Psychological Bulletin
2016, Vol. 142, No. 7, 729–757

© 2016 American Psychological Association
0033-2909/16/$12.00 http://dx.doi.org/10.1037/bul0000043

Relational Categories as a Bridge Between Cognitive and Educational Research

Micah B. Goldwater
The University of Sydney

Lennart Schalk
ETH Zurich

Both cognitive and educational psychology literature strive to investigate human category and concept learning. However, both literatures focus on different phenomena and often use different methodologies. We identify and discuss commonalities and differences between the literatures. This literature comparison reveals that research on relational category learning offers a promising avenue to integration. We suggest that this integration would be especially beneficial to advance our understanding of conceptual change essentially, how complex scientific concepts and categories are acquired and developed in educational contexts elaborating or correcting students' prior conceptions. Furthermore, the focus on relational categories allows us to provide an integrative discussion on how recent lines of research on analogy, memory and category learning, and knowledge restructuring relate to and can inform education. In general, this article advocates the complementary nature of cognitive and educational psychology and identifies viable, and potentially synergistic paths for future research.

Keywords: conceptual change, relational categories, STEM education, analogy, categorization

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Category and concept learning have long traditions in the cognitive and educational psychology research. Despite the similar general aims of understanding how humans learn and reason, cognitive and educational psychologists have largely focused on different phenomena, often using different methods of research, and targeting different communities of readers. The goal of this article is to discuss how cognitive and educational psychologists study category and concept learning to uncover where research approaches have diverged, but more importantly, to also advance a novel integration of the two literatures. We reveal that despite apparent differences, there are great opportunities for each of these literatures to inform the other and that a stronger alignment opens up exciting avenues for future research.

Cognitive psychologists have investigated the learning and representation of categories, and how knowledge of categories serves as the basis for induction and deduction since decades (see Murphy, 2002, for review). Typically, simple perceptual or linguistic stimuli have been used in learning tasks with upward of multiple hundreds of trials. These tasks afford the precise tracking and characterization of learning and performance over time as the learner amasses experience (admittedly, often within the convenient confines of a single hour-long laboratory session). This precise characterization of performance is exceptionally useful as it allows, for example, to develop precise cognitive models to computationally simulate learning processes (e.g., Erickson & Kruschke, 1998; Kurtz, 2007; Love, Medin, & Gureckis, 2004), or to distinguish engagement of multiple brain regions (e.g., Ashby & Maddox, 2005; Davis, Love, & Preston, 2012).

Cognitive psychologists mainly use categories (or category systems) that are entirely artificial. They pursue this approach of learning novel, artificial categories to experimentally control for potential variations and differences in subjects' prior knowledge (though this approach has long been criticized for its incompleteness within the same literature, e.g., Murphy & Medin, 1985, and see below). Any given experiment will typically design the structure of the artificial categories to test the role of component cognitive mechanisms in the learning process, or to test divergent predictions of competing models of category learning. In general, the goal of cognitive psychologists is to reverse engineer acquisition and use of category knowledge to precisely capture the underlying cognitive capacities, mechanisms, and representations.

In contrast, educational psychologists' primary goal is not to reverse engineer, but to understand learning to design improved instructional techniques. In doing so, researchers have focused on materials used (or could be used) in real classrooms to teach complex and educationally relevant topics (e.g., highly abstract, idealized, and generalizable scientific concepts such as "force" in physics). Generally, this complexity makes it difficult to characterize moment-to-moment engagement of cognitive mechanisms as precisely as it has been achieved by cognitive psychologists

This article was published Online First March 7, 2016.

Micah B. Goldwater, School of Psychology, The University of Sydney; Lennart Schalk, Institute of Behavioral Sciences, ETH Zurich.

Australian Research Council Grant DP150104267 supports the research of the MBG. Both authors contributed equally to this work. We first would like to thank editor David Uttal and three anonymous reviewers for their helpful critiques. We also thank Tyler Davis, Arthur B. Markman, Bruno Rütsche, Henrik Saalbach, Michael Serra, and Elsbeth Stern for comments on a draft version of the article, and Jeffrey Loewenstein, Kenneth Kurtz, and Dedre Gentner for influential discussion of these ideas.

Correspondence concerning this article should be addressed to Micah B. Goldwater, School of Psychology, University of Sydney, Brennan MacCallum (A18), Camperdown NSW, 2006 Australia. E-mail: micah.goldwater@sydney.edu.au

(though see, e.g., Bannert, Reimann, & Sonnenberg, 2014; Koedinger, Corbett, & Perfetti, 2012, for advances in this area).

However, instead of ridding the experiment of prior knowledge, a major strength and focus is the characterization of its major influence on learning (e.g., Clement, 1982; Libarkin, 2001, 2008; Opfer, Nehm, & Ha, 2012; Talanquer, 2006). For example, prior knowledge has been shown to greatly affect how learners interpret and learn from their class materials (e.g., Rittle-Johnson, Star, & Durkin, 2009), how well they can implement domain-general problem-solving strategies (see, e.g., Koedinger & Roll, 2012; Zimmerman, 2000, for reviews), how well they are able to generalize knowledge encoded from learning materials to novel cases or problems (e.g., Carpenter, Franke, Jacobs, Fennema, & Empson, 1997), and how well they are prepared for future learning tasks (e.g., Bransford & Schwartz, 1999). Further, the degree with which new concepts conflict with the students' prior knowledge (e.g., their naïve theories) is one of the primary determinants of how difficult a concept is to learn (Chi, Roscoe, Slotta, Roy, & Chase, 2012; Jacobson, Kapur, So, & Lee, 2011).

For cognitive psychologists, a model capturing the relative difficulty of learning different category structures is important evidence that the model is a proper characterization of the learning process (e.g., Kurtz, 2007; Love et al., 2004; Nosofsky, Palmeri, & McKinley, 1994; Shepard, Hovland, & Jenkins, 1961). However, the motivation for education researchers to understand which concepts are difficult to learn means identifying targets for interventions to assist in learning and conceptual change; that is, to study how complex scientific concepts are acquired. Ideally for educators, a cognitive model should inform school learning and aid in designing interventions.

Cognitive and educational psychologists do not only differ with respect to the treatment of prior knowledge but also in the types of categories they have aimed to explain. On the one hand, cognitive psychologists have largely focused on categories that concern the intrinsic features of individual entities. For example, *guitars* have strings, a hollow wooden body, and make music; *dogs* have four legs, fur, and bark. On the other hand, educational psychologists tend to focus on categories that are defined by the relations among entities. For example, the category *catalyst* classifies molecules by their role in effecting changes of state in other molecules; *force* is defined by the multiplicative relationship between mass and acceleration. Critically, in recent years cognitive psychologists' interest in the role of relational categories in everyday thought has increased. For example, the relational category *barrier* is anything that can play a preventative role (that can be a physical barrier, e.g., large rock or a more abstract barrier, poverty); a *pet* is an animal who a person owns to serve a role of companionship (Gentner & Kurtz, 2005; Markman & Stilwell, 2001). Furthermore, cognitive psychologists started to investigate relational categories relevant to education. For example, Goldstone and Sakamoto (2003) investigated the learning of *positive feedback systems* that classify phenomena by their causal structure regardless of content domain (also see Rottman, Gentner, & Goldwater, 2012). However, compared with the focus on feature-based categories, basic and applied research on relational categories is in its infancy. In this article, we purport that a continued focus on relational categories will enable a prolific integration of the (rather) basic cognitive and the (more) applied educational research.

Central to our argument is that the distinction between categories represented either by intrinsic features or by extrinsic relations is not merely a distinction of content. It is a distinction in representational form that has critical implications for mechanisms of learning and reasoning. We will give evidence for and elaborate upon this distinction throughout the article. We acknowledge that cognitive and educational research are aimed to be complementary, with cognitive psychology investigating the mechanisms behind category learning, and educational psychology designing and testing methods to improve learning. However, we argue that without a specific focus on relational categories and knowledge, the degree to which cognitive theories can inform education is ultimately limited. Research into relational categories will foster both the integration of cognitive and educational psychology research, and increase the value and efficacy of cognitive theories for educational application.

