailed description of the tutorial is presented in the methods section of Experiment 1.

In addition to these efforts at general contextualization, each experiment also manipulated an additional aspect of personalization, commonly used to motivate students. In Experiment 1, participants were given discovery experience through a short hands-on activity of detecting signals (diagnosing people as sick versus healthy from their cells) and finding out the outcomes of those decisions. This experience may engage students by allowing them to actually make decisions under uncertainty; this may also motivate the need for a framework such as SDT. Experiment 2 personalized learning by using a more active and engaging tone, simply by using the term “you” in the SDT tutorial. Participants were told that they were the active detecting agent with the “you” pronoun or they were detached from the agent with the pronoun “he.” Experiment 3 examined whether participants benefit from a tutorial that included contextualizing details about a specific and familiar doctor, a character from a popular medical television drama, compared to the same tutorial couched in terms of a generic doctor.

These particular types of personalizing contextualizations were chosen because, although they reflect learner-centered design, their cognitive implications have not been considered as much as their motivational benefits. Also, these personalizations put learners in a particular perspective with respect to the material to be learned. The manipulations of Experiment 1 and 2 place the user in the point of view of the signal detector either directly, through experience, or linguistically. The perspective fostered by Experiment 3, a well-known character, gives students an opportunity to take a perspective that has been portrayed in an entertaining way (on a television show). Although the goal of this tutorial is to teach the larger structure of SDT, our manipulations of personalization emphasize the signal detector’s perspective. If this perspective is helpful for understanding the abstract structure of the system, students may enjoy both motivational and learning benefits of personalization. However, if the perspective of the signal detector limits students’ understanding of SDT, then, although there may be a motivational benefit, learning or transfer may be compromised.
GENERAL PREDICTIONS

Personalization with experiential, conversational, or familiar information may motivate students to give extra effort to comprehending the SDT material. However, if the goal is to understand the general structure of SDT, learning and transfer may benefit from a certain amount of decontextualization as well. If decontextualization and detachment from a specific learning situation are necessary for identifying functionally relevant structure (Goldstone, 2006), then hands-on experience or a particular point of view or familiar details may anchor learning too much to irrelevant (or worse, misleading) past experiences or real-world knowledge (e.g., Rumelhart & Ortony, 1977; Schank & Abelson, 1977). Preconceptions of doctors and their methods of diagnosis may hinder students from appreciating the new insights that a SDT perspective might bring. Additionally, the anchoring influence of personalization may also result in poor transfer to dissimilar (e.g., non-doctor) situations. By this account, there is a possibility for too much contextualization.

There is a tension between the concretely experienced and personally relevant on the one hand and the transportable and general on the other hand that is appreciated by researchers in both cognition and education. Goldstone (2006) argues that hybrids such as “recontextualization,” combining contextualization and decontextualization to create new categories for making predictions and inferences, may be necessary to foster both grounded, connected understanding and flexible transfer. Related notions in education such as “situated generalization” (Carraher, Nemerovsky, & Schliemann, 1995), “progressive formalization” (Freudenthal, 1983) or “action-generalization” (Koedinger, 2002; Koedinger & Anderson, 1998) characterize generalization behavior through a combination of decontextualizing processes and active, concrete, and specific situations. Other approaches suggest interfacing between contextualized and decontextualized learning through coordination classes (diSessa & Wagner, 2005) that emphasize the extraction of invariant information among a wide variety of contextualized situations. One of the benefits of a less active or specific perspective on a doctor diagnosis situation is that by not committing learners to a particularly detailed construal, they can adopt a more general view of the entire system. Personalization may have the consequence of encouraging students to adopt the signal detector’s perspective. However, this may interfere with a more structural understanding of SDT beyond what the detector might know or not know.

EXPERIMENT 1: ACTION-INVOLVING EXPERIENCE

Personal motivation is important to contemporary educators who are concerned with promoting inquiry over the acquisition of factual knowledge, reflecting a classic Piagetian view where true and deep understanding comes from activity at the hands of the learner (Piaget, 1970). Constructivist frameworks endorse rich activities that engage learners rather than merely studying texts (e.g., Savery & Duffy, 1994). By involving students in activities, their learning environment is set up for them to explore and discover principles. Several studies have documented a positive relationship between hands-on activity and test scores found in school settings (Bredderman, 1983; Inagaki, 1990; Stohr-Hunt, 1996).
Unfettered action-oriented experience is often contrasted to direct instruction where a teacher (or text or computer program) guides the activities of a student. Direct instruction may be particularly effective for demonstrating difficult concepts and abstract regularities because students are unlikely to discover these through unconstrained experience. Recently, Klahr and Nigam (2004) found that mastery acquired from directed activity resulted in more sophisticated learning than mastery by discovery. Experience-based learning can be made more effective with teacher guidance (e.g., Brown & Campione, 1994; Cognition and Technology Group at Vanderbilt (CTGV), 1996). Current debates take into consideration the capacity limits of cognitive architecture and suggest that direct instruction (Kirschner, Sweller, & Clark, 2006) or well-supported discovery activity (Schmidt, Loyens, van Gog, & Paas, 2007; Hmelo-Silver, Duncan, & Chinn, 2007) may aid novices by easing cognitive load (Sweller, 1988).

Studying the interaction between guiding information and active experience is a valuable domain for education research given the possible overarching implications for curriculum design and classroom content. Research shows that although active experience is influential and potentially helpful, not all active experiences are equal. Mayer, Mathias, and Wetzell (Experiment 3, 2002) presents a case of active experience being helpful for learning only when it comes before explicit instruction but provides no benefit when it comes after. They suggest that becoming familiarized with a hands-on model first prepares students to understand components that will be critical when learning about the more abstract causal model. Active experiences can change the way students organize and interpret future learning.

Unfortunately, these changes may not always facilitate learning and transfer. DeLoache’s studies (2000) regarding young children’s use of scale models provide empirical evidence that hands-on activity is not always beneficial. In these experiments, a 3-year-old child observes a miniature doll hidden in a scale model of a room. The child is told that a larger version of the doll is in the corresponding location in a larger room and is given an opportunity to search. Some children were given 10 minutes of free play and manual exploration of the model before the hiding event and others were simply shown the hiding event. Children who had an opportunity to play with the model were less effective at using the model as a map of the real room. Furthermore, children in another experiment who had less opportunity to interact with the model (by placing it behind a window) were more effective at searching in the real room. DeLoache (1995, 2000) notes that although experience typically facilitates performance, in this case, experience may have changed the way children viewed the model room. Although a lower level of experience-based salience allowed children to view the model room as a representation, a higher level of interaction with the object prevented this more sophisticated understanding. Hands-on experience induced the children to see the model as a physical toy rather than as a representation for another object. Recently, Uttal and colleagues (Uttal, Bostwick, Amaya, & DeLoache, in preparation) observed that using block letters and numbers to play basketball (throwing numbers into a basket) or blow bubbles (i.e., using the “P” shape as a bubble wand) did not improve performance on letter and number comprehension. In actuality, playing irrelevant games may have actually hurt performance because control children, who did not play with the letter and number toys, showed better comprehension at the end of the study. The lesson from these results is that hands-on experiences can limit or distort a learner’s perspective; these perspectives may change the content and quality of what is learned.

To understand the pedagogical effect of active experience, our experiment takes into consideration the particular perspective induced in our participants through experience. In Experiment 1,
participants were either given the SDT tutorial alone (the Control condition), or a short sequence of active SDT decisions, in which they effectively play the role of signal-detector, followed by direct instruction in the form of the tutorial (the Experience condition). Ideally in educational practice, “experience” and “activity” are much richer than the implementation we have provided here, and experiences can also be tailored to foster an appropriate perspective on SDT. The minimal experience we provide is not designed to be optimal for teaching SDT, but for fostering a signal detector’s perspective. However, it is worth noting that even this limited contextualization might still be an improvement because most SDT tutorials do not implement any active experience at all and merely describe applicable contexts. Another potential benefit of the minimal signal-detecting activity is that this experience is relevant to learning SDT. Actively making responses and examining the consequences might give rise to action-specific insights such as the understanding that hits and false alarms are both possible outcomes of making a “present” response. If these action-specific insights were intentionally guided towards a more general understanding of SDT (such as tallying these active responses into the four decision outcomes of SDT), one might predict that actually experiencing signal detection may be beneficial since transferable learning can draw on activity contexts (Greeno, Smith, & Moore, 1992). Alternatively, the action-specific insights alone may not lead to a general understanding of SDT and instead could bias students to adopt perspectives that turn out to be detrimental to learning and transfer.

METHOD

Participants and Design. Seventy-three undergraduates from Indiana University participated in this experiment for credit. A computer program randomly assigned participants the Experience \((n = 41)\) or Control condition \((n = 32)\).

To quickly read the text presented in the experiment (including tutorial and quizzes) takes about 15 minutes. Four additional participants (from the Control condition) who spent less than 15 minutes to complete the experiment were excluded from analysis. When participants were debriefed at the end of the experiment, they reported how much they previously knew about SDT. All of our participants said they did not know it at all or had heard of it but did not know what it was about.

