EDUCATIONAL PSYCHOLOGIST, *47*(3), 177–188, 2012

ISSN: 0046-1520 print / 1532-6985 online
DOI: 10.1080/00461520.2012.695709

Routledge
Taylor & Francis Group

Seeing Deep Structure From the Interactions of Surface Features

Michelene T. H. Chi
Mary Lou Fulton Teachers College
Arizona State University

Kurt A. VanLehn
School of Computing, Informatics, and Decision Systems Engineering, Ira A. Fulton Schools of Engineering
Arizona State University

Transfer is typically thought of as requiring individuals to "see" what is the same in the deep structure between a new target problem and a previously encountered source problem, even though the surface features may be dissimilar. We propose that experts can "see" the deep structure by considering the first-order interactions of the explicit surface features and the second-order relationships between the first-order cues. Based on this speculative hypothesis, we propose a domain-specific bottom-up instructional approach that teaches students explicitly to focus on deriving the first-order interactions cues and noticing the second-order relationships among the first-order interaction cues. To do so, researchers and instructional designers need to first extract from experienced solvers or experts how they derive such first-order cues. Transfer is assumed to be based on the similarities in the second-order relationships, which are familiar everyday relationships such as equal to, greater than, and so forth.

Transfer can be broadly construed as the ability of individuals to "treat" a new concept, problem, or phenomenon as similar to one(s) they have experienced before. The term "treat" can be used broadly to refer to performing various tasks such as categorizing, deciding, diagnosing, explaining, identifying, learning, problem solving, and analogical reasoning in or across different contexts, concepts, problems, patients, phenomena, or situations. For example, if students have learned to explain one phenomenon correctly, can they then explain another similar phenomenon, thereby exhibiting transfer? Transfer can also be construed more specifically as consisting of two sets of processes: *initial learning* followed by *reusing* or *applying what was learned*. In this article, we propose a new hypothesis for why transfer often fails in the classic "two-problem transfer paradigm," and we suggest an instructional approach that may remediate the "failure-to-transfer phenomenon." We begin by briefly describing these terms next.

Correspondence should be addressed to Michelene T. H. Chi, Mary Lou Fulton Teachers College, Arizona State University, P.O. Box 871908, Mail Code 1908, Tempe, AZ 85287-1908. E-mail: Michelene.Chi@asu.edu

THE FAILURE-TO-TRANSFER PHENOMENON IN THE TWO-PROBLEM TRANSFER PARADIGM

The majority of research on transfer focuses on the procedural task of problem solving, such as, if students have learned to solve one algebra problem, are they able to transfer that knowledge by solving another algebra problem that has the same underlying deep structure? In this section, we briefly describe the *failure-to-transfer phenomenon* in this classic context which has sometimes been referred to as the *two-problem transfer paradigm* (Lave, 1988).

In the typical two-problem transfer paradigm, students are first asked to attempt to learn and solve a source Problem A (such as learning by reading a solution or worked-out example to Problem A, or learning by solving Problem A successfully), and then students are asked to solve a novel target Problem X that has the same deep structure as problem A. By and large, students can successfully solve Problem X only if the surface features are very similar to the surface features of Problem A (they basically copy the same procedure). That is, transfer almost always occurs when the source problem and the target problem are similar in their surface features (Reed, Demster, & Ettinger, 1985). In the context of algebra word problems, surface similarity means that the two problems describe the same surface situation (or are literally similar), such as involving two cars traveling at different speeds, with the goal of finding out how many hours it will take the second car to overtake the first car. So a source problem may list a particular speed and distance, whereas the target problem lists a different speed and/or distance. (We elaborate next the operational definitions of surface features and deep structures, as used in the literature.)

However, the majority of studies using variations of the classic two-problem transfer paradigm show that transfer fails when the two problems are dissimilar at the surface level but similar at the deep structural level. The typical finding seems to be that after students have succeeded in solving a source problem, they cannot then solve a target problem successfully that is slightly different at the surface level (Catrambone & Holyoak, 1989; Gick & Holyoak, 1980, 1983; Reed et al., 1985; Ross & Kennedy, 1990) even though they have the same underlying deep structure.

The point is that regardless of whether one takes a cognitive or an alternative perspective on transfer (see Reed, in press), this *failure-to-transfer phenomenon* needs to be explained rather than dismissed as a contrived lab-based paradigm, especially because many standard texts used in school curricula adopt this mode of assessing transfer. For example, students are often asked to learn how to solve problems by studying worked-out solution examples either presented in their texts or worked-out by their teachers at the whiteboard, and then they are either tested for their understanding or asked to practice as homework by solving additional problems (provided either at the end of the chapter or by the teachers) that are either similar or dissimilar in the surface features. Therefore, this traditionally used laboratory-based two-problem transfer paradigm is a somewhat authentic paradigm. In short, the failure-to-transfer phenomenon is ubiquitous in school settings and needs to be explained.

The failure-to-transfer phenomenon is surprising because the assumption is that students should have been able to "see" (as experts can) that the target Problem X and the source Problem A are similar at a deep level, thus allowing them to retrieve the procedure learned from and used in source Problem A, then apply this retrieved (or could be modified) procedure to the target Problem X. In contrast to the failure-to-transfer phenomenon in problem-solving research, in categorization research, it has been shown that participants can categorize a new instance successfully even when there is no visual similarity with a prior instance, but there is some underlying theory-based similarity. For example, if we know how intoxicated people tend to act, then we might categorize a fully clothed man jumping into a swimming pool as a drunk (Murphy & Medin, 1985), especially if the context is a pool party with lots of available alcohol, even though this particular intoxicated person's behavior does not resemble that of other drunks one has seen. Thus, participants are usually able to "transfer" by categorizing "a clothed man jumping into the pool" as the same as "a man talking loudly and wildly in a bar," even though they look and act very differently and are seen in different contexts.

Thus, broadly conceived, the processes of transfer require that students abstract or understand the deep structure of the first problem and then recognize that the second problem also has the same deep structure, therefore the procedure associated with such a deep structure then applies. Thus, for the purpose of understanding our hypothesis on instruction and learning, we simplify transfer in problem solving as constituting these two broad sets of processes (as stated earlier): The first set of processes can be called *initial learning,* and the second set can be referred to as the processes of *reusing* or *applying* what had been learned. Before we can entertain hypotheses for the failure-to-transfer phenomenon, we need to describe the difference between surface features and deep structures, as used in the literature.