Here is a roadmap for this article: We begin by introducing the standard paradigm of feature-based category learning, delineate some fundamental differences between feature-based and relational categories, and argue how relational categorization is a fundamental problem in education. Next, we discuss how theories about analogical mapping and reasoning help to understand challenges in relational encoding and retrieval. Afterward, we introduce the notion of knowledge transfer, provide an integrative review for why transfer of relational knowledge is difficult, and delineate effective strategies to foster transfer. Subsequently, we focus on the cohering nature of relational knowledge, and explain that the development of expertise in many educationally relevant domains requires building up and revising knowledge systems comprised of relational concepts. In the penultimate section, we discuss how studying this process of conceptual change, that is, the challenge of changing one's conceptual knowledge system(s), should, on the one hand, guide the research of cognitive psychologists and would, on the other hand, benefit from rigorous methods developed by cognitive psychologists, before we provide a concluding summary. Throughout, we will use feature-based categorization as a kind of foil, to highlight the importance of relational knowledge specifically, and provide recommendations for future research bridging cognitive and educational psychology.[1]

[1] We reviewed a range of articles when writing this article. It is beyond the scope of the article to provide systematic reviews for all research fields discussed (e.g., feature-based category learning, relational category learning, and conceptual change). Furthermore, several systematic reviews and meta-analyses already exist for these specific research fields and are referenced in this article. However, to ensure that this article indeed goes above and beyond existing research and theoretical suggestions, we conducted systematic literature searches through ERIC (1966–January 2015), PsychInfo (all years to present), and Web of Science (from 1900 to present) using six conjunctions of search terms (* indicate truncations): (a) feature-based categor* AND concept learn*, (b) feature-based categor* AND conceptual change, (c) relational categor* AND concept learn*, (d) relational categor* AND conceptual change, (e) relational cognition AND category learn*, and (f) relational cognition AND concept learn*. The search was conducted on February 17, 2015. None of the articles directly addressed the present articles main theme. The results of the systematic search are available from the authors.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Feature-Based and Relational Categorization in Cognitive Psychology

Theories of categorization typically assume that categories are (a) represented by a set of features that describes the properties of category members, (e.g., Love et al., 2004; Rosch & Mervis, 1975; Smith, Shoben, & Rips, 1974), and (b) that feature-based categories are arranged into hierarchical taxonomies of generality, such as *dog*, *mammal*, and *animal* (e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Theories differ on whether the set of features is organized around a prototype, exemplars of the category experienced in the past, or subclusters of features that describe category members. Classification of new items is typically proposed to be a function of the similarity of the new item to the category representation (e.g., Goldstone, 1994).

As mentioned above, cognitive psychologists have primarily focused on studying the acquisition and use of artificial categories. These categories are defined by sets of precisely describable features. The precision allows developing highly controlled experiments to study how humans represent and learn categories. Furthermore, it allows comparing predictions of different theoretical accounts which are often evaluated by fitting computational models.

The most typical experimental method is the inductive classification paradigm, or simply "classification learning." In this paradigm, exemplars from a small number of categories are presented one a time and subjects indicate which category they believe the exemplar to be a member of (see Figure 1 for a typical example). At first the subjects are merely guessing, but with more trials they start to develop more accurate hypotheses about the mapping of exemplar features to categories. These experiments typically have two phases: (a) the learning phase and (b) the test or transfer phase. In the learning phase, subjects receive corrective feedback about

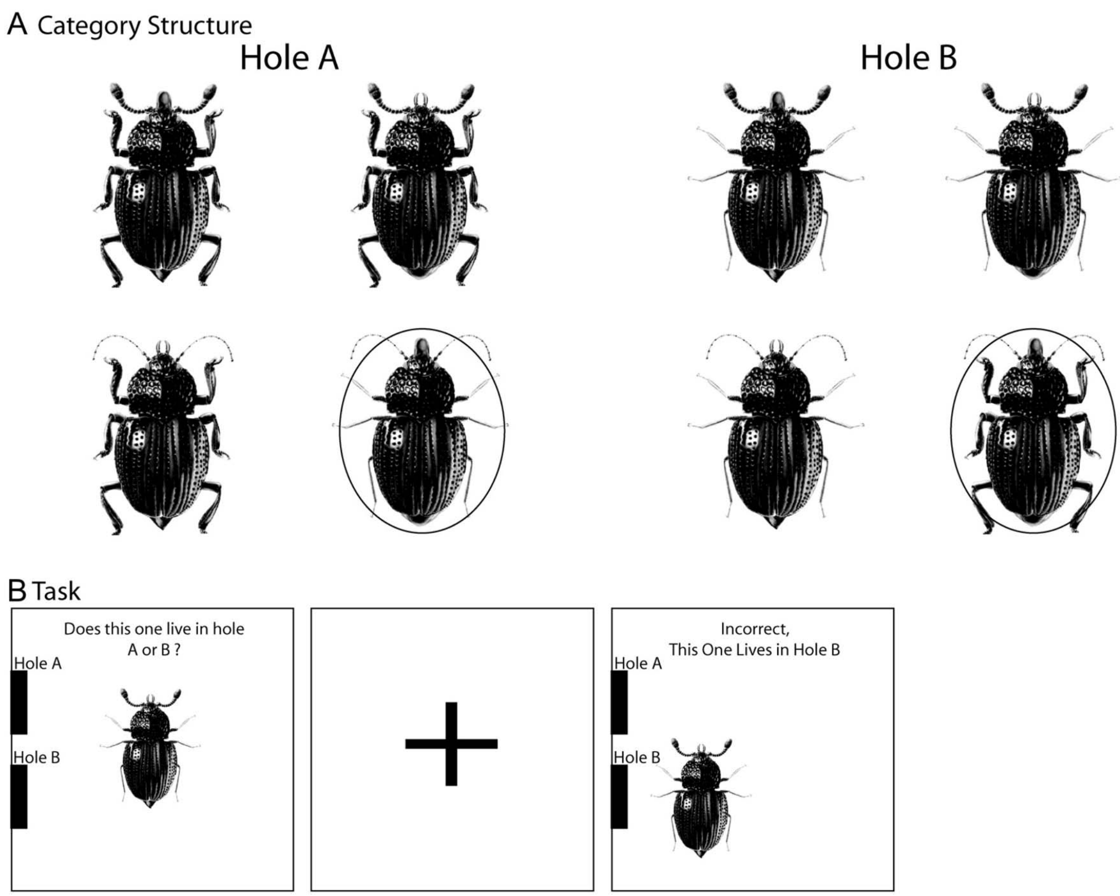

Figure 1. From Davis, Love, and Preston (2012) that examined how computational models of category learning can be used to explain functional magnetic resonance imaging (fMRI) data during category learning. The task was to classify beetles as residents of Hole A or Hole B. The beetles varied along perceptual dimensions such as eyes (green or red), tail (oval or triangular), legs (thin or thick), antennae (spindly or fuzzy), and fangs (pointy or round). For any given subjects, there was a single dimension that could correctly classify [3/4] of the exemplars (here, leg thickness). The circled exemplars violate the rule and appear to be members of the other category. As shown in (B), after subjects were asked to classify a beetle as either Hole A or Hole B, subjects received feedback. From "Learning the Exception to the Rule: Model-Based fMRI Reveals Specialized Representations for Surprising Category Members," by T. Davis, B. C. Love, and A. R. Preston, 2012, *Cerebral Cortex, 22,* p.262. Copyright 2011 by The Authors. Adapted with permission.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

their classifications. Typically, this phase continues until subjects reach a criterion of accuracy. In the test/transfer phase novel exemplars are presented and subjects classify them without receiving feedback. The classification of the novel exemplars serve as a proxy to reveal how the subjects represented the categories they learned.

Within the basic constraints of the inductive classification paradigm, research has revealed many important findings. For example, research has illuminated issues such as the relationship between structure of the learning task and the content of what is learned (e.g., Love, 2003; Markman & Ross, 2003), the role of various component processes (e.g., attention-shifting; Kruschke, 1992), how error feedback shifts representations (Love et al., 2004), and what category structures engage explicit declarative reasoning systems in contrast to implicit procedural learning systems (Ashby & Maddox, 2005). While we will argue that the primary focus on feature-based categories limits the applicability of this research, the basic structure and methodology of the inductive classification task (and related, yet crucially distinct variants; see Markman & Ross, 2003) nevertheless allows for researching critical aspects of concept learning that is directly applicable to education. Further, we will argue that this applicability has so far not been fully exploited. To start this argument, we describe the basic differences between feature-based and relational categories.