Materials and Procedure. All participants read through a computer-based SDT tutorial made up of pictures and explanatory text. The principles of SDT were embedded in the context of a doctor trying to diagnose patients with leukemia by examining blood cell distortion levels. Participants were told that patients with leukemia typically have more distorted cells, and healthy patients have less distorted cells. Although cell distortion is an imperfect indicator of health, the doctor tries to optimize his decisions based on this imperfect evidence.

After a brief introduction to this scenario, participants in the Experience condition were told: “Pretend you are a doctor who has gotten the results of a blood test.” They were shown a graphic of a continuum of cell distortion levels, and were shown that the left and right ends had mostly healthy and mostly sick patients, respectively. Then participants were shown cells one at a time
to diagnose as “sick” or “not sick” as shown in Figure 1. After each diagnosis, the participant received feedback indicating the actual sickness/health of the patient. The diagnosis activity included a total of 40 cells that took most participants less than 2 minutes to complete. Then, they were introduced to a tutorial. Participants in the Control condition were shown the brief introduction to the scenario and then immediately received the tutorial without any categorization trials.

The tutorial has been used in other experiments (i.e., Son, Doumas, & Goldstone, under review) and uses a combination of descriptions and diagrams to teach basic SDT concepts. The tutorial is available online at http://jys.bol.ucla.edu/Perspective (the URL is case sensitive). Pilot experiments teaching students SDT with traditional normal distributions contrasted to other attempts using frequency bar graphs supported the claim that frequency information is far easier to understand than probability information (both in general cognition, Gigerenzer & Hoffrage, 1995; and pedagogy, Bakker & Gravemeijer, 2004). We speculated that the overlapping region of the traditional distributions (e.g., Figure 3) was particularly crucial for understanding SDT but also particularly confusing for students. Because we were not interested in teaching graph reasoning per se, we developed bar graphs that utilized non-overlapping spaces and color codes tailored to represent critical concepts of SDT (see Figure 2). Non-overlapping regions of the screen (i.e., top
These are a few screenshots of the tutorial. The distribution of boxes on top of these screens were outlined in red to indicate the actually sick status of these individual patients whereas the boxes presented on the lower half of these screens were outlined in green to indicate their actually healthy status. Boxes on the right side of these screens were labeled with a red "S" in order to show that the doctor had diagnosed these cases as "sick." Boxes on the left were labeled with a green "H" to indicate that the doctor had diagnosed these as "healthy."

Each case is represented by a distorted cell in a box outline. To show the results of the doctor’s categorization, each case is labeled with a red “S” or green “H.” To show actual category membership, the outline of each case is red or green for the two categories of sick and healthy.

The underlying SDT structures of the tutorial and transfer stories are shown here. The participants never saw this figure.
patients, respectively. The cases are spatially ordered, from left to right, by increasing cell
distortion. If the likelihood of a particular level of distortion indicating leukemia is high, there 
are more red cases in the column than green. If the likelihood of leukemia is low for a particular 
level of distortion, then there are more green cases than red ones. Columns that include both 
category 
types can be seen as analogous to the overlapping regions of traditional SDT distributions because 
the same level of distortion could belong to either category. Although in typical SDT distributions 
there is an actual physical overlap to indicate both categories being associated with the same cell 
distortion level, in our diagrams, occurring in the same column indicates the same meaning as 
the overlap. Pilot experiments with non-overlapping representations seemed more effective than 
traditional SDT distributions, which frequently led to incorrect interpretations such as, “Cases 
in this overlapping area are both sick and healthy.” The cases are then separated by category 
into two different histograms to clearly show the actual status of these patients. This reflects the 
SDT distinction between actual signal and noise distributions, as contrasted with the doctor’s 
decision.

The tutorial was a 47-screen self-paced slide show covering basic SDT concepts such as the 
difference between evidence for a decision, the decision, and the actual status of the case. The 
tutorial was designed to help students understand the relationship between changes in decision 
thresholds and resulting changing patterns of diagnosis errors, and between changes in the actual 
population distribution and patterns of errors. At the end of the tutorial students are expected to 
reason about how decisions and distributions affect outcomes. Toward this end, students were 
shown how a decision boundary could lead to two ways of making the right decision (hits and 
correct rejections) and two ways of being incorrect (misses and false alarms). This was followed 
by two examples where the decision boundary was moved in order to show the relationship 
between these categories. Additionally, participants were shown what would happen if the signal 
distribution (all of the targets) shifted along the evidence continuum.

Due to time constraints, the tutorial was limited to showing basic reasoning with SDT concepts 
rather than the equations and other abstract formalisms typically used in SDT. These principles 
were fully contextualized in narrative form rather than generic statements about SDT. Meaningful 
principles in SDT were couched in terms of a doctor and his patients. For example, the concept 
of hits was translated into the story context as “diagnosed sick, actually sick.” An example of 
moving the decision boundary was presented in the story context as a situation where the doctor 
must avoid misses (“diagnosed healthy, actually sick”) because the undiagnosed disease is fatal. 
When the doctor moves the decision boundary, more patients, both actually healthy and sick, are 
diagnosed as sick. Within the entire tutorial, the category labels (“hits,” “false alarms,” “correct 
rejections,” “misses”) were never explicitly mentioned.

After reading through the tutorial, participants answered eight multiple-choice questions about 
the tutorial’s doctor situation that could be answered correctly by applying SDT principles. There 
was a ninth tutorial question (“Which of these is the worst scenario for the patient?”) that was 
used for exploratory assessment, but was not included in the main analyses because it did not 
have a correct answer. Difficult quiz questions were purposefully used to ensure that participants 
needed to use SDT principles rather than relying on common sense (tutorial quiz questions are a 
subset of the questions included in Appendix A).

After the tutorial quiz, participants received an opportunity to analogically transfer what they 
had learned to a different story context, although they were not explicitly instructed that the two 
scenarios were related. Participants read a few paragraphs (included in Appendix A) presented
on three slides describing a small town that wants to export sweet melons and avoid exporting bitter melons. Sweet melons, laden with juice, tend to be heavier, so this town decides to sort the melons by weight (even though weight is only an imperfect indicator of sweetness). Heavy melons are exported and sold whereas light melons are rejected. However, all of the melons are subject to consumer reports that allow the town to find out which melons are actually sweet/bitter. A nine-question transfer quiz was subsequently administered. The transfer quiz is available on the extended Appendix online, http://jsb.bol.ucla.edu/Perspective.

At the end of the experiment, participants were told that these two stories were analogous and asked to explicitly place elements of the two stories into correspondence with each other in a six-question multiple choice mapping quiz. Each mapping question presented an element from one of the stories, such as “sick patient” (signal) and four possibilities from the other story, two of which are viable answers according to SDT—“sweet melon” (signal) and “bitter melon” (noise). The mapping of sick patient to bitter melon is not necessarily incorrect because the mathematics of SDT binary choice is not affected by which category is called “present” and which is called “absent.” The tutorial and transfer situations were designed to reflect the signal-to-signal mapping because the doctor seeks sick patients for treatment and the town seeks sweet melons for shipping. One reflection of that intention is found in the spatial organization of the tutorial and transfer figures, with cell distortion and melon weight increasing from left to right, and signals appearing on the right side (for a schematic illustration see Figure 3). Because of this spatial alignment, we will call this “sick patient-sweet melon” mapping the structural answer. However, to make this match, one must overcome an appealing semantic match between sick patients and bitter (sickly) melons. Thus the mapping quiz was scored in three ways: a structural score, a semantic score, and a combined score (structural + semantic).

Results

Tutorial and transfer scores were subjected to a quiz-type × condition, 2 × 2, repeated-measures ANOVA and the results are shown in Table 1. We found a significant effect of condition (experience + tutorial versus tutorial only), $F(1, 71) = 8.33$, $p < .01$, and a marginally significant effect of quiz-type (training versus transfer), $F(1, 71) = 3.04$, $p < .10$, but no interaction, $F(1, 71) = 1.38$.

The Control condition (tutorial only) scored $0.48 (SD = .19)$ correct on the quizzes while the additional experience dropped performance down to $0.40 (SD = .27)$. Although the interaction was not significant, there is evidence that the effect of condition was stronger for the transfer quiz than tutorial quiz. Although the Control condition participants significantly outperformed those
that received SDT experience on transfer performance, $t(72) = 11.84$, $p < .001$, there was no difference of initial learning on the tutorial, $t(72) = 2.61$.

One might suppose that transfer differences may have been caused by initial differences in learning such that students in the Experience condition acquired less (or perhaps less accurate) knowledge. Although that may be part of the story, differences in learning do not completely account for the differences on transfer. This is evidenced by an ANCOVA with transfer score as the dependent variable which found that even though tutorial score was a significant covariate, $F(1, 69) = 15.36$, $p < .01$, condition still exerted a significant influence, $F(1, 69) = 8.88$, $p < .01$, with transfer performance better for participants in the Control condition than the Experience condition. There was no significant interaction between condition and the tutorial score, $F(1, 69) = 1.49$. If tutorial learning was the primary factor for differences in transfer, there should be no additional influence of condition. This supports the notion that an advantage of decontextualization goes beyond helping students learn more.