Surface Features Versus Deep Structures

Researchers pretty much agree that "surface features" refer to literal objects, concepts, or entities explicitly described in a problem statement or a situation, sometimes also referred to as the cover story. However, many different ways of defining "deep structure" are offered in the context of different studies and different domains. In the problem-solving literature, deep structure often refers to the procedure for solving a problem. For example, in probability problems, whether a problem involves a combination principle or a permutation principle has been considered the deep structure, whereas the cover stories that described the problems in the context of marbles and cars are usually considered the surface features (Ross, 1987). Thus, traditionally, for problem-solving research, if two problems or solution steps are generated by the same rule, then they share the same deep structure. In such a definition, a rule can be an equation such as $\text{rate}_1 \times \text{time}_1 = \text{rate}_2 \times \text{time}_2$ (Reed et al., 1985) for a math distance problem, or a principle such as Newton's Second Law for physics problems (Chi, Feltovich, & Glaser, 1981).

Besides rules, there are many other ways to define deep structure. One other way to define a deep structure is in terms of a schema. In the classic Gick and Holyoak (1980, 1983) studies, when a problem solver successfully solved the fortress problem (in which a general needs to capture a fortress by dividing the army and converging on the fortress from many sides), the hope was that the solvers had induced a "convergence" schema that could be reused to solve a new tumor problem (in which rays can be divided into weaker strengths so they will not kill healthy tissue, yet they can converge and destroy the tumor). Another alternative way to define deep structure in problem solving is to consider not just rules that generate a solution (such as a formula to compute density) but rules that are more conceptual and abstract, such as learning rules that *density is invariant under transformation* (Schwartz, Chase, Oppezzo, & Chin, 2011).

In non-problem-solving research, deep structure also has been defined in other ways. For example, in understanding stories, the surface features would refer to the setting and objects, whereas the deeper structure would refer to the structure of the causal plot (Gentner, Rattermann, & Forbus, 1993). In learning studies, deep structure has been defined as the mental models that students have constructed. For example, in learning about the human circulatory system (Chi, de Leeuw, Chiu, & LaVancher, 1994), students' deep understanding can be assessed as the correctness of the mental model that can be depicted to represent their understanding of the circulatory system. Similarly, in work on students' learning of science processes, two schemas were described as relevant to students' understanding: an "emergent" schema and a "direct" schema. These schemas would constitute the deep structures for various science processes (Chi, Roscoe, Slotta, Roy, & Chase, 2012).

In categorization research, deep structure has also been defined in terms of the first-order relationships between two objects, things, locations, arguments, or more generally, two entities (Gentner & Kurtz, 2006; Gentner & Markman, 2006). Thus, two objects have the same deep structure if they invoke the same relationship even though they can be superficially dissimilar. For example, a bridge or bridging is a relationship that connects two things, and the things can be two locations, two concepts, or two entities. This means that a wooden plank has the same relational structure (because it connects two locations) as a dental bridge (that connects two teeth). Thus, a dental bridge and a wooden plank share this relational similarity whereas they share little, if any, intrinsic similarity, or similarities of features of the entities. An intrinsic feature of the wooden plank bridge might be that it is made of wood or it is wide enough for people to walk on, whereas an intrinsic feature of a dental bridge is that it permanently joins adjacent teeth or dental implants.

In sum, there are many ways to identify and define the deep structure of concepts, categories, entities, problems, phenomena, and situations. Its main difference from surface features is that the surface features can usually be perceived whereas the deep structures often cannot be directly perceived.

Lacking Deep Initial Learning of the Source Problem

One obvious explanation for the discrepancy in the failure of transfer in the context of the two-problem transfer paradigm compared to the success of transfer in other tasks such as categorization is that initial learning was not deep. Let's call this the *lacking-deep-initial-learning* hypothesis. There is agreement among multiple perspectives (such as a cognitive, a situative, or an embodied perspective) and multiple researchers on this explanation. For example, from a cognitive perspective, Ross (1987) stated,

> Assume that novices are trying to make an analogy between the current and past problem, but that they *do not have a good understanding of the appropriate problem structure* [emphasis added], . . . in this case, novices may rely on superficial similarities of the problems to decide how to set up the correspondences between problems. (p. 630)

From a situated perspective, Engle (2006) said, "First consider the crucial issue of whether students' *initial learning* [emphasis added] of the relevant content was successful enough to provide a substantive basis for them to have transferred what they learned to new contexts" (p. 453). Similarly, Lobato (2006) suggested that "transfer from the actor-oriented perspective is the influence of learners' prior [*initial learning*] activities on their [*subsequent*] activity in novel situations" (p. 437). But more explicitly, Lobato's opening sentence stated that "a central and enduring goal of education is to provide learning experiences that are useful beyond the specific conditions of *initial learning*" (p. 431). Likewise, Marton (2006) stated that "transfer is about how what is learned in one situation [*initial learning*] affects or influences what the learner is capable of doing in another [*subsequent*] situation" (p. 499). In short, regardless of perspectives, it is commonly assumed that transfer first requires deep initial learning of the source Problem A before one can expect successful transfer to the target Problem X.

This lacking-deep-initial-learning hypothesis should predict that if students did learn the source problem deeply, then they would be able to transfer more successfully than if they did not learn it deeply. One way to substantiate this hypothesis is to consider how participants were asked to learn the source problem. In many classic studies using the two-problem transfer paradigm, it becomes apparent that students were not required to learn the initial problem deeply. For example, students were often asked to either *read* a provided solution to the source problem (Gick & Holyoak, 1980) or *attempt to* solve and then *read* a provided solution (Reed et al., 1985). Reading a solution clearly does not imply that students have understood the example. In a framework that Chi (2009) introduced that differentiates various overt modes of engagement with learning materials as either *passive* (*receiving* information only), *active* (*selecting* information for emphasis), *constructive* (*generating* new information), or *interactive* (*collaboratively constructing or co-constructing* new information), reading is considered a *passive* overt engagement activity in which students only receive information, whereas generating a solution or self-explaining each solution step is considered a *constructive* engagement activity, which engenders greater learning. This ICAP framework (with the ICAP hypothesis that *interactive-I* mode of learning is better than *constructive-C*, which is better than *active-A*, which in turn is better than *passive-P*) thus suggests that reading a solution example alone is not a good enough learning activity to engender deep learning (Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

The ICAP hypothesis can also interpret the results of Gick and Holyoak's (1983) third experiment. The results of the third experiment show that students manifested greater transfer if they *provided their own solution* to the source problem rather than merely *reading* the solution to the source problem, because reading a solution is a *receiving* mode of learning, whereas providing a solution is a *generative mode.* Therefore, providing their own solutions led to deeper initial learning, and deeper initial learning did foster greater transfer, thus providing direct evidence in support of the lacking-initial-deep-learning hypothesis.