In a series of articles, Gentner (1981, 1982, 1983) developed the first line of research in cognitive psychology to focus on the qualitative distinction between the intrinsic features of objects and the extrinsic relations among objects (though this conceptual distinction has its roots in classical Greek philosophy). She theorized that objects and relations comprise a natural partition in the world. Identifying and representing object features is relatively easy because objects are stable over time and space, so they are more easily individuated. This increased perceptual cohesiveness also allows for directly perceiving the similarities among objects, so they are more easily categorized. In contrast, relations among objects are less perceptually cohesive; they are dynamic and often short-lived. For example, compare the stability of a person and a house's existence to the spatial relation of a person being in front of a house. Further, because the specific instantiation of a relation involves concrete objects and situations, recognizing common relations across sets of objects is more difficult than learning about the objects themselves. For example, again compare the spatial/physical barrier of a large rock to the more abstract barrier of poverty, or just consider the spatial/physical differences between two examples of the relation *on* (referring to physical support) in what it means for pants to be *on* (one's legs) and for a painting to be *on* a wall. The increased difficulty leads to two consequences relevant for the current discussion. First, the meanings of relational words are more contextually mutable than words referring to objects (Gentner & Asmuth, 2008; Gentner & France, 1988); and second, relational words are generally learned later in development than words for objects across cultures (Gentner & Boroditsky, 2001).

Building on Gentner (1981); Markman and Stilwell (2001); Gentner and Kurtz (2005) and Goldwater, Markman, and Stilwell (2011) have laid out a cognitive framework that distinguishes feature-based from relational category representations (and see Markman, 1999, for extensive review). The framework further differentiates between two kinds of relational categories: schema-governed and role-governed categories. Schema-governed categories are represented by whole relational systems (e.g., *family*, *catalysis*), and role-governed categories are represented by the roles within such systems (e.g., *aunt*, *reagent*). It is argued that these different kinds of categories require different cognitive mechanisms to be encoded, represented, and reasoned with.

Several empirical studies support the theoretical value of distinguishing between these different categories. For example, Gentner and Kurtz (2005) demonstrated that participant generated exemplars of relational categories are less similar to each other than exemplars of feature-based categories. To illustrate this finding, compare the similarity of a large rock and poverty as examples of the relational category *barriers* to celery and carrot as examples of the feature-based category *vegetables*. Regarding the distinction between role-governed categories and feature-based categories, Goldwater and colleagues (2011) showed that people are more likely to refer to properties extrinsic of category members when listing properties of role-governed categories (e.g., guests are polite to others), while they list more intrinsic properties for feature-based categories (e.g., knives are made of metal). Moreover, the proportion of extrinsic and intrinsic properties listed predicted patterns of labeling and describing images uploaded to a photo-sharing website (flicker.com) that highlights these results' external validity.

Consistent with this framework, recent work in cognitive neuroscience has examined whether and to what extent relational cognition requires additional brain areas on top of the processing individual objects. For example, Cohen (2009) argues that the cortex parses the world into objects, and that the hippocampus is critical for remembering the relations among them. Cohen presents evidence that hippocampal amnesiacs exhibit massive deficits in memory for the relations among objects, yet show little deficit in memory for relations (of parts) within objects. Additionally, increased demands for integrating relations among entities across a variety of tasks is related to increased activation of the frontopolar cortex, compared with reasoning about features (e.g., Christoff et al., 2001), or categories of objects (e.g., Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006), or the general memory demands of holding multiple things in mind (Parkin, Hellyer, Leech, & Hampshire, 2015). Thus, these integration processes seem quite explicit. Indeed, Dominey, Lelekov, Ventre-Dominey, and Jeannerod (1998) demonstrated that the specific elements within a sequence can be learned under implicit task conditions, while explicit reasoning is needed for learning the sequences' abstract relational structure (though see Day & Gentner, 2007; Day & Goldstone, 2011, and discussion below).

Further, people typically do not only reason about single relations, but about relations between relations. Gentner (2003) and Penn, Holyoak, and Povinelli (2008) argue that the ability to represent such structured systems of relations that are composed of not just relations among objects (first-order relations) but by multiple or higher-order relations among the relations is the defining quality of human-unique cognition. In a similar vein, Halford, Wilson, and Phillips (2010) argue that relational knowledge is at the center of all higher-level cognition. Not coincidentally, the frontal pole (aka Broadman's Area 10) is extremely large in humans, potentially the largest single anatomical structure of the frontal lobes (Christoff et al., 2001), and twice the relative size in comparison to chimpanzees (Semendeferi et al., 2001; see integrative discussion in Burgess, Gilbert, Okuda, & Simons, 2006).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Taken together, it should come as no surprise that relational concepts play a central role in education.

Relational Concepts in Education

Most of the concepts and principles taught in schools, colleges, or universities relate several roles or variables to each other. Of course, purely perceptual, feature-based classification systems do play a certain role in education. For example, children learn that wooden objects float, or learn to distinguish different basic level categories of natural kinds in elementary school.[2] Nevertheless, today's focus in education is on acquiring relational concepts, schemata, and complex combinations of concepts (Resnick, 2010) often in the form of formalized representational systems. Central topics dealt with in STEM-education (science, technology, engineering, and mathematics) such as evolutionary theory in biology, Newtonian laws in physics, Coulomb's law in chemistry, but also concepts in the social sciences and humanities like complex systems principles in sociology, the principles of category learning in psychology, grammatical and syntactical rules of languages, or the golden ratio in arts, are all examples of relational concepts that are capable of describing a wealth of superficially dissimilar situations.

The productive use of these kinds of relational concepts necessitates recognition of a new problem or text as an instantiation of the concept. Recognition of disparate novel exemplars is one of education's greatest challenges. In the case of scientific concepts,[3] this recognition process mainly requires learners to identify relational similarities. Sometimes superficial similarity concurs with relational similarity (e.g., two problems about forces affecting the orbits of different planets; see Bassok, Chase, & Martin, 1998; Mayer, 1981). When this is the case, the conjunction supports students' memory retrieval and problem solving (e.g., Mayer, 1982).

Frequently, however, superficial features are not diagnostic for whether or not a scientific concept applies. To make it even more complicated, one often has to look past superficial features to successfully generalize scientific concepts (e.g., to solve a problem about forces on the molecular level). As an example, consider the work of Chi, Feltovich, and Glaser (1981) who asked novices and experts to sort physics problems. The problems could either be sorted based on the underlying physics concepts (e.g., problems about the conservation of momentum, Newton's 3rd law, etc.) or based on superficial similarities (e.g., the presence of pulleys). Experts recognized the underlying concept and sorted the problems accordingly. In contrast, novices classified the problems based on their superficial similarity. Researchers have successfully replicated this difference between novices and experts and have termed the transition from novice to expert performance a "relational shift" (e.g., Chi et al., 1981; Gentner & Rattermann, 1991; Rottman et al., 2012; Stains & Talanquer, 2008). Thus, experts are able to see the deep structure in the interactions of surface features (Chi & VanLehn, 2012) and/or to ignore misleading superficial similarities of the problems. Instead, they rely on the important relations defined in scientific concepts to recognize or organize problems.

In a meta-analysis of interventions to improve scientific problem solving, Taconis, Ferguson-Hessler, and Broekkamp (2001) showed that successful interventions focused on improving the schematic representations of problem structures. These interventions "allow patterns or configurations to be recognized as belonging to a previously learned category and which specify what moves are appropriate for that category (p. 446, as quoted from Sweller & Cooper, 1985, p. 60)." In contrast, unsuccessful interventions focused on general problem-solving strategies while ignoring domain-knowledge and problem representation. Often the challenge for novices is not how to solve a problem, but how to classify a problem which would allow the novice to apply the right problem-solving procedure (e.g., Quilici & Mayer, 1996; Rohrer & Pashler, 2010).

Taken together, most of the scientific concepts dealt with in educational settings can be conceived of as relational categories. Even the most basic scientific concepts studied early in educational careers like evolutionary theory in biology or basic concepts in physics (like floating and sinking) require that learners develop highly flexible relational knowledge representations to be able to successfully use these concepts. To become expert in a scientific domain, learners need to acquire a highly interrelated set of concepts and principles that classify phenomena, problems, and situations by their deep (common) relational structure and not (only) by superficial features.

Analogical Mapping and the Challenges of Relational Learning

How do people find commonalities in relational structures? The prevailing theories were first developed to explain how people understand analogies (Gentner, 1983; Holyoak & Thagard, 1989). In these theories, analogies are defined as comparisons based on commonalities in relational structures between two analogs (e.g., objects, situations, concepts, and so on). Consider the oft-cited analogy of the Rutherford atomic model (see Figure 2): the atom was thought to be like the solar system because there are several smaller objects (electrons and planets) revolving around a single larger object (nucleus and the sun) in both. Commonalities are determined by a mapping process. Researchers refer to this process as "structural alignment" (Gentner & Markman, 1997).