One simple explanation of the Experience condition’s poor performance might be that the participants in this condition had more work, an extra initial task, so they were simply tired or not trying as hard as those in the Control condition. We examined this hypothesis by looking at the total time spent on the experiment. Although overall the Experience condition seemed to spend a bit more time on the entire experiment ($M = 23.25$ minutes, $SD = 5.81$) than the Control condition ($M = 21.85$, $SD = 5.40$), this difference was not significant, $t(72) = 1.12$. The categorization task was very short and integrated into the beginning of the tutorial to seem less like a separate task.

This is a counterintuitive finding in some ways because detecting signals would give participants an experiential distinction between being right and being wrong in different ways (e.g., correct rejects, false alarms, hits, and misses), an insight that is relevant to SDT. To understand these results better, we examined the individual questions and found a clue. One of our questions asked, “Which of the following decision strategies will ensure that the doctor maximizes the number of actually healthy people he diagnoses as healthy?” In SDT terms, this asks how the doctor can maximize correct rejections. The correct answer offered is “diagnose everyone as healthy” (the rest of the answer choices are shown in Appendix A) and while fourteen control participants gave correct responses, no one in the Experience condition made this choice. This difference was statistically confirmed by a chi-square analysis on the four answer choices, $\chi^2(3) = 16.61$, $p < .001$. From a doctor’s point of view, a strategy that diagnoses everyone as healthy is unreasonable and perhaps even preposterous.

If the results of our study were only caused by a few odd questions, the effect of SDT experience may not be as negative as our initial analyses suggest. Given that the individual questions asked in these quizzes differed widely in their mean accuracies, a relevant analysis concerns whether experience exerted an impact across all 17 questions (8 tutorial, 9 transfer). If the effect of experience was a general one, across many questions, we should see the effect of condition even in an item-based analysis but no effect of particular items. To answer this question, here and in the subsequent experiments, we report item analyses to compare conditions. The effect of item-type (tutorial or transfer) and tutorial condition (Control, Experience) were evaluated in a $2 \times 2$ repeated-measures ANOVA, with item-type as a between-item factor. Even in this item analysis, there was a significant main effect of condition with the control performing better than the Experience condition, $F(1, 15) = 4.91$, $p < .05$, but no main effect of item-type, $F(1, 15) = 3.00$, and no interaction between condition and item-type, $F(1, 15) = .01$. A paired $t$-test
revealed that this effect of condition was due to a significant advantage of the Control condition over the Experience condition (a difference of $M = .08, SD = .15$), $t(16) = 2.29, p < .05$. This result is consistent with the hypothesis that the disadvantage of experience was found throughout the tests rather than being explained by one or two damaging questions.

Recall that there were two types of mappings, structural or semantic, that could have been made on every mapping question. The six mapping questions were analyzed relative to these two mapping types and two conditions in a $2 \times 2$ repeated-measures ANOVA and the results are shown in Table 2. Although there was no reliable effect of condition, $F(1, 10) = 2.69$, there was a significant effect of mapping type, $F(1, 10) = 10.02, p < .01$. A paired $t$-test revealed that semantic mappings were more frequent than structural ones, mean difference of $.26 (SD = .72), t(72) = 3.05, p < .01$. Additionally, there was a significant interaction between condition and mapping type, $F(1, 10) = 30.62, p < .001$. Paired $t$-tests showed that control participants made significantly more semantic mappings than experience participants, $t(5) = 4.00, p < .01$. The Experience condition made significantly more counterintuitive, structural mappings than the Control condition, $t(5) = 4.36, p < .01$. The structural mappings are more true to SDT principles than the semantic mappings because the structural mapping respects that the signal is defined by what the detector is looking for and the noise is what interferes or is confusable with that signal. However, this tendency for the Experience condition to map structurally was accompanied by worse, not better, performance on the tutorial and transfer quizzes.

**Table 2**

<table>
<thead>
<tr>
<th></th>
<th>Structural mapping</th>
<th>Semantic mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sick Patient–Sweet Melon</td>
<td>Sick Patient–Bitter Melon</td>
</tr>
<tr>
<td></td>
<td>Signal–Signal</td>
<td>Signal–Noise</td>
</tr>
<tr>
<td>No Experience</td>
<td>.22 ($SD = .27$)</td>
<td>.57 ($SD = .35$)</td>
</tr>
<tr>
<td>Signal-Detecting Experience</td>
<td>.39 ($SD = .29$)</td>
<td>.48 ($SD = .32$)</td>
</tr>
</tbody>
</table>

**Discussion**

Why is it that, even across questions, there is a negative effect of a brief diagnosis experience? Or (perhaps put more optimistically) why is there a benefit of having no experience? Could it be because of the induced perspective? The situation where a doctor diagnoses patients and experiences hits, misses, false alarms, and correct rejections is structurally identical to the town making decisions about the sweetness of melons—but only from a perspective that emphasizes the overall structure of SDT. To understand these situations from an SDT viewpoint, some of the ideas about doctors and fruit farming must be ignored. From an SDT perspective, the decision criterion and its relationship to the underlying signal and noise distributions is more important than doctors using common sense in their diagnoses or caring deeply about the well-being of their patients. Part of the difficulty of learning SDT after experiencing the categorization task is that participants may have attended to the patient-diagnosing aspect of this task, which may recruit knowledge extraneous to SDT, such as ideas about practicing medicine and how people deal with illness.
This doctor-based perspective and the resulting patient-based construal may have competed with an SDT perspective that emphasizes certain aspects of the decision making process: signals, thresholds, decisions, hits, misses, false alarms, and correct rejections. Sometimes the goals of a doctor and the solutions offered by through the application of SDT may not match. When discrepancies between the two construals arise, concrete experiences from the doctor’s point-of-view may tilt the balance in favor of attending more to the doctor-specific factors rather than the more formal SDT perspective.

Experience condition participants do seem to develop a notion of signal as “that which is being looked for.” This characterization of targets leads to structural matches, and the successful ignoring of semantic features. This perspective converges with the structure of SDT that educators generally want students to learn. Many psychology classes use this kind of basic feedback experience as demonstrations of experimental paradigms presumably based on the intuition that “putting students in the experiment participant’s shoes,” prepares them to learn the organizing concepts. The implicit pedagogical expectation might be this: by showing students what an experimental participant would see and respond to, students might learn more about experimentation in general. However, our mapping results suggest that experience leads to SDT-compatible interpretations, but the quiz results also indicate that experience leads to competing doctor-specific interpretations. So although there are some benefits, they may not be powerful enough to overcome irrelevant interpretations that come along as baggage.

Experiment 1 examined the effect of a hands-on perspective by giving participants actual signal-detecting experience. When the learner feels like part of the system (the detector), this can lead to different learning outcomes than a perspective that fosters a more distant, observer viewpoint. Providing active experience fosters an active personal perspective for our learners, but there are other ways of manipulating perspective. Can this perspective be induced by a more minor manipulation such as a change in textual voice? Experiment 2 used a personalized conversational narration to further investigate the role of induced perspective on learning.

EXPERIMENT 2: CONVERSATIONAL NARRATION

There are two main reasons to posit benefits of personalized narratives: they play critical roles in engagement and mental models. The first is that personalized conversation engages learners and readily connects them to the learning content. Comments that are directed at the learner, such as strategically placed questions in computer-based tutorials (i.e., “Which do you think . . . ?”), might engage students to think more critically at crucial junctures (Moreno & Mayer, 2004). Contextualizing the learning with personal goals (i.e., “your mission,” “your journey”) also facilitates learning (Moreno & Mayer, 2004). Even with subtle grammatical cues such as personal pronouns (“your lung”) versus a more generic framing (“the lung”), this personalized text aids learning (Mayer, Fennell, Farmer, & Campbell, 2004). Some researchers recommend the second-person pronoun for textbooks and computer-based tutorials because, “it connects the reader to the mathematics because it speaks to the reader directly” (Herbel-Eisenman & Wagner, 2005).

It is important to note that while engagement is often equated with excitement and interest, this connectedness to the material seems to go beyond that. Although vividness and interest have been linked to active perspectives (Fernald, 1987; Velasco & Bond, 1998), interest may not be
necessary for enhanced learning. For example, Mayer et al. (2004) showed increases in learning from the pronoun manipulation but did not show parallel increases in several measures of interest. This suggests that readers are not simply more interested by conversational style but can be engaged with the content in particular ways. This leads to the second influence of conversational narration—textual voice can change how the reader understands the text by changing the mental models that people build. Theories in discourse processing have claimed that through text comprehension, agents and situations are mentally constructed (Chafe, 1994; Graesser, Bowers, Bayen, & Hu, 2000). Bergen (2006) showed that pronoun-introduced perspectives, such as “you,” which engage participants in the activity, are influential in mental modeling, a process often implicated in high-level reasoning tasks (Johnson-Laird, 1983).