The preceding paragraph suggests that deep learning cannot be guaranteed by having students *read* the solution steps. Deep learning also cannot be accurately assessed merely by seeing if students can solve the first problem successfully, because the correct solution can arise from copying an example solution, or from retrieving a similar solution. Both would be classified as the *active* mode of learning, in which students *copied* a solution or *retrieved* the steps of a prestored solution (neither of which is *generative*). There is considerable evidence showing that being able to solve a problem without error (i.e., knowing the procedure) does not necessarily imply that one has understood the deep structure of a problem. For example, Kim and Pak (2002) found no correlation between problem-solving ability and understanding of concepts such as force and acceleration after students have solved on average 1,000 mechanics problems. Similarly, Nurrenbern and Pickering (1987) found a lack of relationship between solving numerical chemistry problems and understanding the molecular concepts underlying the problems. Thus, the results of many studies suggest that successfully solving an initial problem (in a transfer paradigm) so that students have acquired the procedure for solving an initial source problem does not guarantee that students have actually learned the underlying deep structure of the initial problem, such as the principle or concepts underlying the procedure.

Consistent with the lacking-deep-initial-learning hypothesis, studies that have directly assessed students' initial understanding of the source problems also have shown conclusively that the representation (or understanding) that students had of the initial problem determined the degree of transfer. This was shown clearly in Gick and Holyoak's (1983) fourth experiment. In that experiment, they provided two source stories and asked participants to write down ways in which the two stories were similar, and then the quality of their descriptions were rated in terms of good, intermediate, or poor. Good means that the basic idea of convergence was present. The results were very pronounced and clear: 91% of the participants who wrote a good description could solve the target problem even without a hint, as compared to 30% of the participants who wrote a poor description. This result seems unequivocal in showing that if participants understand source problems deeply, then they are more able to transfer. Similar results were also obtained by Novick and Holyoak (1991). Thus again, there seems to be direct evidence in support of the lacking-deep-initial-learning hypothesis.

As just pointed out, because there is general agreement among researchers that failure-to-transfer reflects a lack of deep initial learning, and there is evidence to show directly that deeper initial learning leads to greater transfer, the problem of failure-to-transfer becomes a problem of how to foster deeper initial learning. The next section focuses on methods that have been used to foster deeper initial learning.

Instructional Methods That Fostered Deeper Initial Learning Successfully

Many well-known concrete instructional methods that succeeded in fostering transfer did so basically by enhancing initial learning. One method is to provide two source examples, and furthermore to ask students to explicitly compare and contrast the two source examples so that they are more likely to be able to abstract the underlying structure (Gick & Holyoak, 1983; Schwartz et al., 2011). Another method is to ask students to self-explain a worked-out solution example of the source problem (Chi et al., 1989). A third method is to require students to identify, for every written step in a solution, the deep principle that generated it (Min Chi & VanLehn, 2010; VanLehn & Chi, 2012). In all these cases, transfer successes improved.

The successes of these methods at enhancing transfer can be explained generally by the ICAP framework, in that the students were asked to be more *constructive* in the initial learning processes. Recall that *constructive* activities refer to activities undertaken by students that generate knowledge beyond the information given, such as drawing a diagram when none was provided by a problem, self-explaining a worked-out solution, identifying the deep principle that justifies every step, or comparing and contrasting two problems, as such comparisons and contrasts produce similarities and differences that were not explicitly stated in the two source problems (Chi, 2009). Constructive activities typically foster deeper understanding and learning, which then lead to greater transfer.

Despite the successes of these general instructional methods that encourage *constructive/generative* activities, these instructional approaches do not explain how deep initial learning is achieved. More specifically, these constructive/generative methods do not address one crucial dilemma, which is that experts or experienced problem solvers can "see" the deep structure of a problem but novice solvers cannot. Being able to see the deep structure obviously enhances transfer. Might there be a more specific instructional method that can enhance students' deep initial learning to the point that they can "see" a problem's deep structure and thereby transfer? We address this dilemma next.

HOW CAN EXPERTS "SEE" THE DEEP STRUCTURE?

The dilemma is, when given the exact same problem statement, experts can "see" the deep structure, whereas novice learners cannot (Chi et al., 1981). We assume that the ability to "see" the deep structure of a problem is an outcome of having learned the materials with deep understanding. The question is, how can experts "see" the deep structure and how can instruction facilitate novice students to "see" the deep structure as well? In this section, we explore an alternative hypothesis and suggest how it might translate to instruction. We start by describing novice students' success at "seeing" the relevant surface features, in contrast to their inability to "see" the deep structures.

Students Can See the Relevant Surface Features

Although everyone agrees what surface features refer to in a problem statement, there is very little discussion about whether students can pick out the *relevant* from the less relevant or *superficial* surface features because obviously not all surface features are relevant or important for the problem solution. Suppose a math problem is about how fast a pickup truck is traveling (Ross, 1987); in such a problem, the surface features are the truck, the speed at which the truck is traveling, the time it started, and so forth. Obviously the pickup truck is a superficial surface feature whereas the speed and time are relevant surface features.

We have some evidence to suggest that students are capable of differentiating relevant from irrelevant (or superficial) surface features. The evidence shows that novice students (assuming they have shallower understanding) are just as competent as experts (assuming they have deeper understanding) at distinguishing relevant from superficial surface features. We base this conjecture on the following finding. In Study 8 of Chi, Glaser, and Rees (1982), six expert students (graduate students) and six novice students (who had completed one course in mechanics with at minimum a B grade) were asked to first judge (rate on a 1–5 scale) how difficult each of the 20 physics problems presented to them were (these problems did have different underlying structures in terms of the physics principles), and then to simply circle the key words or phrases that helped them reach their decisions on a problem's difficulty.