Aligning relational structures is a complex process that is governed by three primary (structural) constraints determining the quality of the match (also see Holyoak & Thagard, 1997, for discussion of goal-oriented constraints). First, the "one-to-one mapping-constraint" governs that one representational element can only be put into alignment with exactly one other representational element from the other relational structure. Second, the "parallel-connectivity-constraint" governs that objects are put in correspondence based on their relational role. Accordingly, the first, second, and so forth roles of the first relational structure are placed in correspondence with the first, second, and so forth roles of the second relational structure (respectively). Third, the

[2] Sakamoto and Love (2010) give a convincing demonstration of how the feature-based category learning literature can be applied to aid primary schoolchildren's learning of different biological categories, such as types of sharks.

[3] Even though, relational concepts and categories can be found in all kinds of educational arenas, the majority of research in educational psychology is focused on learning STEM concepts. Thus, we will use the generic term "scientific concepts" in this article to refer to this kind of research.

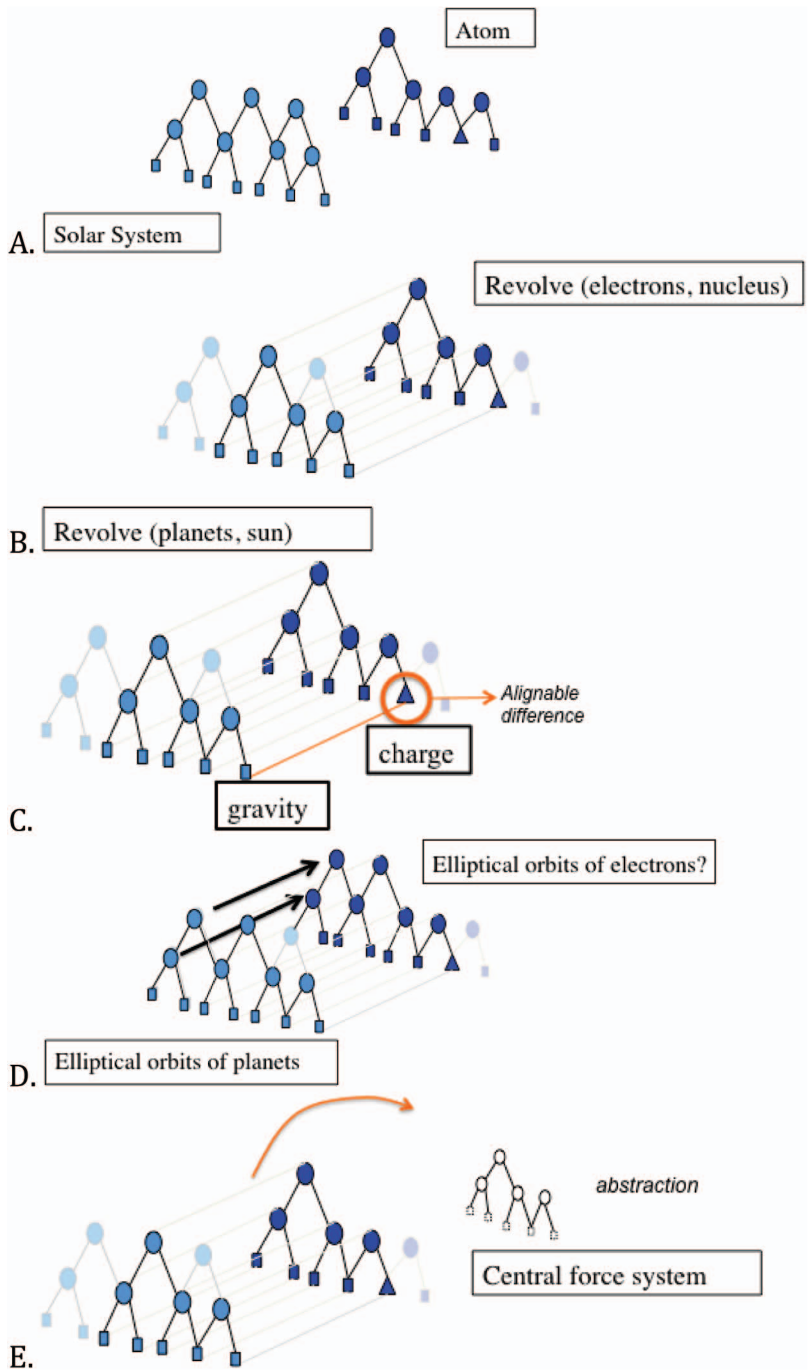

Figure 2. Comparison as Structural Alignment, using the "atom is like the solar system" analogy. (A) Knowledge represented as hierarchical relational structures: elements are bound by how they relate. (B). Comparison highlights structural commonalities. (C) Comparison highlights structure-relevant or "alignable" differences. (D) Candidate inferences are proposed for target structure from base to then be evaluated. (E) Structural commonalities are abstracted to form a new relational concept. From Gentner, D. & Smith, L. (2012). Analogical Reasoning. In V. S. Ramachandran (Ed.) *Encyclopedia of Human Behavior* (2nd Ed.). p. 132. Oxford, UK: Elsevier. Copyright 2012 by Elsevier. Adapted with permission. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

"systematicity-constraint" governs a preference for matches between deeply nested relational structures. Nested structures are structures which consist of higher-order relations that govern corresponding lower order relations. The Rutherford atomic model displays all three constraints, one-to-one mapping, parallel-connectivity, and systematicity. Specifically, electrons map only on the planets, the nucleus maps only on the sun, and the alignment of the nucleus with the sun is based in the higher-order causal relation that links two lower-order relations. Namely, that the sun is *more* massive than the planets, and that this difference in mass *causes* the planets to *revolve* around the sun. Research has repeatedly shown that learners can project analogical inferences across domains when the correspondence between the analogs is based in matching highly systematic relational structures (e.g., Clement & Gentner, 1991; Markman, 1997). This systematicity in their correspondence is more important for inferential power than overlap in objects or domain-specific features (e.g., Gentner, Rattermann, & Forbus, 1993).

The notion of systematicity highlights a crucial point about relational concepts: relational knowledge is rich and full of content. It may be tempting, because of relational categories' more abstract nature (e.g., Gentner & Kurtz, 2005), to think relational categories are just like feature-based categories, but with all the perceptual stuff thrown out. However, this is not the case. First, consider that feature-based categories themselves differ in abstractness, for example, *dog*, *animal*, and *thing*. Second, the natural imperfect correlations between abstractness and relationality do not explain the results distinguishing the feature-based and relational categories. Gentner and Asmuth (2008; and see Jamrozik Sagi, Goldwater, & Gentner, 2013) equated their lists of relational and feature-based categories for imageability. While the relational categories were rated as reliably less imageable than the feature-based categories in Goldwater and colleagues (2011), a closer analysis is quite telling. As a reminder: more extrinsic properties were listed for relational categories, while more intrinsic properties were listed for feature-based categories. The imageability ratings mediated the difference in the number of intrinsic properties listed between relational and feature-based categories, but the imageability ratings did not mediate (i.e., were unrelated to) the difference in number of extrinsic properties listed. That is, increasing abstractness in the feature domain seems to just be about shedding more and more features (e.g., all animals share fewer features than all dogs). However, relations are not just about a lack of stuff, they are conceptual content in and of themselves.

This conceptual richness is what gives relational concepts their power. Finding across-domain relational commonalities fosters novel understanding and insights. For example, when recognizing a commonality in the feedback-based causal structures in the climate and in the economy, this can increase understanding of each area (e.g., Rottman et al., 2012). Likewise, in "everyday cognition," Sagi, Kaufmann, and Clark (2009) showed that the meanings of relational words have changed more over history than meanings of feature-based nouns, and Jamrozik and colleagues (2013) demonstrated that relational words have more conventionalized metaphoric uses than feature-based nouns. That is, these findings show that relational meanings can be extended and applied to new domains and contexts, similar to how scientific principles can be generalized to new problems, phenomena, or domains. However, this conceptual power comes at a cost to the learner.