More specifically, first-person (“I am a doctor . . .”) and second-person (“You are a doctor . . .”) narration might unite the reader with the actions of the character or narrator. Third-person narration (“He is a doctor . . .”) is considered a detached, omniscient observer in formal literary study and often invisible to naïve readers (Duchan, Bruder, & Hewitt, 1995; Graesser et al., 1996). There may be a particular superiority that comes from relating the material to the self because such information might be easier to process (Spiro, 1977).

d’Ailly, Simpson, and MacKinnon (1997) applied these ideas about mental modeling and personalization to grade school children solving word problems. Children were asked to solve different types of comparison word problems. There were problems where the second-person language identified participants with the known anchoring quantity for the comparison (i.e., the you-Known problems, “You have 3 balls. You have 2 more/less balls than Bob. How many balls does Bob have?”) or identified them with the unknown quantity (i.e., the you-Unknown problems, “Bob has 3 balls. Bob has 2 more/less balls than you. How many balls do you have?”). When the pronoun “you” was substituted in the known position and served as the anchor, children were better able to solve these word problems compared to problems with other names in them.

However, there was no facilitation when the pronoun occurred in the unknown position. d’Ailly and colleagues hypothesized that children tend to anchor the referent according to the self term (“you”), and when they need to reverse the anchor (as in the case of the you-Unknown problems), they had difficulty doing so. Similar studies with adults solving problems about linear ordering suggest that the self-term easily becomes the point of reference (d’Ailly, Murray, & Corkill, 1995).

These results suggest that personalized conversation may lead to better interpretations—but not in all situations. Perhaps for cases such as “your cloud” and “your plant” (Moreno & Mayer, 2004) personalization provides the best interpretation. However, the work by d’Ailly and colleagues (1997) shows interpreting the “you” perspective as the anchor may not always be optimal for solving the problem. Experiment 1 provides some evidence that anchoring the participant to the signal detector’s perspective may not be optimal for learning SDT and may in fact be detrimental. So, to promote a personalized perspective without manipulating experience, we changed the wording in the tutorial to reflect a second-person perspective (“you”). If learners tend to anchor their perspective at the self term, this manipulation puts the learner in the role of the signal-detecting doctor. To provide a control condition that had a perspective more like an outside observer, we used the third-person perspective (“he”).

Although this may seem like a modest manipulation, if “you” helps, this can be readily applied in computer-based tutorials. However, if it does not, this may be a warning that we need to be careful when applying such personalizing contextualizations.
Method

Participants and Design. Fifty-five undergraduates participated for credit. Participants were randomly assigned to be in the Third-person condition \((n = 27)\) or Second-person condition \((n = 28)\). Four participants (2 from the Third-person condition, 2 from the Second-person condition) were excluded from analysis because they spent less than 15 minutes on the experiment. Participants reported having no previous knowledge of SDT.

Materials and Procedure. Participants read through a computer-based SDT tutorial similar to the one used in Experiment 1. Afterward, they took a tutorial quiz worded in the corresponding manner (second-person or third-person voice) of their tutorial. The transfer scenario description and quiz followed. We designed additional questions for a total of 12 tutorial and 12 transfer questions (these have been included in Appendix A). The same mapping quiz used in Experiment 1 was also administered.

For the Second-person condition, all of the tutorial and the corresponding tutorial quiz questions were modified to indicate that the participants themselves were the doctor. For example, the introduction in the Second-person tutorial read, “Imagine that you are a doctor who looks at blood samples to check if your patients have leukemia, a cancer of the bone marrow. Since bone marrow produces blood cells, you can look for distorted blood cells to diagnose your patients.” In the matching Third-person tutorial, participants were encouraged to “Imagine a doctor who looks at blood samples to check if his patients have leukemia, a cancer of the bone marrow. Since bone marrow produces blood cells, the doctor can look for distorted blood cells to help him diagnose his patients.”

When participants were asked to imagine alternative scenarios in the Second-person condition, they were instructed to imagine placing themselves in the pertinent position. For example, participants were told: “Changing the decision boundary is something you can do to change what kinds of mistakes you make. There are some things that are out of your control that also affect how good your diagnosis is. Consider a situation where it becomes even harder for you to diagnose your patients because everyone started taking vitamins that make distorted cells caused by cancerous cells look better. So now sick people have less distorted cells than they used to.” The corresponding Third-person version read, “Changing the decision boundary is something the doctor can do to change what kinds of mistakes he makes. There are some things that are out of his control that also affect how good his diagnosis is. Consider a situation where it becomes even harder for the doctor to diagnose his patients because everyone started taking vitamins that make distorted cells caused by cancerous cells look better. So now sick people have less distorted cells than they used to.”

Results

Subject analyses showed no significant main effects of test, \(F(1, 53) = .97\), and no interaction between condition and test, \(F(1, 53) = .77\). However, there was a marginally significant effect of condition, \(F(1, 53) = 2.81, p < .10\) with second-person perspective tending to produce worse performance than the third-person perspective. Given this tendency for worse “you” performance
for both tutorial and transfer measures, it may have merely been that they learned less overall from
the tutorial. An ANCOVA on transfer performance revealed tutorial performance as a significant
covariate, \( F(1, 51) = 24.55, p < .001 \), no significant effect of condition, \( F(1, 51) = 1.54 \), and no
interaction, \( F(1, 51) = .13 \), suggesting that the learning differences induced by these perspectives
may have led to the subtle differences seen in transfer. Although the “you” versus “he” perspective
difference was a more modest manipulation than the experience-induced perspective, the more
personalized training still results in somewhat reduced learning.

To more finely assess whether the “you” perspective acted in similar ways to the Experience
condition, we examined the individual quiz items. Because the quizzes in Experiment 2 included
all of the questions used in Experiment 1, we were able to analyze the same question about max-
imizing the number of actually healthy people diagnosed as healthy (tutorial question #7). Recall
the correct answer is an unintuitive one from a doctor’s point of view (“diagnose everyone as
healthy”). Eleven out of 27 participants in the “He” condition were able to make this choice versus
7 out of 28 in the “You” condition. The most popular answer for the “You” condition was a caution-
ary and practical one, but not one particularly informed by SDT (“look more carefully at the cell
distortion levels before your diagnosis”). This choice was made by 10 out of 28 “You” participants
compared to the single person that made this choice in the “He” condition. A chi-square analysis on
the four answer choices confirmed these differences between conditions, \( \chi^2(3) = 10.90, p < .05 \).
These results suggest that the “You” condition prompted participants to use general knowledge
about medical diagnosis rather than solely relying on SDT principles. If general medical knowl-
edge were aligned with the inferences of SDT, then perhaps such contextualization would result in
better learning. Here, however, the common sense answer competed with the choice informed by
SDT.

The slight disadvantage to the “You” condition seen in the subject-analysis may have been
caused by a few difficult questions. So like Experiment 1, we performed an item-analysis over the
two types of items (those in the tutorial and transfer) and the two perspective conditions (“He”
and “You”). This 2 \times 2 repeated-measures ANOVA showed a significant main effect of condition,
\( F(1, 22) = 19.83, p < .001 \). There was no significant effect of item-type, \( F(1, 22) = .12 \), nor was
there a significant interaction, \( F(1, 22) = 1.59 \). This suggests that over all the questions, those
in the “You” condition showed limited learning and transfer. Also, this analysis shows that the
tendency for worse “You” than “He” performance suggested by the subject-analysis is significant
when analyzed by item. Collapsing over the two quizzes, the “He” condition scored an average of
.57 (SD = .20) while the “You” condition averaged .49 (SD = .21). The results, organized
according to quiz, are shown in Table 3.

### TABLE 3

<table>
<thead>
<tr>
<th>Tutorial (12 Questions in Tutorial)</th>
<th>Transfer (12 Questions in Transfer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal</td>
<td>Sick patient</td>
</tr>
<tr>
<td>Noise</td>
<td>Healthy patient</td>
</tr>
<tr>
<td>“He”</td>
<td>.55 (SD = .20)</td>
</tr>
<tr>
<td>“You”</td>
<td>.49 (SD = .20)</td>
</tr>
</tbody>
</table>
TABLE 4
Tutorial and Transfer Mappings from Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Structural Mapping</th>
<th>Semantic Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sick Patient–Sweet Melon</td>
<td>Sick Patient–Bitter Melon</td>
</tr>
<tr>
<td>Signal–Signal</td>
<td>Signal</td>
<td>Signal–Noise</td>
</tr>
<tr>
<td>&quot;He&quot;</td>
<td>0.23 (SD = .26)</td>
<td>0.62 (SD = .31)</td>
</tr>
<tr>
<td>&quot;You&quot;</td>
<td>0.19 (SD = .16)</td>
<td>0.64 (SD = .26)</td>
</tr>
</tbody>
</table>

Mapping quiz performances were analyzed relative to perspective condition as well as type of mapping (semantic versus structural). The results are shown in Table 4 and although there were no differences between conditions, \( F(1, 10) = .02 \), there was a significant difference in prevalence between mapping types, \( F(1, 10) = 29.10, p < .001 \), such that semantic mappings were made more often over all six questions than structural ones.

Discussion

Research from embodied cognition shows that particular text-based perspectives influence memory (Abelson, 1975; Pichert & Anderson, 1977), comprehension (Bower, Black, & Turner, 1979), and category judgments (Barsalou & Sewell, 1984). These studies suggest that a participant’s adopted perspective can influence cognitive performance. Experiment 2 indicates that these findings are important to pedagogy because grammatically induced perspective can also influence learning and transfer of structural information. Particularly if the particular vantage point offered to students is non-optimal or limiting, learning opportunities consistent with an active perspective may have negative consequences. Our results suggest that the detector perspective elicited by the second-person pronoun can hinder learning and transfer of SDT.