The surprising finding was that there were no major differences, in that both the experts and the novices were equally facile at picking out the important or relevant surface features in the statements of the physics problems, such as "horizontal force," "frictionless," and "move together." For 19 of the 20 problems, the expert and novice students circled the same set of words or phrases in the problem statement. Only in 7 of the 20 problems did the experts identify 1.6 additional features, whereas in 13 of the 20 problems, the novices identified an additional 2.1 features as important. In other words, it was not the case that experts picked out one set of surface features as important or relevant, whereas novices picked out a different set of literal features.

This finding presents an important dilemma. That is, if the experts' superiority is in "seeing" the deep structure of problems, but novices were equally facile at identifying relevant surface features, then what allowed the experts to "see" beyond the surface features? We propose a novel hypothesis to address this dilemma.

Deep Structure Is Derived From Perceiving Interactions of Surface Features: A Novel Hypothesis

The hypothesis we propose here is that experts can "see" the deeper cues by considering the *interactions* of the explicit surface features. Again, we use the term "features" here to refer to words, concepts, or phrases that are explicitly stated in the problem and can be picked out successfully by students as relevant (as just described), and the term "cues" to refer to deeper structural concepts, rules, or principles that must be derived, inferred, or computed. This hypothesis is based on insights that can be gleaned from analyzing what domain experts see when they solve problems. The insights are suggested by both the results of Study 7 in Chi et al. (1982; also reported as Study 4 in Chi et al., 1981) and Study 8 (Chi et al., 1982).

In Study 7, two novice students (who had completed a basic college mechanics course with an A grade) and two expert physicists (who had frequently taught introductory mechanics) were asked to state the "basic approach" they would take to solve the same 20 problems (without defining for them what "basic approach" meant). Moreover, they were also asked to state the problem features that led them to their choice of basic approach. In contrast to the task described previously of asking students to circle the relevant key features in the problem statements (Study 8, Chi et al., 1982), for which there were no differences between novices and experts, in this task, there were overwhelming differences between the experts and novices in the key features they verbally cited as contributing to their decisions about the basic approach to the solution of the problems. In fact, none of the key features cited by the novices and experts overlapped. Not surprisingly, novices again mentioned the literal surface features such as objects and concepts that were explicitly stated in the problems, such as "pulleys," "friction," and "gravity." Experts, on the other hand, stated deeper structural cues such as, "it's a before-and-after situation." These have been called *derived cues* (see Table 11 in Chi et al., 1981).

There are a few different ways to characterize the differences in the derived cues and surface features that the experts and novices "see." One general way is to say that the experts identified and "saw" more *process* cues than novices (74 process cues for experts vs. 2 for the novices), and novices saw more *entity* cues than experts (39 entity cues for novices vs. 21 for experts; Chi, 2011). Examples of *process* cues cited by both experts were a "before-and-after situation" and "inelastic collision," and examples of *entity* cues cited by both novices were "friction" and "spring." *Processes* and *entities* are distinct ontological categories that are difficult for students to shift across, thereby reinforcing the difficulty of helping students "see" a process cue (Chi, in press).

However, characterizing the cues that experts see (e.g., "a before-and-after situation") as a *process* cue does not tell us how experts see them. We propose that an additional way to characterize process cues is that they represent *interactions* among the surface features or the literal objects and entities. We now describe five findings that can be reinterpreted as showing the skill of seeing the interactions among features. First, when physics experts mention a cue such as a "before-and-after situation" in the study just described, not only is this cue not explicitly stated in the problem description, but such a cue must be derived from surface features contained in a problem description such as, "The cart starts from rest and rolls down the ramp" and "When the cart reaches the bottom of the ramp, it is moving at 2 m/s." Somehow, the surface features embodied in the mentally conceived situation (or mental model) can generate the energy information needed for two conditions, when the cart was at rest (the initial condition) and when the cart reached the bottom of the ramp (the final condition). Thus, experts can derive the energy information at the initial and final conditions (what is referred to as first-order cues). Once derived, these first-order cues allow the expert to "see" or know that the mechanical energy (kinetic + potential) at one time point is *the same* as the mechanical energy at the other time point. (Later, this equivalence relationship will be referred to as a second-order cue.) The equivalence of the two energy quantities must have led to the cue of a "before-and-after situation," meaning that the energy *before* the cart rolls down is the same as the energy *after* the roll. The point is that it is the interaction between two surface features that yield the first-order and second-order cues. We unpack this example further next.

A second example comes from the domain of chess. When novice chess players see a chess board in the middle of a game, they do see the relevant pieces (such as the Queens and the Kings), where they are located on the board, and maybe which other pieces are adjacent and nearby. However, when an expert sees a chess board in a middle game position, they also see the complicated interactions among the pieces. Chase and Simon (1973, p. 234), for example, showed that an expert player can see "abstract" attack interactions that involve a combination of pieces of the same color converging on the opponent's King position. Novice players presumably do not see these "abstract" attacks that depict complicated interactions. This interpretation is confirmed by the results of an eye movement study (Reingold, Charness, Pomplun, & Stampe, 2001).

This third example is a developmental study that illustrates more directly that seeing the interactions of dimensions leads to more successful problem solving. In Siegler's (1978) study, he asked children of different ages to predict which side of a balance scale would go down when varying weights are placed on pegs at varying distances on two sides of the fulcrum. Younger children tended to make predictions based on either the weight or the distance dimension (i.e., using Rules I and II). However, older children were able to consider both the weight and the distance dimensions (using what he called Rule III), even though they did not know how weight and distance dimensions interacted until they got older, which was when they can use Rule IV. Rule IV usage showed that children could compute the torques on each side by multiplying the amount of weight on each peg by the peg's ordinal distance from the fulcrum. In short, this study showed clearly that success at solving this balance scale problem depended on children's ability to consider the interaction of two dimensions. In using Rule III, even though they could consider only the interactions of weight and distance in a vague subjective way, causing them to simply "muddle through" when the two dimensions conflicted (Siegler, 1978, p. 114), children using Rule III were nevertheless more advanced than children using Rules I or II in that they were more prepared to advance to Rule IV. The point here is that simply considering the interaction between two dimensions, even without knowing exactly how they relate, made these children more advanced than younger children who only considered one dimension at a time.