Challenges in Relational Encoding

Structure-mapping is a powerful mechanism, but its computational complexity taxes learners' cognitive resources when encoding and retrieving information. Both cognitive and educational psychologists have demonstrated these challenges. For example, using a scene comparison task, Goldstone and Medin (1994) showed that finding the commonalities in the relations among objects simply takes more time than finding the commonalities in the objects themselves. Moreover, competing commonalities in object features often prevent children from finding relational commonalities (Gentner & Rattermann, 1991; Richland, Morrison, & Holyoak, 2006). This prevention can potentially be explained by the more rapid recognition of the features of objects compared with recognizing the relations among them (Simms, 2013). In educational settings, relative processing speeds can explain many consistent patterns of incorrect answers across students in physics problem solving. For example, when students misuse velocity in replace of acceleration to calculate force, it is often because velocity can be detected on a graph that focuses on a single dimension (Heckler & Scaife, 2015). On the other hand, acceleration (change in velocity over time), is often graphically represented as the slope of a line. Analyzing the slope to extract information about acceleration is slower because this entails processing the relation between two dimensions represented by the graph (velocity and time).

Further, the increased processing speed for recognizing relations reflects increased working memory demands. Waltz, Lau, Grewal, and Holyoak (2000) have shown that finding commonalities in objects across scenes can be done with relatively little working memory resources, while finding commonalities in the relations-between-objects suffers under cognitive load (see Figure 3). Indeed, Halford, Wilson, & Phillips' (1998) Relational Complexity Theory defines working memory capacity specifically by the number of entities that can be actively related to each other in memory, and not just by the number of entities maintained in memory themselves. That is, when entities are related, they must be considered simultaneously. When entities are not being related, they can be considered sequentially.

To better understand how relations specifically tax working memory, consider that you had to explain the effects of two, three, or four variables on a psychological measure like reaction time (RT). Compare the load on working memory for generating an explanation in two different scenarios. In the first scenario, there are only main effects and no interactions. Because each main effect is independent, they can be considered in sequence, one at a time. Thus, there is only a minimal increase in working memory load to generate your explanation from two to three to four variables. In the second scenario, however, all variables interact. When trying to generate an explanation of either a two-way, three-way, or a four-way interaction, the task's difficulty strongly increases because one can only consider the effect of each variable in relation to the others. As researchers, we all have experienced the increased strain of understanding three-way interactions in comparison to two, and then the near impossibility of trying to understand four-way interactions (see Halford, Baker, McCredden, & Bain, 2005, for empirical demonstration).

Educational psychologists have identified the same principle. The relational complexity of a new concept or problem defines the difficulty of for a learner to learn the concept or solve the problem.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

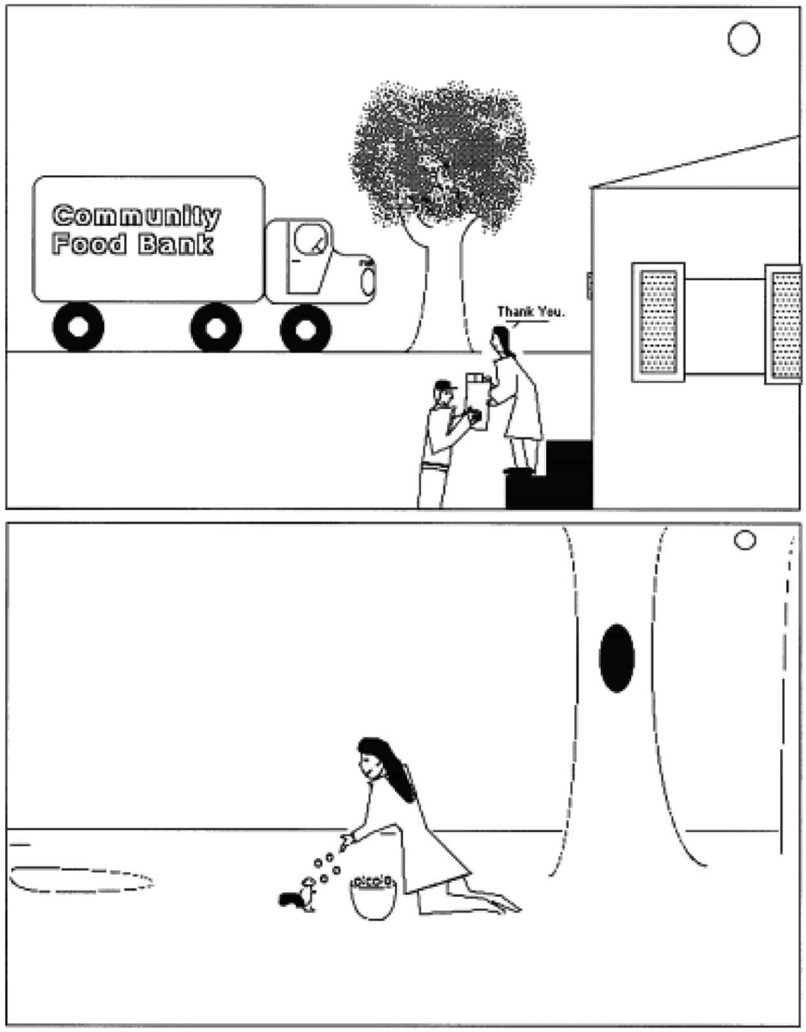

Figure 3. From Markman and Gentner (1993a), stimuli were also used by Waltz et al. (2000). Noticing the perceptual commonalities of the two women in both scenes is relatively automatic. Noticing the relational commonalities between two scenes, that the woman in the top scene is playing the same food-recipient role as the squirrel in the bottom scene, uses working memory resources. From "Structural Alignment during Similarity Comparisons," by A.B. Markman and D. Gentner, 1993, *Cognitive Psychology, 25,* p. 436. Copyright 1993 by Elsevier. Reprinted with permission.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Sweller's Cognitive Load Theory (Sweller, 2010) calls this a problem's or concept's "element interactivity." This refers to the number of elements that interact with each other and thus must be considered simultaneously (and see, e.g., Hegarty, Mayer, & Monk, 1995 on the similar challenge of understanding relational statements in arithmetic word problems). Recent research focuses on whether concepts varying in their element interactivity benefit from different kinds of educational interventions and support procedures (Chen, Kalyuga, & Sweller, 2015; Leahy, Hanham, & Sweller, 2015; and see below).[4]

Challenges in Relational Retrieval

Even if a learner successfully processed and encoded the relational structure of a concept, a learning problem, or an expository text, relational retrieval remains difficult. Quite frequently when presented with a new example (with identical relational structure) after delay, the previous example will not be cued in memory (Gick & Holyoak, 1980), and instead retrieval will often be based on or at least biased by superficial similarity (see Ross, 1984, 1987, 1989; Ross & Kennedy, 1990, for full exploration of similarity-based retrieval in problem solving). To understand why, first recall the complexity of the structure-mapping process. Then consider that the comparison processes that operate over feature-based representations are much simpler. For example, Tversky's (1977) Contrast Model

[4] Cognitive Load Theory argues that the central problem of education is the difficulty of the student to learn complex concepts given the limits on human working memory. Sweller (2011) summarizes a variety of instructional techniques that many years of research has shown to be effective in reducing the student's cognitive load.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

compared feature-based representations and calculated their similarity by counting the number of common and distinct features, together with a relative weighting of these counts based on contextual factors. Unlike the complicated and quite lengthy algorithmic implementation of structural alignment (e.g., see Falkenhainer, Forbus, & Gentner, 1989; Goldstone & Day, 2013; Halford, Andrews, Wilson, & Phillips, 2012; Hummel & Holyoak, 1997), iterations of the Contrast Model are simple enough to be represented as a single formula.

Critically, the differences in computational cost between considering overlap in either features or relational structure predict and explain dissociations between analogical comparison and analogical retrieval (in addition to several behavioral differences between relational and feature-based processing discussed throughout this article). Finding commonalities in relational structure is computationally expensive, and so it is greatly facilitated by simultaneous presentation of the base and the target analog, making memory retrieval unnecessary.

However, retrieving information from long-term memory has many fundamental differences from the consideration and manipulation of information in working memory. Considering and manipulating information in working memory is often a slow and serial process. In contrast, retrieving information from long-term memory often seems to happen in an instant: you walk into a new space, or are reading a new text and related memories can just jump to mind. Unlike in reasoning situations when both analogs are available for simultaneous consideration, overlap in superficial features greatly influences retrieval when only one analog is available and a second must be recollected from memory (e.g., Gentner et al., 1993; Holyoak & Koh, 1987). In contrast to searching for structural commonalities, searching memory for mere overlap in content (similar to the comparisons of Tversky, 1977) is quick and easy (Forbus, Gentner, & Law, 1995; Hummel & Holyoak, 1997).