Our negative result combined with other, typically positive, results of personalization found by other researchers begs the question: When should personalizing texts be used? We suspect that it largely depends on the content to be learned. In SDT we are interested in an overall system, and through personalization, the learner becomes positioned as a participant inside of the system. The goals, structures, and affordances of an individual within a system may be different than information available from an outside observer’s perspective. Wilensky and Resnick (1999) have observed that novices to a domain often have trouble using multiple perspectives. Students imitating agents have been known to experience difficulty abandoning the agent’s perspective in order to understand systems-level phenomena (Penner, 2001). Applying this tension between individual and system-level perspective to our situation suggests that our Second-person condition may have emphasized the individual level to the detriment of the system level. As an example, information that is present in the system but unavailable to the doctor (e.g., actual sickness/health of a patient) seems to be underemphasized for students in the second-person perspective.

The participant perspective induced in our experiments was that of a doctor. This seemed to activate practical knowledge of medical diagnoses. Background knowledge of doctors and patients can help contextualize abstract SDT patterns, rendering them more comprehensible. Our results suggest a caveat: Contextualization could be misleading if the structural content is embedded in the medical context so that it detracts students from using relevant SDT principles. Instead,
students who learn from the doctor’s perspective seem to prefer their common sense notions of making reasonable diagnoses as opposed to SDT’s focus on optimal decisions over many cases. The SDT framework is often not enough to decide what to do in real life medical situations because there are other pressures that may affect the decision boundary (i.e., false alarms may be financially costly or systematic misses could cost lives). However, SDT does provide clear and useful principles for evaluating and predicting medical situations that sometimes contradict typical common sense notions. Less contextualized construals or a more detached learning situation may draw more attention to structural SDT principles, resulting in better learning and transfer.

The overall pattern of results between Experiments 1 and 2 suggests that the active detector/doctor point of view can be detrimental to learning and transfer of SDT principles. This active participant point of view can be instantiated in multiple ways. Hands-on activities may be a particularly strong manipulation and conversational style might be more subtle. There have been a number of findings that show imagined actions (e.g., Glenberg et al., 2004; Schwartz & Black, 1999), as well as watching someone else perform the action (i.e., Lozano, Hard, & Tversky, 2006), influence thinking in ways similar to first-hand action.

At the very least, it is important not to assume that student’s active engagement with learning material necessarily translates to their deep understanding of it. Casual observation of people’s behavior at science museums suggests that an engaging, interactive display may attract participants, but does not guarantee that people will take the time to learn the scientific principle demonstrated by the display. In fact, the display may be so engaging that people choose to engage in the display instead of learning the principle. Harp and Mayer (1998) have also shown that irrelevant “seductive details” reduce learning of main ideas and generation of problem-solving solutions. Similarly but more subtly, the benefit of Experiment 2’s third-person point of view may have been its relative neutrality and lack of engagement. Experiment 3 provides a more direct test of this seductive detail hypothesis.

EXPERIMENT 3: SPECIFIC, FAMILIAR, POPULAR CHARACTER

Thus far, we have seen how two common approaches to personalization do not necessarily provide the best perspective for learning. In Experiment 3, we want to examine another popular approach to personalization, using elements that are familiar and popular with students. Using personally relevant details—information from background experiences, friends and teachers as characters, favorite topics—has been shown in numerous studies to increase interest, understanding, and reasoning (Anand & Ross, 1987; Lepper & Cordova, 1992; Ku & Sullivan, 2002; López & Sullivan, 1991, 1992). Even characters and objects that come from an interesting fantasy context (i.e., pirates, spaceships) show similar boosts to motivation and learning (Cordova & Lepper, 1996; Parker & Lepper, 1992). Beyond increasing interest, personalized details may also have a significant effect on learning through cognitive impact. Familiar contexts and details can relieve cognitive load and boost learning because they may be easier to reason with (Ross, 1983; Wason & Shapiro, 1971). Also it may be easier for learners to interrelate information into a coherent model given more familiar and compelling information (Ross, 1983; Anand & Ross, 1987).

On the one hand, the addition of interesting details seems like an obvious way to make learning more engaging and meaningful. On the other hand, these additional details can also seem like an obvious addition of seductive details, interesting but ultimately distracting from learning. Much
of the evidence regarding personalized details is positive, but there are some limitations. Some researchers question whether the motivating effects of personalized details are due to novelty and would wear off if used systematically (Ross, Anand, & Morrison, 1988). There are also several cases where personalization seems to have no significant improvement on actual test scores (Bates & Wiest, 2004; Ku & Sullivan, 2000; Wright & Wright, 1986). There are also cases in which personalization can distract or mislead. Renninger, Ewen, and Lasher (2002) showed that when the domain is personalized to student’s own interests, such as training dogs or ballet turnout, this can potentially limit strategies and mask a lack of knowledge. Part of learning in any domain is distinguishing between relevant and irrelevant information, the addition of extraneous information that might be confused for relevant information can easily increase demand (Muth, 1984).

There are multiple ways to implement familiar or interesting details but it seems that the best conditions for personalization are: (a) when irrelevant details are easily distinguished as separate from the problem, and (b) when they do not increase demand (at least in terms of amount of material). In Experiment 3, we focus on how interesting human stories, such as anecdotes, individual cases, and richly developed stories can be incorporated into classroom learning through highly developed characters/environments as opposed to generic characters/situations. A geometry problem about a basketball player’s position relative to the hoop can also be instantiated with a popular player such as Kobe Bryant or a generic character. A character could also become more familiar through the learning series. A well-developed example of teaching with character-specific narratives is the Jasper series for teaching mathematics (CTGV, 1992), where students learn about Jasper in a detailed way, with his motorboat on Cedar Creek, stopping at Larry’s dock, and so on. Students engaged in corresponding projects involving problem identification in these rich stories have shown improvements on achievement tests (Vye et al., 1989). Another example of using a well-known character to promote interest is the use of Young Sherlock Holmes to enrich a series of lessons in language arts and social studies (Bransford, Kinzer, Risko, Rowe, & Vye, 1989).

Experiment 3 incorporated a currently popular television character into the SDT tutorial to create two versions, one with an interesting, familiar doctor and the other with a generic doctor. If this personalization through a well-known character has the effect of drawing users to be involved and take a personal active stance, we may see effects parallel to that of Experiments 1 and 2. If the familiarity and background experience with this specific doctor organizes the learner’s understanding, participants may learn SDT with doctor-specific construals rather than the more general SDT principles.

Method

Participants and Design. Sixty-four undergraduates participated for credit. Participants were randomly assigned to be in the Generic (n = 31) or Specific (n = 33) condition. Nine additional participants who spent less than 15 minutes to complete the experiment were excluded from analysis (3 from the Generic condition, 6 from Specific condition). When participants were debriefed at the end of the experiment, they reported how much they previously knew about SDT. Two additional participants told us they already knew SDT and were excluded from analysis (both from the Generic condition). The rest of our participants did not know SDT at all or had heard of it but did not know what it was about.
**Materials and Procedure.** Participants read through a computer-based SDT tutorial similar to the ones used in the previous two experiments. Undergraduate students from an upper-level research class and undergraduate research assistants in two laboratories (not in the subject pool) were informally polled for popular doctors on television. The Specific condition was constructed around the most popular of these suggestions, a well-known character on a currently popular television drama, Dr. Derek Shepherd, a neurosurgeon on “Gray’s Anatomy.” The tutorial used in the previous two experiments was modified such that the doctor is now a neurosurgeon detecting operable versus inoperable tumors. In the Specific condition, the tutorial starts off with this paragraph of background information about Dr. Shepherd: “Derek Shepherd is a highly successful neurosurgeon. He works at a prestigious research hospital called Seattle Grace and he has performed numerous complicated, risky procedures including a stand-still operation, a double-barrel brain bypass, and spinal separation surgery for adult conjoined twins.” Interspersed with this paragraph were four pictures of the specific doctor in surgery situations, taken from the television show.

The Generic tutorial contained the same information only framed for neurosurgeons in general: “Neurosurgery is a complex and difficult field. The procedures involved are often risky, dangerous, and require great skill and training. Some of the more difficult procedures including stand-still operations, double-barrel brain bypasses, and the brain/spinal separation surgery.” There were four pictures of unknown neurosurgeons performing operations. Other than these opening paragraphs and periodic mentions of the name “Dr. Shepherd” in the Specific condition or “the doctor” in the Generic condition, the tutorials were identical. Both Generic and Specific conditions received the same tutorial quiz where there was no mention of “Dr. Shepherd.”