For a fourth example, we apply the same interpretation to a transfer study described by Schwartz, Chase, and Bransford (in press). In their clown study, middle-school students in both conditions were given the same pairs of problems to either practice applying a formula to find the density of a pair of buses that exhibited the same ratio (the apply condition) or invent a crowdedness (density) index by comparing and contrasting these same pairs of buses (the invent condition). Students in the invent condition exhibited far greater transfer in terms of their understanding of the ratio structure that comprises density. To invent the density index, students in the invent condition must have considered the relationship between the number of clowns and the number of bus compartments (or volume).

Finally, the failure of novice students to consider interactions can be explicitly shown by the finding in Study 5 of Chi et al. (1982). In that study, four physics experts (two college professors, one postdoc, and one 5th-year graduate student) and four novice students (who had completed the introductory physics course with a B grade, using the Halliday & Resnick, 1974, text) were simply asked to summarize Chapter 5 on particle dynamics of Halliday and Resnick (1974). Because this chapter introduced Newton's three laws, all participants mentioned the three laws so that we can compare what they say about each law. The textbook was available to them while they summarized out loud for 15 min. There were no differences between the experts and novices in terms of quantitative measures such as the length of their summaries or the number of quantitative relations mentioned. However, a content analysis revealed important differences. We illustrate with an example of their summary of Newton's Third Law.

In the 1974 version of the text, Newton's Third Law was stated as, "To every action there is always opposed an equal reaction; or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts." We decomposed this Law into five major components with the important concepts italicized as follows: (a) the reaction is *opposite in direction*, (b) the reaction is *equal in magnitude*, (c) action-reaction involves the mutual actions of *two general bodies*, (d) action-reaction are general forces *extended by each body on the other*, and (e) the direction of action-reaction is *a straight line.* Because seven out of the eight participants did not mention component (e), we ignore this fifth component. Among the first four components, we might agree that components (c) and (d) are relevant to interactions, whereas components (a) and (b) are not relevant to interactions (because they refer to notions of *opposite* and *magnitude*). The results show that both the expert and novice students (all eight of them) mentioned component (a), and all four experts and three of the novices (totaling seven) mentioned component (b). However, none of the novice students mentioned either components (c) or (d), whereas three of the four experts mentioned both components (c) and (d). In short, the findings show clearly that to understand this Third Law deeply, one must understand the interaction aspects of the Third Law, as the experts were able to articulate.

Collectively, the findings of these five studies all suggest that to understand some principles or solve some problems successfully, students must consider the interactions among some literal features. Thus, the hypothesis we propose here is that experts can "see" the underlying principle or deep structure of a problem because they can derive the higher order cues based on the *interactions* of the surface features, in which the surface features can be directly perceived (e.g., the weights and distance in the case of the balance scale).

This hypothesis differs from alternative hypotheses that have been proposed. For example, for Siegler's developmental findings, the superior ability of older children to use Rule III and IV had sometimes been attributed to children's mental capacity that develops with age (Case, 1974). In our own earlier work (Study 8; Chi et al., 1982), the fact that experts can "see" the deep structures of problems was hypothesized to arise from experts having acquired causal inference rules that related explicit problem features directly with derived cues and principle, in a linearly causal way, without considering the interactions of the literal features. For example, we had assumed that literal cues such as "frictionless" would lead directly to the derived cue of "not having dissipative forces" and then subsequently to the principle of Conservation of Momentum/Energy, without assuming that "frictionless" may interact with other literal features in the problem. The next section proposes an instructional approach based on our hypothesis.

AN INSTRUCTIONAL APPROACH BASED ON THE HYPOTHESIS OF PERCEIVING INTERACTIONS

Our hypothesis of the importance of focusing on interactions of literal features may suggest alternative methods of instruction. In contrast to the general constructive/generative instructional methods described earlier, such as comparing and contrasting or self-explaining, which are effective overall for all types of tasks and domains, the general constructive/generative instructional approaches can be enhanced in a way that is tailored specifically to our hypothesis. A direct instructional implication of our hypothesis is to teach students explicitly to focus on the interactions among the relevant literal surface features as a way to "see" and understand the deep structure. This requires that students learn to *derive* first-order cues.

Deriving First-Order Interaction Cues

Our instructional proposal can be illustrated in the case of the clown study again (Schwartz et al., in press). In that task, eighth-grade students were told to either apply a formula to find the density of each bus company that carries clowns or invent a density index for the buses. For the apply condition, they were given the formula that density equals the number of objects divided by the volume (which was a given quantity). For the invent condition, students were asked to come up with a procedure to find out how crowded the buses for a company were. An instructional approach that is intermediate between these two approaches (apply vs. invent) is the one proposed here: In addition to being asked to invent a crowdedness index, students receive the hint to look for interactions among relevant dimensions. That is, they would be asked to think about how the relevant dimensions relate to each other. Based on our earlier results showing that novice physics problem solvers could in fact pick out the relevant surface features, we surmise that students will not have difficulty picking out the relevant dimensions (in this case, number of clowns and number of bus compartments), but what they must look for in addition is the way those dimensions interact. (Using Siegler's balance scale task as an example, this hint is analogous to scaffolding students to consider interactions of two dimensions, the clowns and the bus compartments; that is, to push them to consider Rule III.) Although this instructional method should be even more effective than the invent instructions without the hint to look for interactions, the point of the example is merely to illustrate how instruction can draw students' attention to focus on interacting surface features, thus expediting acquisition of first-order cues.

For another example, consider again the cart-ramp physics problem:

> A cart starts from rest and rolls down the ramp. When the cart reaches the bottom of the ramp, it is moving at 2 m/s. Neglecting all sources of friction (air, rolling, etc.), what is the height of the ramp? (Assume that the mass of the cart is 5 kg.)

The surface features are the cart, the ramp, air friction, and so forth, and we assume that novice students are capable of imagining the dynamic situation of a cart gaining speed as it rolls down a ramp. This dynamic image can be conceived of as the student's mental simulation of the situation described in the problem. (Note that because this dynamic situation is relatively simple, students can form an accurate mental model of it readily, in contrast to our earlier study, in which students had to learn to form a complicated correct model of the human circulatory system; Chi et al., 1994.)