Gentner and colleagues (1993) demonstrated this dissociation between memory retrieval and the recognition of structural similarity. Subjects read a series of short narratives, including: (a) a target story, (b) a superficially dissimilar, but structurally similar story, and (c) a superficially similar, but structurally dissimilar story. After reading all of the stories, subjects looked past the superficial differences and rated the structurally similar story as a more apt basis for drawing inferences and generalizing information to the target. However, when subjects read only the target story, and had to indicate which story it reminded them of, the subjects reported the superficially similar, but structurally dissimilar story. That is, despite the inferential power of the structurally similar story, it was not readily cued in memory.

The distinction made between the content overlap-based retrieval and structural alignment-driven inference highlights a key aspect of our argument. We claim that the distinction between features and relations is a formal one, not just one of content. Forbus and colleagues (1995) simulated empirical findings such as by Gentner and colleagues (1993). In these simulations, memory retrieval and structural alignment act on the same representational content. However, during memory retrieval the structure is not taken into account. That is, a higher-order relation and an intrinsic object feature are both represented in a similar feature-like representation, and all the search does is to look for content similarity. In full structural alignment, however, the relational structure matters, and determines what inferences are made. Consequently, the same content can be represented in different forms, and it is the form, and not the content, that predicts which psychological processes operate. This difficulty in analogical retrieval has proved critical in educational research.

Knowledge Transfer (and Lack Thereof)

Within a few years of education, young students have to develop reading, writing, and arithmetic skills, but also grasp complicated, highly generalizable scientific concepts that took scientists centuries to discover. Given the argumentation in the previous sections, successful learning of scientific concepts does not only mean that the concepts can be reproduced for materials that have been studied before (thus, with an overlap in both superficial features and relations). Rather, a learner should be able to use and apply the concept in novel contexts (often lacking superficial similarity) and thus, can generalize and make productive use of her knowledge (De Corte, 2003). The application and generalization of prior knowledge to a novel example or context is known as “knowledge transfer.”

While knowledge transfer is a challenge for all students, demonstrating knowledge transfer to novel examples is a critical distinction between high and low performing students. For example, in a study on the use of self-explanations in educational materials, Chi, Bassok, Lewis, Reimann, and Glaser (1989) considered “good” students those who related examples to concepts (like natural laws) resulting in example-independent knowledge. In contrast, “poor” students relied heavily on the examples themselves. Likewise, McDaniel, Cahill, Robbins, and Wiener (2014) have shown that there are consistent individual differences across learning tasks: some learners abstract relational concepts (indicated by their ability to transfer those rules to new cases) while others simply learn the exemplars (and thus, fail to transfer). McDaniel and colleagues related these individual differences to working memory capacity, which is also predictive of differences in category learning speed (independent of learning strategy, see Craig & Lewandowsky, 2012; also see Little & McDaniel, 2015, for differences in the use of rule-learning strategies without differences in cognitive ability).

Studying the question of “how knowledge acquired in one situation applies (or fails to apply) in other situations” (Singley & Anderson, 1989, p. 1), has a long tradition in psychological research (of which analogy is only one part). From the first experimental work of Thorndike and Woodworth (1901) to recent studies on how cognitive skills are transferred (e.g., Taatgen, 2013), the theories developed in this tradition all describe, explain, and predict transfer performance by some kind of similarity between learning and transfer tasks (e.g., Bransford & Schwartz, 1999; Chi & VanLehn, 2012; Day & Goldstone, 2012; Nokes, 2009; Nokes-Malach & Mestre, 2013; Singley & Anderson, 1989). It is beyond the scope of this article to provide the details of these theories, but we recommend some of these other articles for discussion of the roles of, for example, metacognition and motivation, which are crucial for a complete theory of knowledge transfer. Instead, we continue our focus on analogical reasoning theories as they are most suitable for our specific integration of the cognitive and educational psychology research; and they have recently been specifically extended to learning in educational settings like classrooms (see Vendetti, Matlen, Richland, & Bunge, 2015).

We suspect that educationalists took up analogical reasoning theories because these theories are suitable to capture higher-order thinking that modern education requires (see Richland & Simms, 2015, for a recent overview of learning by analogy in mathematics, science, and history education). Furthermore, educational researchers almost regard it as a truism that prior knowledge rather than intelligence or other interindividually varying constructs is the best predictor for future learning (e.g., Schneider, Körkel, & Weinert, 1989; Walker, 1987; Weinert & Schneider, 1999 and see Hambrick, 2003, but also see Hambrick & Meinz, 2011, and Hambrick et al., 2012, for a more complex pattern). Analogical reasoning theories offer a way to describe how prior knowledge is retrieved and used by learners to solve novel complex tasks. However, as we have emphasized, this is a fraught task. Relational knowledge can be difficult to encode and may be even harder to retrieve. Thus, transfer failure may be the most consistent finding in education research. How can learners successfully transfer their knowledge, and how can instruction help?

Effective Instructional Strategies to Foster Transfer Part 1: Comparing Examples to Find Relational Commonalities

One of the most successful and straightforward ways to improve relational category learning and transfer is to guide learners to compare two or more learning examples. The benefits of comparison for relational learning has been demonstrated in children, adults, and in both the education and cognitive literatures (e.g., Christie & Gentner, 2010; Gentner, Anggoro, & Klibanoff, 2011; Gentner, Loewenstein, & Thompson, 2003; Gick & Holyoak, 1983; Goldwater & Markman, 2011; Jung & Hummel, 2011; Kurtz, Boukrina, & Gentner, 2013; Rittle-Johnson & Star, 2011; Tomlinson & Love, 2010).

When first encountering exemplars, relations are often represented in fragmented, and exemplar-specific manners (Doumas, Hummel, & Sandhofer, 2008; Gentner, 2010; Rein & Markman, 2010). However, the comparison process induces a structural alignment of the exemplars that highlights their common relational structure, helping to represent the relations in a more general manner. For example, Kurtz and colleagues (2013) showed that simultaneous presentation of pairs of exemplars increased learning of the category-defining spatial-relations which then enabled transfer to exemplars that shared virtually no featural overlap (see Figure 4). Put differently, when two exemplars are analogs, they can also be considered to be exemplars of the same relational category. If exemplars are aligned successfully, their common relational structure can then be used to classify later and more disparate exemplars. Indeed, Quilici and Mayer (1996) show that comparisons of statistics problems that emphasize shared deep structure help students classify problems by their proper solutions (e.g., which problems require a t test, and that require a χ^2). Further, Reed (1989) argues that problem comparisons can actually fail to help learning when there is no higher-order relational concept for the comparison to aid in discovering.

The benefits of comparison are clear and compelling. Alfieri, Nokes-Malach, and Schunn (2013) conducted a meta-analysis showing that example comparison is a highly effective tool for learning with evidence both from the cognitive and education literatures. Rittle-Johnson and Star (2011) have reviewed the many kinds of comparisons that help mathematics learning; for example, comparing different solution procedures to the same exact problem, and comparing different problems with similar relational structures to each other. The instructional method of "Invention with Contrasting Cases" tasks students to generate the underlying principle that governs a set of examples through comparison