The 12-question tutorial quiz from Experiment 2 was modified to reflect the new tutorial situation about neurosurgery. The transfer situation and 12-question transfer quiz were the same as those used in Experiment 2. The six-question analogical mapping quiz included modified tutorial elements. Note that the mapping situation has changed and the tutorial signal is “operable tumor” (which has a known but difficult treatment, surgery) and noise is “inoperable tumor” (which does not have a known treatment). A signal-to-signal mapping (what the doctor is looking for and what the farmers are looking for) is an operable tumor to a sweet melon, both a structural (they are both targets) and semantic mapping (they are both good). Because of this overlap, this mapping was more intuitive than Experiment 1 and 2’s sick patient to sweet melon mapping. Thus, each of Experiment 3’s mapping questions has one correct answer, the structural and semantic match. We created this congruent alignment to confirm that contextualization is not only influential in incongruent alignments (Experiments 1 and 2). Ideally we would test matching and mismatching versions of the tutorial and transfer stories with each manipulation of contextualization but this would be prohibitively expensive in terms of time and participants, so we chose to implement one experiment with an intuitively appealing structural match.

At the very end of the experiment, participants were asked with a multiple choice question how familiar they were with this particular television show. Participants who reported watching “a few” to “a lot” of episodes of the show were placed into the familiar category and those that reported “not at all familiar” and “not too familiar” with the show were put in the unfamiliar category.
Deviating from previous results, subject analyses showed no significant main effect of condition, $F(1, 62) = 1.98$, no significant interaction, $F(1, 62) = 1.65$, yet showed a highly significant effect of test, $F(1, 62) = 13.98$, $p < .001$. With performance on the transfer quiz being significantly greater ($M = .62, SD = .22$) than the tutorial quiz ($M = .50, SD = .21$), $t(63) = 3.7$, $p < .001$. This boost in transfer that we have not seen in the previous two experiments may be due to the better alignment of the learning and transfer contexts. The match between structural and semantic mappings also resulted in a high prevalence of signal-to-signal mappings made by both Specific ($M = .78, SD = .23$) and Generic conditions ($M = .71, SD = .25$); these were not significantly different from each other, $t(63) = 1.53$. Both groups made very few signal-to-noise mappings (shown in Table 5) and did not differ from each other, $t(63) = 1.58$.

Although there was no effect of condition on individual subjects, there is still the possibility that for each question, one of the training conditions may have found it a bit easier to answer. An item analysis over the two types of questions (those in the tutorial and transfer quizzes) and the two tutorial conditions (specific and generic) resulted in a $2 \times 2$ repeated-measures ANOVA that did show a significant main effect of condition, $F(1, 22) = 6.53$, $p < .05$, but no significant effect of item-type, $F(1, 22) = 2.69$. The interaction was marginally significant, $F(1, 22) = 2.81$, $p < .10$. Across all 24 items (12 items from tutorial quiz and 12 from the transfer quiz), participants in the Generic condition significantly outperformed those in the Specific condition. These results are shown in Table 6. These results reflect a disadvantage for the specific perspective, consistent with the other manipulations of contextualized perspective. As with Experiment 2, the results are significant by the item, but not participant, analysis.
Curiously, the item analysis did not reveal any difference across the two types of test questions. Further examination revealed that the generic group did not show significant differences in their tutorial and test quiz performance, $t(23) = .96$, while the specific group did, $t(23) = 4.54$, $p < .05$. In even closer detail, one of the tutorial questions sets up an ambiguous situation and asked the participants what the doctor would do if he had to choose between making a diagnosis according to SDT principles (using the decision boundary) versus “trust[ing] his experience as a doctor...” (as well as two other foils—available in Appendix A). Half of the participants in the Generic condition (16/31) chose the SDT solution compared to 5/33 in the Specific condition who preferred to trust the doctor’s opinion (15/33) more than those in the Generic condition (9/31) and this difference was reflected in the chi-square analysis, $\chi^2(3) = 11.12$, $p < .05$. A salient principle character may detract attention away from SDT structure or other cogent parts of the problem and highlight qualities of the specific person.

We have mainly construed the between-participants manipulation in terms of a specific versus generic doctor, but another way to think about this is that we included a familiar (famous) doctor as well. Because this categorization would be most appropriate for participants who are actually familiar with the particular television show, we categorized participants according to their reports of familiarity. The scores from the specific doctor condition, shown in Table 5, were analyzed according to a $2 \times 2$ repeated-measures ANOVA with respect to item type and show familiarity. Although there were no main effects of item type or familiarity, $F(1, 22) < 2.91$, there was a significant interaction, $F(1, 22) = 6.21$, $p < .05$. Paired $t$-tests showed that the participants in the Specific condition who were familiar with the show had poorer performance on the tutorial quiz, scoring 10% ($SD = .17$) less than unfamiliar participants, but this difference was only marginally significant, $t(11) = 1.98$, $p = .07$. However on the transfer quiz, the difference trended in the opposite direction with familiar participants doing 6% ($SD = .14$) better than those who were unfamiliar with the show, $t(11) = 1.52$, $p = .15$. This is an odd pattern of results because better performance on the tutorial is typically expected to produce better transfer performance. In the current case, those participants that are familiar with the television show do not initially seem to show good learning but their transfer performance does not suffer. This result suggests that educators should examine the role of contextualization in assessment (also recommended by Bates & Wiest, 2004) as well as during learning.

Familiarity with the TV show did not generally hurt or help performance because when the scores from the Generic condition were analyzed with regard to item type and show familiarity, there were no effects of familiarity, $F(1, 22) < .17$, only an effect of item type, $F(1, 22) = 4.61$, $p < .05$. These results are also shown in Table 5.

**TABLE 7**

Tutorial and Transfer Results of Experiment 3 Broken Down by Familiarity with this Particular Television Show

<table>
<thead>
<tr>
<th>Condition</th>
<th>Tutorial</th>
<th>Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic Condition</td>
<td>.53 ($SD = .18$)</td>
<td>.61 ($SD = .19$)</td>
</tr>
<tr>
<td>Familiar with TV show</td>
<td>.55 ($SD = .21$)</td>
<td>.64 ($SD = .19$)</td>
</tr>
<tr>
<td>Unfamiliar with TV show</td>
<td>.52 ($SD = .26$)</td>
<td>.56 ($SD = .14$)</td>
</tr>
<tr>
<td>Specific Condition</td>
<td>.42 ($SD = .22$)</td>
<td>.62 ($SD = .15$)</td>
</tr>
</tbody>
</table>
Discussion

These results reflect a slightly different pattern from the other two experiments so far, where transfer scores were usually close to tutorial scores and participants in the more richly contextualized condition suffered in transfer. In the current experiment, the transfer scores were appreciably better than tutorial scores. Although the same SDT pictures and graphs were used across all of the experiments, the current tutorial had a significant overall change to the story context, from a doctor examining distorted blood cells in order to diagnose leukemia to a doctor examining tumor density to diagnose operability. This change may have resulted in a generally increased similarity to the melon farmers examining melon weight to decide sweetness. Tumors and density may provide a better match to melons and weight than cells and distortion. These changes to the story resulted in more consistent mappings between the tutorial and transfer situations. Experiments 1 and 2 had a signal-to-signal match of sick patients to sweet melons—a counterintuitive alignment. However, Experiment 3’s change resulted in a signal-to-signal match between operable tumor and sweet melons—a more intuitive, semantically congruent mapping. Previous research has suggested that variations in the similarity and alignability between two situations can have a strong effect on transfer (i.e., Gentner & Toupin, 1986; Gick & Holyoak, 1983; Ross, 1987; Son, Doumas, & Goldstone, under review).

Recall the effect of contextualization in Experiments 1 and 2—participants defaulted to domain-specific knowledge about medical diagnosis and doctors rather than the use of SDT principles. This may have reduced learning of structural principles that could be used in future transfer situations (Bransford & Schwartz, 1999; Schwartz, Bransford, & Sears, 2005). However, in Experiment 3, the familiar character personalization did not result in poorer transfer results despite poorer performance on the tutorial quiz. This may have been due to the difference between pretending to be a doctor when the learner is not actually a doctor versus learning about an actual doctor who presumably knows more about medicine and has more experience than our participants. Instead of triggering schematic doctor knowledge, this manipulation may have activated specific doctor knowledge—a perspective that may be detrimental for answering questions that probe participants’ knowledge of SDT in the context of this doctor, but not for learning SDT principles that can be applied to different domains.

GENERAL DISCUSSION

Contextualizing elements such as activities, personal perspectives, and concrete examples are prevalent in education. Perhaps this is, in part, a reflection of the intuition that well-grounded understandings come from concrete experience. However, the literature on the effect of actual concrete experience on transfer is mixed. There are examples of concrete experiences that foster transferable knowledge (e.g., Inagaki, 1990; Mayer, Mathias, & Wetzell, 2002) but there are also examples of knowledge that is fixed to particular aspects of the learning situation (e.g., DeLoache, 1995; Lave, 1988). Our experiments suggest that it is not that concrete experiences, activities, and demonstrations are generally good or bad for transfer, but rather these manipulations cause particular construals that affect learning and transfer.

As a summary, the ways we personalized instruction hurt performance on quizzes that hinged on knowledge of abstract principles. More specifically, perspectives fostered by active experi-
ences as well as personalized language were detrimental to learning and subsequent transfer (Experiments 1 and 2). Experiment 3 tested whether personalization through specific and familiar details produced the same effects. The specific contextualization seemed to hinder learners from revealing what they learned about SDT in questions about the specific celebrity doctor. Even though the personalization provided in these experiments was relatively subtle, all three experiments showed similar detriments for personalization. However, these analyses were only significant when analyzed according to item; only Experiment 1 showed significant detriment of personalization with subject analysis.