We postulate that students will not have as much difficulty learning to ascertain from their mental simulations that there are two time points: The initial time when the cart is at rest at the top of the ramp, and the final time after the cart is moving and is at the bottom of the ramp. The surface features that are relevant at each time point are the cart's mass, velocity, and height. Because the height at the final time is zero (because the cart is at the bottom of the ramp), the energy at the final time is a function of the cart's mass and its velocity (both are given quantities), which are related by the equation $E = m^*h^*g + 0.5^*m^*v^2$. In essence, this equation allows students to derive the first-order interaction cue of the final energy. But even before considering the interactions of the relevant features, students are often required to infer some of their values. For example, students need to infer that the height is zero at the final time because *the cart reaches the bottom of the ramp* and to infer that the initial velocity of the cart is also zero because the *cart starts from rest.* For the initial time, even though the value of height is unknown, students can still be taught to derive it by considering the interactions among the cart's velocity, mass, and height, as specified in the same equation. As indicated earlier, these energies are referred to as the first-order interaction cues, and students can be taught to derive them from the surface features.

The hypothesis is that novice students should be taught explicitly to focus on deriving first-order cues based on the interactions of the surface features described in the problem statement. Such an approach may have other benefits, such as helping students to overlook similarities in the surface features, when the deep structures differ. This is a situation in which two problems look almost identical but in fact have very different deep structures. In such a case, students need to be able to overlook surface similarities. The same instructional approach of deriving interacting cues can facilitate such overlooking. For example, in Study 2 of Chi et al. (1981), physics problems were intentionally designed so that they looked similar at the surface level (the diagram looked similar, the problem statement contained the same concepts and entities and described the same situations) but required different deep principles for solutions. In fact, the entire description of the problem statement was identical with the exception of the final question. For example, both problems began with

> A man of mass M_1 lowers himself to the ground from a height X by holding onto a rope passed over a massless frictionless pulley and attached to another block of mass M_2. The mass of the man is greater than the mass of the block.

One problem then asked students to find *the tension on the rope* (thus making it into a force problem), whereas the other problem asked students to find *the speed* that the man hit the ground (making it an energy problem). Thus the two problems required very different underlying solution principles based on the question that was asked at the end. Experts treated these two problems as different (by sorting them into different categories) whereas novices treated them as the same. The puzzle is, how is it possible that experts can see a different deep structure when the surface features were essentially identical? That is, by what processes do experts "see" the deep structure? When an expert reads "find the tension in the rope," the expert considers the interaction of the surface cue "tension" with the gravitational force on the block, the mass of the block, and the acceleration of the block (i.e., Newton's Second Law applied to the block) and its interaction with the gravitational force on the man, the mass of the man, and the acceleration of the man (i.e., Newton's Second Law applied to the man). On the other hand, when the expert reads "find the speed of the block when it hits the ground," the expert considers the interaction of the surface feature "speed" with the height of the block at ground level (which is zero), the mass of the block, and the total mechanical energy of the block (i.e., definition of mechanical energy applied to the block). In short, depending on which surface cue is in the "find . . ." statement, the expert notices different first-order cues for the whole solution.

To implement the instructional approach of teaching students about cues requires that researchers and instructional designers first elicit from experts the cues that they use, as opposed to our earlier focus on the principles that they used (Chi et al., 1981). Such a knowledge elicitation approach requires more than just collecting and *analyzing problem solving protocols*, as the protocols may not reveal the cues of the experts, because experts can perceive the derived cues for routine problems implicitly without explicitly mentioning or even consciously reasoning about them. Such an elicitation method is also different from merely asking domain experts for their *subjective opinions* on how to solve or how to instruct students to solve problems. Experts' subjective opinions can often be misleading. Although this instructional approach is domain specific in the sense that how experts derive cues must be known before instruction can be designed, the approach itself is applicable to multiple domains for which interaction cues must be derived in order for novice students to "see" the underlying structure.

Are First-Order Cues Sufficient for Transfer?

Having recommended teaching students more explicitly how to derive the first-order cues or quantities that relate the interaction of several surface features, the question now is whether successfully deriving the first-order cues is sufficient for transfer, and if not, what else is needed. Our conjecture is that transfer is based on second-order cues, where second-order cues are relations among first-order cues.

In our cart-ramp problem, once the first-order cues are derived (i.e., the energy at the initial time and the energy at the final time), learners need to infer that they are *equal.* This equality relates two first-order interactions, so it is a second-order cue. There are two ways to infer this second-order cue of equality: a top-down way and a bottom-up way. Experts would know that the Conservation of Total Mechanical Energy principle applies in this problem because there is no friction or heating. That is, the surface feature or condition of *no friction* elicits the Conservation of Total Mechanical Energy principle. And if this principle applies, this means that the second-order relationship of energies at the initial and final times is equal. However, we assume instead that naïve students have not understood this principle; that is, they have not learned that *no friction* is an important *condition* that should invoke the Conservation principle, which in turn dictates the *equality* of total energy at two time points. What they have learned basically is to associate an equation with the name of the principle. Thus, novice students cannot solve this problem in a top-down way.

This assumption that they have not learned all the conditions that invoke a principle is based on our prior analyses of novice students' self-explanations when they studied worked-out physics examples. From their self-explanations, we coded their acquisition of physics principles such as Newton's First and Second Laws (Chi & VanLehn, 1991, p. 94). We found that the acquisition of Newton's Second Law, for example, consisted of learning several unidirectional inference rules, each with specific conditions (including inducing an incorrect inference rule), such as

1. If the forces acting on a body do not sum to zero, then the body will move.
2. If a body is accelerating, then its net force must not be zero.
3. If a body has acceleration, then it must experience a net force.
4. If a body has a net force, then it is accelerating.
5. If a body is not at rest, then the net force will not equal to zero (incorrect).

This analysis suggests that learning a principle such as Newton's Second Law (that the sum of Forces = ma) requires that students first acquire multiple unidirectional inference rules with various specific conditions. Although we surmise that experts treat all these inference rules as simple variations on a principle and know all the conditions that can elicit the correct principle, for novice students, complete understanding of the principle may require acquiring all the various unidirectional conditional rules related to the principle, weeding out incorrect unidirectional rules, then consolidating the remaining correct unidirectional rules into a single bidirectional principle (e.g., F = ma). In short, we assume that novice learners cannot readily invoke the correct principle (e.g., Conservation of Energy) from a condition stated in the problem statement (such as *no friction*), to then conclude in a top-down way that the energies at the two time points are equal.