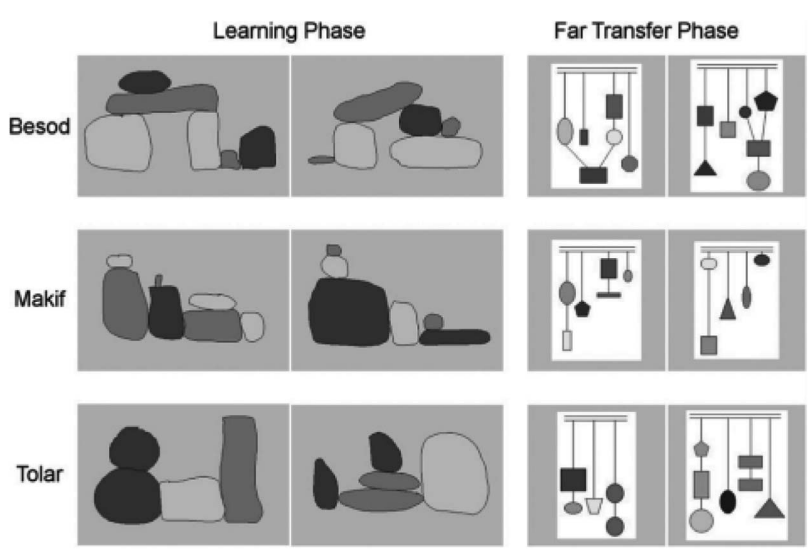

Figure 4. Exemplars from the three relational categories from Kurtz and colleagues (2013) defined by the spatial relations of the rock piles (during learning), and analogous spatial relations among superficially dissimilar elements among transfer visual displays. From "Comparison Promotes Learning and Transfer of Relational Categories," by K. J. Kurtz, O. Boukrina, and D. Gentner, 2013, *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* p. 1304. Copyright 2013 by the American Psychological Association. Reprinted with permission.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

before presenting the accurate principle to the students (e.g., Schwartz & Bransford, 1998; Schwartz, Chase, Oppezzo, & Chin, 2011). This comparison-driven method has increased knowledge transfer over typical "tell and practice"-based instruction wherein students are given a principle before they apply it to a series of problems sequentially with minimal across-problem comparison (see also the related approach of designing for productive failure by Kapur, 2008). Belenky and Nokes-Malach (2012) showed how Invention with Contrasting Cases can induce a mastery-orientation in students, a motivational state predictive of academic success (e.g., Pintrich & Schunk, 2002).

Other research focuses on how to improve the effectiveness of aligning exemplars to foster transfer of relational category knowledge. For example, alignment and relational discovery is aided by first comparing exemplars that also share superficial features to facilitate mapping (e.g., Braithwaite & Goldstone, 2014 for work on learning mathematical combinatorics; Loewenstein & Gentner, 2001 on children's spatial reasoning), and by relational language highlighting the structural commonalities (e.g., Catrambone & Holyoak, 1989; see Gentner, 2010, for a full discussion). Further, Doumas and Hummel (2013; on learning artificial categories of bacteria), and Goldwater and Gentner (2015; on learning abstract causal systems) show how greater learning outcomes can be achieved with separate tasks to first ensure proper exemplar representations before alignment *and then to* directly scaffold the alignment of pairs of exemplars. The two kinds of tasks work together because the utility of the alignment is increased by high quality exemplar representations, and the opportunity to align high quality exemplar representations makes the category knowledge more transferable to disparate contexts.

Critically, representing relational structures abstractly (or at a category-general level), helps to overcome some of the reliance on superficial features in analogical retrieval (e.g., Gentner, Loewenstein, Thompson, & Forbus, 2009; Goldwater, Sibley, Gentner, LaDue, & Libarkin, under revision; Kurtz & Loewenstein, 2007; Novick, 1988; Loewenstein, 2010). Comparison helps learners construct abstract relational schemata and their systematic and rich relational content devoid of superficial features. Such schemata improve retrieval and direct access to prior knowledge based on shared relational structure alone.

Effective Instructional Strategies to Foster Transfer Part 2: Comparing Examples to Find Key Differences

In addition to abstracting relational commonalities, another way structural alignment improves learning is through the highlighting of key differences (see Figure 3) related to the structural commonalities called "alignable differences" (Markman & Gentner, 1993b). For example, because a coconut and a hotel are very dissimilar, it is hard to actually generate many differences between them because very few conceptual dimensions are relevant to each. In contrast, it is quite easy to list differences between a hotel and a motel because the shared relevant dimensions highlight many alignable differences. There are now several demonstrations that indicate that aligning highly similar representations increases learning about their differences using both artificial stimuli (e.g., Gentner, Loewenstein, & Hung, 2007; Sagi, Gentner, & Lovett, 2012), and real world stimuli such as bone structures (Kurtz & Gentner, 2013), model buildings (Gentner et al., 2015), or radiographs (Kok, de Bruin, Robben, & van Merriënboer, 2013). Ming (2009) showed that comparing two distinct methods of graphing data aids learning both methods when the examples used are highly similar to each other. Likewise, Jee, Uttal, Gentner, Manduca, Shipley, and Sageman (2013) showed that comparisons of highly similar rock formations increase geoscience students' ability to identify faults.

However, the overall pattern of the benefits of finding commonalities and differences is more complicated. While both have clear benefits, there has been growing research directly contrasting their benefits, and trying to understand if comparing and contrasting are differentially beneficial depending on contextual factors (see Carvalho & Goldstone, 2014 for review). Specifically, this research has focused on within-category comparisons (to find commonalities), and between-category comparisons (to find differences).

Kornell and Bjork (2008) showed that spacing out presentation of exemplars of multiple "abstract" perceptual categories (e.g., of painting styles) increased learning and generalization of each category in comparison to blocking the presentation of each category's exemplars in a series. While the initial interpretation emphasized that the temporal delay was critical in generalization (making connections to the benefits of spaced practice in skill acquisition (e.g., Donovan, & Radosevich, 1999) Rohrer and Pashler (2010) have provided evidence that the benefit came (at least partly) from interleaving exemplars from different categories. Interleaving enables between-category comparisons, which help discriminate between categories. Rohrer (2012) discusses how interleaving among highly similar categories in math learning (e.g., finding the volume of different kinds of solids) is beneficial because often mistakes in problem solving stem from confusing superficially similar, but conceptually distinct categories for each other. That is, students miscategorize a problem, and consequently, apply the wrong solution procedure. Interleaving helps learning via improving category discrimination.

The benefits of interleaving have led some to, essentially, recommend interleaving universally (e.g., Bjork, Dunlosky, & Kornell, 2013). However, the empirical pattern is again more complex. Higgins and Ross (2011) and Carvalho and Goldstone (2014, 2015) show that between-category comparisons (like those supported by interleaving) are particularly beneficial for categories that are highly similar to each other, and have a high similarity among their own exemplars, but that within-category comparisons are crucial to learning categories that have low similarity among their exemplars (cf. Rohrer, Dedrick, & Burgess, 2014, and see below). One primary way relational categories differ from feature-based categories is that the exemplars of relational categories share less similarity (Gentner & Kurtz, 2005). This lack of similarity (as discussed above) is one reason for why transfer in education is so challenging. Thus, it seems that interleaving may not be optimal for all concepts important in education. In addition, we just wrote extensively about the benefits of analogically comparing exemplars earlier, and most of that work focused on exemplars of the same relational categories.

Further complicating this empirical pattern, Rawson, Thomas, and Jacoby (2014) have shown that the inclusion of a relational concepts' definition modulates the benefits of interleaving. That is, in a task to classify illustrative examples of relational concepts (e.g., classifying descriptions of phenomena illustrating Pavlovian conditioning by the principle label "Pavlovian conditioning"),

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

denying learners access to concept definitions during classification created a benefit for interleaving, but this effect reversed when definitions were available. Carvalho and Goldstone (2014) showed analogous results with perceptual categories. They demonstrated that blocking exemplars is beneficial when exemplars are presented along with category labels, while interleaving was beneficial when learners had to actively classify each exemplar (and received feedback). Carvalho and Goldstone argued that the more passive task of encoding exemplars with their labels encourages within-category comparisons, and so the blocking structure is beneficial as it allows for within-category comparisons across successive trials. Likewise, providing a relational concept definition may encourage a learning-mode of searching for within-category structure wherein learners align exemplar structure with the abstract definition.

In summary, comparisons to discover and abstract within-category structure, and contrasts to highlight key differences between categories both clearly benefit learning. It is a critical avenue of future research to continue to understand the different conditions that optimize the benefits of the different kinds of comparisons. Having a unified model to explain these various effects will constitute a theoretical advance in cognitive psychology, and will allow for clear recommendations for educators when instruction should encourage and afford between-category or within-category comparisons.

Effective Instructional Strategies to Foster Transfer Part 3: Practicing Retrieval and "the Testing Effect"

Next, we consider the benefits of retrieval practice, aka "the testing effect." After initial study of learning materials, retrieving the materials (e.g., on an exam) has several learning benefits over additional study time, most notably the long-term retention of the learning materials (see Roediger, Putnam, & Smith, 2011, for review and discussion of additional benefits). However, as we have made clear, this article is most concerned with improving the generalization (i.e., the transfer) from learning materials to novel materials based on commonalities in relational structure. Here, the evidence is not as clear.