Our results should not be taken as opposing contextualization, personalization, or learner-adapted approaches. Instead, our experiments show that the “one size fits all” approach, where personalized contexts are simply grafted onto contents, could have negative cognitive consequences. Undoubtedly, participating in activities and evoking real world knowledge is influential and can result in effective activity-specific encoding. For example, warehouse drivers in a dairy organize information according to warehouse location and pallet size whereas consumers typically encode by general categories (Scribner, 1985). Their activities and perspectives allow them to selectively encode relevant information. However, this context-bound encoding leads to potential pitfalls. Trying to answer questions about SDT after hands-on activity or through an active participant perspective seems to lead to construals based on common sense knowledge about doctors, rather than SDT decision making more generally. A decontextualized understanding may be beneficial for learning structural principles by leading to a broad, detached understanding of a situation rather than being guided by a particular perspective.

Even though a high degree of contextualization may be detrimental to transfer of system-level knowledge, research suggests that being able to adopt or simulate an active view might be important for other learning situations. For example, Barab et al. (in press) has found that students learn to apply scientific methodology better using avatars in a fully immersed virtual world (the Quest Atlantis project: Barab, Thomas, Dodge, Squire, & Newell, 2004; Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005) than in the context of a third-person story problem clicking through hyperlinks. Skilled understanding requires actions to transfer across contexts and active perspectives may be a necessary part of skill acquisition. Certain problem-solving skills have been shown to benefit from rich problem-based environments (Cognition and Technology Group at Vanderbilt, 1992). Evidence from cognitive science (Needham & Begg, 1991) and education (Barron et al., 1998) show that active problem-solving or project-based perspectives are good for understanding information relevant to these skills. Framing contexts should not just be seen as scaffolding for static knowledge but as relevant perspectives for particular actions.

Activities and frames that are fine-tuned to the learning situation at hand could be an important educational contribution because it can be a significant improvement over real-world experience. When students are simply thrown into real-world experience, the appropriate perspective to take is not immediately apparent. Additionally, combinations of active and less active perspectives can be highly effective. For instance, Chi, Roy, and Hausmann (2008) have shown that watching interactive tutoring is as effective as actually being tutored. Projection into and identification with an actor can be a useful pedagogical tool for fostering particular perspectives. Students can be asked to imagine themselves in the place of an observer, participant, or protagonist, fostering a multi-perspective view of a situation. Just as some situations are easier to imagine than others (Hegarty, 1992), some perspectives may be easier to adopt than others. If an active perspective is relatively easy to adopt, then perhaps educators should focus on perspectives that students are
less likely to spontaneously adopt. An important skill to acquire might be the ability to choose the right perspective in order to extract relevant information or to integrate information from several perspectives. DiSessa and Sherin (1998) argue that “Shifting the means of seeing, a fortiori, is the core problem of conceptual change.” The role of coordinate classes is to unify apparently disparate situations and construe them in a common manner. For example, in the case of applying SDT principles, perspectives that focus on entities such as targets and signals are more apt than those framed in terms of doctors and patients. Framing tutorials to teach these SDT-appropriate perspectives has benefits over simply assuming that emphasizing elements of the specific doctor scenario will promote effective learning.

Students need meaningful understanding and perhaps activity-based contextualization or familiar contexts provide meaning. However, particular perspectives lead to certain aspects becoming highlighted relative to others. If there is a conflict between perspectives, such as seeing individual trees or a unified forest, making trees salient can detract from understanding the forest. Scientific and abstract understanding often requires appreciating high-level organization, so in order to emphasize system-level meaningfulness, one must have a perspective that can take the entire system into account. Activities, personal perspectives, and well-known specific characters can sometimes produce a limited or skewed understanding of the system. Models targeting levels that are relevant for the problem at hand are more effective for explaining phenomena (Gentner & Stevens, 1983; Wilensky & Resnick, 1999). If we aim to foster this kind of perceptual reorganization (Goldstone, 2006), we may want strategically to encourage students to create such perspectives.

CONCLUSION

If cognition is for the purpose of action (e.g., Glenberg, 1997), one’s perspective is vitally important. However, just because we experience events through first-person, action-oriented, detailed perspectives does not mean that emphasizing these perspectives is the optimal way to learn about generalizable structure. Instead, pedagogical efforts to make learning more intrinsically motivating and personalized to students should consider the cognitive effects of such perspectives.

Cognitive research can also learn an important lesson from educational efforts: The influential role of perspective and actions in learning should be considered in theories about learning and transfer. There are technically no context-free, perspective-free learning opportunities, so how do we acquire and impart seemingly context-independent skills and knowledge? To examine the intersection between transportable and generalizable structures with real-life, richly encased in stories, details, and activities, is an important ongoing goal for both cognitive and pedagogical research.

REFERENCES


Barsalou, L. W., & Sewell, D. R. (1984). Constructing representations of categories from different points of view. Emory Cognition Project Report #2, Emory University, Atlanta, GA.


APPENDIX A

Full tutorial, transfer, mapping quizzes, as well as the tutorials used in Experiments 1–3 are available online (http://jys.bol.ucla.edu/Perspective). The quiz questions from Experiment 2 are provided here as examples (Experiment 1 is a subset of these questions and Experiment 3 has a slightly modified cover story).

Tutorial Quiz Questions for Experiment 2 ("You" Condition)

1. Why do you make mistakes?
   a. The cell distortion evidence overlaps between sick and healthy patients.
   b. When I move my decision boundary, I tend to make more mistakes.
   c. I do not diagnose enough people as sick.
   d. If I tried harder, I could reduce many of my errors.

2. The number of actually healthy and sick people are the same two months in a row. However, in the second month, you are diagnosing more patients as sick when they are actually sick and more people as sick when they are actually healthy. What must have changed in the second month?
   a. You must be diagnosing people with purer cells as sick.
   b. You must be diagnosing people with more distorted cells as sick.
   c. You must be diagnosing more people who are actually sick as healthy.
   d. You must have become better at diagnosing sick people.
3. For a particular level of cell distortion, you know from your experience this month that there is a 50% chance that this level of distortion indicates cancer. What does this mean?
   a. 50% of people with leukemia have this kind of cell.
   b. 50% of all the patients you have seen this month have leukemia.
   c. You have seen equal numbers of people with leukemia and people with distorted cells this month.
   d. **You have seen equal numbers of sick people with this level of distortion and healthy people with this level of distortion.**

4. You are looking into a new blood test for finding distorted cells. How can you find out whether this new test is better than the old one?
   a. You change your decision boundary and diagnose more sick people as sick.
   b. You change your decision boundary and diagnose more pure cells as healthy.
   c. You do not change your decision boundary and diagnose more healthy people as healthy.
   d. You do not change your decision boundary and diagnose more distorted cells as sick.

5. If you move your decision boundary all the way to include even extremely pure cells as evidence for sickness, it means:
   a. you are generally more accurate because you are able to make less errors.
   b. you never diagnose people as sick when they are actually healthy.
   c. you always diagnose people as healthy when they are actually sick.
   d. sick people become more common so you get more experience diagnosing them.

6. This month, each sick person’s cells get a little more distorted while healthy people’s cell distortions do not get better or worse. You do not know this information. If you do not change your decision boundary, how does this change in the population help you?
   a. you increase the number of actually healthy people you diagnose as healthy.
   b. **you decrease the number of actually sick people you diagnose as healthy.**
   c. you increase the number of actually healthy people you diagnose as sick.
   d. sick people become more common so you get more experience diagnosing them.

7. Which of the following decision strategies will ensure that you maximize the number of actually healthy people you diagnose as healthy?
   a. diagnose everyone as healthy.
   b. look more carefully at the cell distortion levels before your diagnosis.
   c. examine the ratio of sick patients with distorted cells to sick patients with pure cells before your diagnosis.
   d. examine the ratio of patients with distorted cells to patients with pure cells before your diagnosis.

8. Which is most likely to lead to inaccuracy in your diagnoses?
   a. Sick people that develop extremely distorted cells.
   b. **Sick people and healthy people have similar distortion levels.**
   c. The people diagnosed as sick have similar distortion levels.
   d. Distorted cells are more common among sick people.

9. Very distorted cells are often caused by protein bundles. Knowing this, your accuracy can:
   a. improve at detecting who is actually healthy and sick.
   b. improve at detecting who is actually sick.
c. improve at detecting who is actually healthy.

d. **not improve based on this information.**

10. Which decisions should you try to maximize to make sure you are accurately diagnosing both sick and healthy populations?

   a. You should focus on increasing the number of sick people diagnosed as sick and reducing the number of sick people diagnosed as healthy.

   b. You should focus on increasing the number of healthy people diagnosed as healthy and reducing the number of healthy people diagnosed as sick.

   c. **You should focus on increasing the number of sick people diagnosed as sick and reducing the number of healthy people diagnosed as sick.**

   d. You should focus on increasing the number of sick people diagnosed as sick.

11. If you set a very high decision boundary where you only diagnose very distorted cells as sick:

   a. You are likely to make many errors in general.

   b. **You are likely to increase the error of diagnosing sick people as healthy.**

   c. You are likely to increase the error of diagnosing healthy people as sick.

   d. You are likely to increase the number of sick people diagnosed as sick.