If novice students have not learned all the conditions that can invoke a relevant principle to apply, then how can they possibly transfer, as transfer is often based on "seeing" the deep principle? With respect to the cart-ramp problem, this means that after having derived the first-order cues of the mechanical energy at the initial and final times, how will novice students know that they should be equal?

We propose that there is a bottom-up way to learn the principle, and that is for novice learners to be taught to *notice* the second-order cues, which consist of simple relationships such as *equal to, greater-than, less-than,* and so on. Continuing with the previous example, suppose novices see how the problem was solved for the height of the ramp (the unknown quantity) by setting the two energies to be equal (such as from studying a work-out example); this allows them to compute the total mechanical energy at the two time points and therefore finding the unknown height quantity. Through such exercises, novices can *notice* that the two expressions are equal. Essentially the outcome of noticing is comparable to applying the Conservation of Energy principle. However, our bottom-up approach provides a way for students to learn to induce and transfer by the second-order relationship of *equality*, without assuming that they already know and can invoke the principle in a top-down way. In fact, we surmise that this bottom-up approach may be the route by which students eventually acquire an operational version of the Conservation of Total Mechanical Energy. Thus, this is a bottom-up approach, in contrast to the experts' top-down approach of applying a principle that is already known and invoked.

To recap, we are proposing that the difficulty for novice learners in transfer is not realizing that (a) they must derive first-order cues from the interactions of objects and entities directly stated in the problem, and then (b) they must look for and notice the second-order cues of equal to, or greater than, and so on. Our assumption is that "seeing" the second-order relationships is not difficult because the relationships are commonsense everyday ones, whereas first-order cues may involve complicated and unfamiliar interactions. Nevertheless, students must be told that they need to notice the second-order cues, after they have learned to derive the first-order cues.

We can provide another example to point out that the nature of second-order cues are familiar everyday relationships such as *simultaneous, independent, all,* and so on (Chi et al., 2012). For example, for many science concepts of processes taught in school, it is often mandatory for learning and correct understanding that students can understand an emergent kind of process (such as *diffusion of ink in water*). One way to understand an emergent process is to notice a set of relevant second-order cues among the first-order interactions. For example, to understand *diffusion flow*, students have to notice that the collisions of molecules (a first-order interaction) can occur *simultaneously* (a second-order relationship). In other words, one pair of molecules colliding can occur *simultaneously* with another pair of molecules colliding. Another second-order relationship is that the pairs of molecules can collide *independently* of each other pair's collisions. Again, we assert that the meaning of these second-order relationships (e.g., *simultaneous, independent*) themselves are not difficult to understand. The problem is that students need to be told to notice these second-order relationships, and to recognize that they are the same across situations or problems, thereby allowing them to transfer.

In sum, although deep structure in a problem can refer to both first-order and second-order cues, we propose that transfer is based on the second-order relationships. But second-order relationships can only be perceived when first-order interactions are derived. Therefore, instruction needs to focus on both deriving first-order interaction cues and noticing second-order relationships.

Similarities and Differences With Analogies

The first-order types of cues described here in the context of learning to solve problems are similar to those described by Gentner (1983) in the context of solving analogy problems or understanding analogies such as "The atom is like the solar system." In this solar system example, the entities in a solar system are the planets and sun. The first-order relationships (statements about entities) are that the planets *revolve* around the sun and that the planets and the sun *attract* each other. The second order relationships (statements about statements) are that the attraction of the planets and the sun *causes* the revolution of the planets around the sun. Thus the second-order cue is a causal relationship, whereas in the case of the cart-ramp example, the second-order relationship is that of *equality* between the initial and final energy, so in both cases, we assert that the second-order relationships (causal or equivalent) are everyday relationships that students are familiar with and can understand. It is simply a matter of *noticing what they are.*

However, there are many differences between learning to solve problems (as in physics) and learning about atoms from analogizing to the solar system. First, the first-order interaction cues of the planet-solar analogy (revolving, attracting) are already known and familiar to students while learning about atoms, whereas students must learn to derive the first order relationship of mechanical energy at both time points.

A second difference between the planet-solar analogy example and the cart-ramp problem-solving example has to do with our assumption about transfer. In analogy, transfer is based on the success of mapping. As Gentner pointed out, when students are told, "The atom is like the solar system," making correct inferences about the atom requires that students realize not only that the electrons are like the planets and the nucleus is like the sun (mapping the entities) but also that they map over the first-order relationships (revolving, attracting) and the second-order relationships (attracting causing revolving). We postulate instead that transfer in problem solving is not based on mapping either the entities or the first-order relationships, as they can be so different between problems, but rather transfer is based on noticing the similarity in the second-order relationships.

To illustrate the role of the second-order relationships on transfer in physics, consider the following problem:

> A block of mass M is dropped onto a spring with a spring constant k. When the spring is fully compressed, the block is H meters below where it started. How much was the spring compressed? Neglect friction and heating of the spring.

To transfer the solution of the cart-ramp problem to this block-spring problem, students must notice that the energy at the initial time equals the energy at the final time for both problems (i.e., the second-order cues are the same for the two problems). From this example, we can see more clearly the differences between problem-solving transfer and analogy transfer. First, in analogy transfer, the nature of the first-order interactions are identical in order for transfer to occur (both situations involve *revolving* or *attracting*), whereas the first-order interactions are not identical in the case of the physics problems; that is, the way energy is computed in the cart-ramp case is quite different from the way energy is computed in the block-spring case. The cart case involves the interaction of the cart's velocity, mass, and height, whereas the spring case involves computing the interaction between the blocks' velocity, the spring's compression distance, and its spring constant.

A second difference between the analogy of the solar system to the atom and the analogy of the cart-ramp system to the block-spring system is that the solar system is more explicitly taught and thus better understood than the artificially constructed cart-ramp situation. Thus, the solar system is often used to teach students about the atom because students have explicitly been taught the solar system's first-order relationship (attracting, revolving) and second-order relationship (attracting cause revolving). In the case of problem solving in physics, and in the cart-ramp problem as an example, students are taught a mathematical derivation of its solution but they are not explicitly taught how they need not only to derive the first-order interactions but also to notice the second-order relationships. Thus, students are not able to use the cart-ramp problem to understand how to solve the block-spring problem.