12. Two patients come into your office today. Patient A seems very weak but has a cell distortion level of 3. The other one, patient B, seems normal but has a distortion level of 6. Over many years, you’ve found that setting the decision boundary to a cell distortion level of 4 or more (range from 1–7) minimizes errors. What should you do?

   a. You should use your decision boundary; then diagnose patient A as healthy and diagnose patient B as sick.

   b. You want to make sure patient A gets treatment so change your decision boundary to include level 3 only for patient A; then diagnose patient A as sick and patient B as healthy.

   c. You want to make sure patient B does not get exposed to harmful radiation for no reason diagnose both patients as healthy without changing your decision boundary.

   d. You should trust your experience as a doctor and diagnose patient A as sick and patient B as healthy because cell distortion should not be the only evidence that you base your decision on.

---

**Tutorial Quiz Questions for Experiment 2 (Both Conditions)**

1. (Use provided graph.) Remember that the farmers of Chanterais do not know which melons are sweet and which are bitter before they export them. They only know how much the melons weigh. The farmers of Chanterais shipped all melons that met the minimum weight of 1500 grams. How many 1750-gram melons did they export?

   a. 200

   b. 400

   c. **600**

   d. 1200
2. (Use provided graph.) Approximately what percentage of all 1000 gram melons (1 kg) are sweet?
   a. 10%
   b. 25%
   c. 33%
   d. 50%
   e. 66%

3. (Use provided graph.) There was a very bitter shipment of melons last year so the townspeople wanted to be extremely careful this year. They set a 1750 gram minimum weight but they do not know which are sweet or bitter. How many melons that weighed 1500 grams were rejected?
   a. 300
   b. 450
   c. 500
   d. 750
   e. 950

4. (Use provided graph.) With the minimum weight for the pluma melon set at 1750 grams, how many bitter pluma melons are rejected?
   a. 400
   b. 950
   c. 1550
   d. 2000
   e. 2600
5. Some of the farmers in Chanterais debate over using a high-tech digital scale in place of their old-fashioned analog scale. What would be evidence that the high-tech scale is a better diagnostic?
   a. Chanterais changes their required weight and exports more sweet fruit.
   b. Chanterais changes their required weight and rejects more bitter fruit.
   c. Chanterais does not change their required weight and rejects more bitter fruit.
   d. Chanterais does not change their required weight and rejects more light-weight fruit.

6. For 1750 gram melons, Chanterais knows from last month that there is a 25% chance that these melons are sweet. What does this mean?
   a. 25% of sweet melons will weigh 1750 grams.
   b. 25% of the melons will be sweet.
   c. 75% of the melons will be bitter.
   d. 25% of 1750 gram melons will be sweet.

7. If Chanterais lowers their minimum weight, which of the following would happen?
   a. They will export more sweet melons and fewer bitter melons.
   b. They will never export bitter melons.
   c. They will export less sweet melons and more bitter melons.
   d. They will export more sweet melons and more bitter melons.
8. In a particular year, there is plenty of rainfall and all the melons get about 250 grams heavier. The prior year Chanterais exported melons that weighed 1500 grams or more. If they do not change their policy:
   a. Chanterais will ship more heavy melons that are sweet and fewer melons that are bitter.  
   b. Chanterais will reject more light melons that are sweet.  
   c. **Chanterais will ship more melons.**
   d. Chanterais will reject more melons.

9. How does this graph support the idea that melon weight is a good predictor of sweet melons?
   a. **There are fewer heavy melons that are bitter than are sweet.**
   b. There are fewer light melons that are bitter than are sweet.
   c. There are fewer light melons than heavy melons.
   d. There are more sweet melons than bitter melons.
   e. There are more heavy melons than light melons.

10. Which decisions should the farmers of Chanterais try to maximize to make sure their melons are of high quality?
    a. They should try to export more sweet melons and reject fewer sweet melons.
    b. They should try to reject more bitter melons and export fewer bitter melons.
    c. **They should try to export more sweet melons and export fewer bitter melons.**
    d. They should try to export more sweet melons.

11. If the town only sells extremely heavy melons:
    a. They will make fewer errors in general.
    b. **They will reject more sweet melons.**
    c. They will export more bitter melons.
    d. They will export more sweet melons.
12. Why does Chanterais export bitter melons?
   a. Because they previously had no quality control procedure.
   b. They do not reject enough melons.
   c. Whenever they change their minimum melon weight, they tend to make more errors.
   d. **Sweet and bitter melons sometimes have the same melon weight.**

Analogy Questions for Experiment 2 (Both Conditions)

*Italics* indicate a structural mapping; ** indicate a semantic mapping.

1425 a. A patient diagnosed as sick but is actually healthy is like what?
   b. A bitter melon that is rejected.
   c. *A bitter melon that is accepted.*
   d. A sweet melon that is rejected.**
   e. A sweet melon that is accepted.

1430 What in the doctor story is most analogous to a heavy melon?
   a. A patient who is sick.
   b. A patient who is healthy.
   c. *A patient with distorted cells.*
   d. A patient with pure cells.**

1435 A melon that is sweet but was rejected is analogous to:
   a. a *sick patient who had been diagnosed as healthy.*
   b. a sick patient who had been diagnosed as sick.
   c. a healthy patient who had been diagnosed as sick.**
   d. a healthy patient who had been diagnosed as healthy.

1440 1. What in the melon export story is most analogous to the sick patient in the doctor scenario?
   a. a *sweet melon.*
   b. A bitter melon.**
   c. An exported melon.
   d. A rejected melon.

1445 2. The patient with leukemia who has been diagnosed as sick is most like:
   a. A melon that is rejected and sweet.
   b. A melon that is rejected and bitter.**
   c. A melon that is exported and bitter.
   d. a *melon that is exported and sweet.*

1450 3. An exported melon is like:
   a. a patient who has been given low distortion test results.
   b. a patient who has been given high distortion test results.
   c. *a patient who has been given a sick diagnosis.*
   d. a patient who has been given a healthy diagnosis.**
# APPENDIX B

## % Correct Results for Each Quiz Question

**TABLE A1**

Accuracy on Each Question Broken Down by Experiment and by Condition Within Experiment

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experience</td>
<td>“You”</td>
<td>Generic</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>Perspective</td>
<td>Specific</td>
</tr>
<tr>
<td></td>
<td>Perspective</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tutorial Questions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.93</td>
<td>1.0</td>
<td>.54</td>
</tr>
<tr>
<td>2</td>
<td>.61</td>
<td>.50</td>
<td>.54</td>
</tr>
<tr>
<td>3</td>
<td>.61</td>
<td>.69</td>
<td>.64</td>
</tr>
<tr>
<td>4</td>
<td>.41</td>
<td>.38</td>
<td>.32</td>
</tr>
<tr>
<td>5</td>
<td>.71</td>
<td>.56</td>
<td>.79</td>
</tr>
<tr>
<td>6</td>
<td>.37</td>
<td>.31</td>
<td>.36</td>
</tr>
<tr>
<td>7</td>
<td>.34</td>
<td>.0**</td>
<td>.39</td>
</tr>
<tr>
<td>8</td>
<td>.95</td>
<td>.88</td>
<td>.89</td>
</tr>
<tr>
<td>9</td>
<td>.63</td>
<td>.63</td>
<td>.50</td>
</tr>
<tr>
<td>10</td>
<td>.29</td>
<td>.26</td>
<td>.29</td>
</tr>
<tr>
<td>11</td>
<td>.68</td>
<td>.63</td>
<td>.58</td>
</tr>
<tr>
<td>12</td>
<td>.29</td>
<td>.19</td>
<td>.52</td>
</tr>
<tr>
<td><strong>Transfer Questions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-Exp. 1</td>
<td>.37</td>
<td>.79</td>
<td>.84</td>
</tr>
<tr>
<td>1-Exp. 2, 3</td>
<td>.44</td>
<td>.61</td>
<td>.68</td>
</tr>
<tr>
<td>2</td>
<td>.39</td>
<td>.13*</td>
<td>.61</td>
</tr>
<tr>
<td>3</td>
<td>.51</td>
<td>.31</td>
<td>.57</td>
</tr>
<tr>
<td>4</td>
<td>.34</td>
<td>.38</td>
<td>.32</td>
</tr>
<tr>
<td>5</td>
<td>.54</td>
<td>.69</td>
<td>.79</td>
</tr>
<tr>
<td>6</td>
<td>.10</td>
<td>.06</td>
<td>.61</td>
</tr>
<tr>
<td>7</td>
<td>.34</td>
<td>.25</td>
<td>.57</td>
</tr>
<tr>
<td>8</td>
<td>.59</td>
<td>.75</td>
<td>.75</td>
</tr>
<tr>
<td>9</td>
<td>.53</td>
<td>.56</td>
<td>.55</td>
</tr>
<tr>
<td>10</td>
<td>.32</td>
<td>.22</td>
<td>.29</td>
</tr>
<tr>
<td>11</td>
<td>.79</td>
<td>.67</td>
<td>.84</td>
</tr>
</tbody>
</table>

*Note.* *Significant difference between the conditions, *p* < .05; ** *p* < .01.