CONCLUSION

In this article, we present the novel hypothesis that one reason transfer often fails in the two-problem transfer paradigm is that students have not learned to derive deeper first-order cues from attending to the interactions of the surface features and to notice the relationships between the first-order interactions (i.e., the second-order cues). Based on this hypothesis, we propose an instructional approach that focuses on explicitly teaching students how to derive first-order interaction cues and notice second-order relationships. We showed how such cues could be useful in learning to transfer when the surface features are different but the deep structures of second-order cues are the same, and conversely in preventing students from being misled by situations where the surface features are predominately similar but in fact the deep structures are different. Although most methods of teaching for transfer mention surface features and deep structures, we advocate instruction to focus on deriving first-order cues. Because both the first-order and second-order cues can be inferred from percepts, they explain the dilemma of understanding how experts can "see" the deep structures from the same surface features available to novices.

Our instructional proposal can be characterized as more of a bottom-up approach. It is bottom-up in that it focuses on teaching students how to derive interaction cues and to notice second-order relationships between the first-order interaction cues. Once these second-order relationships are noticed, learners can then apply more straightforwardly the relevant equations or principles. Deriving these cues and noticing their relationships characterize what we believe is the skill of "seeing" the underlying structure.

This bottom-up approach is different from the assumptions of a typical top-down approach. A top-down approach takes the perspective of the experts and assumes, for the problem just described, that success is determined by already knowing the principle, which is invoked from specific surface features, such as that *friction is zero implies that the Conservation of Total Mechanical Energy principle applies (thereby the energies at the two points are equal)*. However, we have shown that novice students, even those who have received a grade of B or better, have not yet acquired a complete understanding of the principles. As we previously described, this can be seen by two kinds of evidence: first, in their incomplete acquisition of unidirectional conditional rules related to the Second Law (including the acquisition of incorrect conditional rule), and second, in their incomplete articulation of all the components of a principle (missing especially the components related to interactions of the Third Law). Such evidence suggests that even after having the opportunity to learn the relevant materials (such as Chapter 5 of Halliday & Resnick, 1974) or having completed a course on mechanics, students did not have complete understanding of the principles embedded in the chapter. Therefore, a bottom-up approach of first learning to derive the first-order interaction cues followed by noticing second-order relationships may lead to transfer and subsequently to a more complete understanding of the relevant principle. Thus, we are suggesting that deep structure may be the second-order relationships between the first-order derived cues.

Although the examples used in this article to illustrate our interaction hypothesis focus on the domain of physics, we believe that this approach is generalizable to other domains as well. That is, in many other domains, it is also the first-order and the second-order cues that must be derived and attended to in order to understand problems, analogies, or processes correctly. We illustrated the importance of first- and second-order cues in physics problems, in scientific analogies, and in understanding emergent processes.

ACKNOWLEDGMENTS

We are grateful for the support provided by the Institute of Education Science to the first author (Grant Number R305A110090 and R305A110811) and the National Science Foundation (REESE Grant Number 0910221) to the second author, for the preparation of this article. We appreciate and benefited tremendously from the reviews and comments provided by Sam Day and two anonymous reviewers.

REFERENCES

Case, R. (1974). Structures and strictures: Some functional limitations on the course of cognitive growth. *Cognitive Psychology*, *6*, 544–574.

Catrambone, R., & Holyoak, K.J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 1147–1156.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York, NY: Academic.

Chi, M. T. H. (2009). Active-Constructive-Interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.

Chi, M. T. H. (2011). Expertise and perspective shift. In Y. Li & G. Kaiser (Eds.), *Expertise in mathematics instruction: An international perspective* (pp. 17–39)., New York, NY: Springer Science+Business Media.

Chi, M. T. H. (in press). Two kinds and four sub-types of misconceived knowledge, ways to change it and the learning outcomes. In S. Vosniadou

(Ed.), *International handbook of research on conceptual change* (2nd ed.). New York, NY: Routledge.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.

Chi, M. T. H., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 7–76). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Roscoe, R., Slotta, J., Roy, M., & Chase, M. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science*, *36*, 1–61.

Chi, M. T. H., & VanLehn, K. A. (1991). The content of physics self-explanations. *Journal of the Learning Sciences*, *1*, 69–105.

Chi, Min, & VanLehn, K. A. (2010). Meta-cognitive strategy instruction in intelligent tutoring systems: How, when and why. *Journal of Educational Technology and Society*, *13*, 25–39.

Engle, R. A. (2006). Framing interactions to foster generative learning: A situative explanation of transfer in a community of learners classroom. *The Journal of the Learning Sciences*, *15*, 451–499.

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

Gentner, D., & Kurtz, K.J. (2006). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the lab* (pp. 151–175). Washington, DC: American Psychological Association.

Gentner, D., & Markman, A. B. (2006). Defining structural similarity. *The Journal of Cognitive Science*, *6*, 1–20.

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, *25*, 524–575.

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

Gick, M., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38.

Halliday, D., & Resnick, R. (1974). *Fundamentals of physics* (2nd ed.). New York, NY: Wiley.

Kim, E., & Pak, S. J. (2002). Students do not overcome conceptual difficulties after solving 1000 traditional problems. *American Journal of Physics*, *70*, 759–765.

Lave, J. (1988). *Cognition in practice.* Cambridge, UK: Cambridge University Press.

Lobato, J. (2006). Alternative prespectives on the transfer of learning: History, issues, and challenges for future research. *The Journal of the Learning Sciences*, *15*, 431–449.

Marton, F. (2006). Sameness and differences in transfer. *The Journal of the Learning Sciences*, *15*, 501–538.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

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

Nurrenbern, S. C., & Pickering, M. (1987). Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education*, *64*, 508–510

Reed, S. (2012). Learning by mapping across situations. *The Journal of the Learning Sciences*, *21*, 353–398.

Reed, S. K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *11*, 106–125.

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*, 48–55.

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

Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 42–55.

Schwartz, D. L., Chase, C. C., & Bransford, J. D. (in press.) Adaptive transfer: The tension of routines and novelty. *Educational Psychologist*.

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

Siegler, R. S. (1978). The origins of scientific reasoning. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 109–149). Hillsdale, NJ: Erlbaum.

VanLehn, K., & Chi, Min (2012). Adaptive expertise as acceleration of future learning: A case study. In P. J. Durlach & A. Lesgold (Eds.), *Adaptive technologies for training and education* (pp. 28–45). Cambridge, UK: Cambridge University Press.