Jacoby, Wahlheim, and Coane (2010) have extended the retrieval practice paradigm to category learning, and showed that it improves both the learning of individual category exemplars and the category-level generalization(s). However, they used natural feature-based categories (e.g., birds), and so it does not entail successful relational generalization (and see Rohrer, Taylor, & Sholar, 2010 for benefits of retrieval for another form of transfer). Butler (2010) found that practicing retrieval increases analogical transfer, but he only tested with transfer problems that contained explicit retrieval cues. As Butler recognizes, transfer is quite hard without such cues (but again, for education, spontaneous transfer is the true measure of success). Goldwater, Smith, and Hinze (in preparation) could not find the benefits of retrieval practice of learning materials over repeated study in supporting spontaneous analogical transfer of a relational schema. Furthermore, Tran, Rohrer, and Pashler (2015) showed that retrieval practice does not aid deductive inference.

On the other hand, and critical for our argument, McDaniel, Thomas, Agarwal, McDermott, and Roediger (2013) have examined the benefits of retrieval practice on the spontaneous transfer of studied relational concepts. In their experiment they examined how quizzing middle school science students in the middle of a semester on a previously taught relational concept (e.g., competition for resources in ecology) affected classifying new examples of the concept on the final exam at the end of the course. The midsemester quiz tested students on either the spontaneous classification of a novel example of the concept, which required retrieving their past knowledge of the concept definition and applying it to this new item, or just testing their knowledge of the concept's definition directly. If the quiz item merely required the students to retrieve the definition of the concept, then there was no benefit in fostering transfer to novel items on the final exam (in contrast to the item not being quizzed at all), but practicing relational retrieval, transfer, and categorization fostered further transfer on the final exam. That is, transfer begot transfer.

In contrast to the work of McDaniel and others, work in Cognitive Load Theory by Sweller and colleagues have challenged the benefits of testing (Chen et al., 2015; Leahy et al., 2015). These researchers have argued that testing is most effective with low element-interactive items (i.e., items with low relational complexity). Cognitive Load Theory has championed the use of "worked examples" that comprehensively describe how to solve example problems. This method reduces the learner's working memory load by structuring problems into a series of subgoals and steps to be completed in sequence (e.g., Sweller & Cooper, 1985; Sweller, van Merrienboer, & Paas, 1998; and see Catrambone, 1995, 1998; Duncan, 2010). Sweller and colleagues argue this method taxes working memory less than typical "tell and practice instruction" that first describes how to solve a problem and then has students practice problem solving. They further argue that problem-solving can be considered a form of retrieval-practice and testing. They demonstrate that problem-solving may be beneficial over studying worked-examples for low complexity items, but that the benefits reverse for high complexity items. With high relational complexity items, they argue the only way to ensure successful learning is a strategy to minimize working memory load. Further, Sweller argues against the whole notion of a "desirable difficulty" (i.e., an increase in taxing working memory with a learning procedure that leads to improved long-term learning outcomes), promoted by Bjork and colleagues (e.g., Bjork et al., 2013), by suggesting that the benefits of instructional designs based on desirable difficulty are limited to low complexity items. With high complexity items, and given limited working memory capacity, Sweller argues the best chance learners have is to minimize cognitive load, not to add to it.

However, it is important to note that proponents of testing and the desirable difficulties framework have countered these arguments, point to successful outcomes of testing with high complexity items, and suggest that there is no clear a priori way to measure element-interactivity, questioning its practical use (Karpicke & Aue, 2015).

In summary, the testing effect can apply to spontaneous relational transfer, but only when the test itself requires spontaneous relational transfer (McDaniel et al., 2013). However, there is a current controversy about whether these effects are limited to items that do not apply a high cognitive load at the point of encoding. Clearly, more research is needed.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Effective Instructional Strategies to Foster Transfer Part 4: Classification Training and Perceptual Learning Training

The form of successful testing shown by McDaniel and colleagues (2013) directly trained relational categorization of example phenomena. Other research has found similar success in similar direct trainings. For example, Mestre, Docktor, Strand, and Ross (2011) have shown that an effective way to teach physics is to train students to explicitly categorize problems by their governing abstract scientific concept (see Fox & Sullivan, 2007, and Gurlitt, Dummel, Schuster, & Nuckles, 2012 for related approaches). Rohrer and colleagues (2014) have provided evidence that interleaving benefits learning by emphasizing category-level thinking (in addition to improving category discrimination). Unlike most previous demonstrations of the benefits of interleaving, which relied on the discrimination between similar problem-types, this research demonstrated interleaving also benefits problem categories less similar to each other. Rohrer and colleagues have shown that interleaving problem types is beneficial, as learners cannot apply a solution without first classifying the problem. In contrast, when learners practice a series of problems of the same category in a row, they can apply the solution without first classifying the problem. Thus, interleaving is beneficial because it reinforces exemplar to category connections. Likewise, Kurtz (2013) argues that framing relational principles as categories that have exemplars strengthens the status of the principle in the student's mind. Taken together, these lines of research strengthen our argument that relational categorization processes are a critical part of successful educational outcomes.

Cognitive psychologists have used inductive classification methods to examine basic mechanisms of relational category learning (using methods developed by the feature-based category literature in adults). For example, several researchers have analyzed the relative difficulty of learning relational category structures, such as contrasting probabilistic family resemblance structures with fully deterministic rule-based structures (e.g., Kittur, Hummel, & Holyoak, 2004; Jung & Hummel, 2011).

Corral and Jones (2014) have examined the learning of six logically distinct kinds of relational category structures. By examining the relative difficulty of the six category structures, Corral and Jones were able to determine how well several kinds of relational category learning models could characterize the overall pattern (i.e., they presented novel empirical data and novel models to fit the data). This explicit comparison of theoretical models showed that throughout learning there were two independent key processes continuously at play. First, schema-refinement is a process that removes information from a category defining relational structure. Second, schema-elaboration is a process that adds information to that structure. We believe characterizing these two learning processes and their interaction will be critical in advancing our understanding of learning in educational contexts (see below).

Using similar methods to the supervised inductive classification, Kellman and colleagues (e.g., Kellman & Massey, 2013) have emphasized the importance of perceptual learning in STEM (and see Goldstone, Landy, & Son, 2010), as visuospatial representations (e.g., graphs, symbolic characters, etc.) are ubiquitous. Consider again the findings of Heckler and Scaife (2015) wherein students misused velocity for acceleration because of the relative ease of extracting velocity from a graph. Kellman and colleagues have taken the approach of directly training the extraction of perceptual relations in a variety of STEM domains with great success. When learners improve their perceptual extraction abilities, this enables the use of "higher-level" problem solving skills to operate over these high quality perceptual representations. Without such perceptual training, novice learners are often mired in the attempt to properly interpret and visually parse graphical and formal representations.[5] Similar to the benefits of interleaving, this perceptual training shows that recognizing relational structure in a new problem is a critical first step in solving it.

However, to train relational recognition, there is a tension between the use of concrete, grounded, and perhaps familiar instantiations of the relational structure, or the direct presentation of idealized and formal representations (such as rules or principles). This tension has inspired educational researchers to examine and directly compare concrete and idealized external representations (see Belenky & Schalk, 2014, for an overview). To this aim, researchers have evaluated whether studying only an idealized representation of a concept suffices to support transfer (e.g., Kaminski, Sloutsky, & Heckler, 2008; De Bock, Deprez, Van Dooren, Roelens, & Verschaffel, 2011). A consistent pattern has emerged. While single concrete instantiations may help students' initial understanding, an idealized representation seems to better support transfer. Perhaps an ideal method to ensure both initial understanding and transfer is "concreteness fading." Accordingly, the perceptual richness of concrete instantiations is gradually reduced over time, leaving idealized and schematic representations (e.g., Fyfe, McNeil, Son, & Goldstone, 2014). For example, Goldstone and Son aimed to help learners understand the concept of "competitive specialization." This concept captures how an undifferentiated and homogenous system can self-organize into a structure of systematically differentiated functions through the interactions of its parts without any kind of "central planner." Goldstone and Son (2005) had learners interact with simulations of ants foraging. At first, the ants looked like ants, but then their appearance slowly changed into little dots. This fading enabled both high levels of domain understanding (of ants foraging) and cross-domain transfer of the concept of competitive specialization.

Given the earlier discussion on the challenges of relational retrieval, why would direct categorization and perceptual training improve analogical transfer? That is, how do these data fit with the idea that transfer is based on the deliberate analogical retrieval and mapping? It is important to note that transfer can be achieved differently. Sometimes transfer is achieved by deliberate analogical retrieval, for example with explicit analogizing in novel scientific theorizing and engineering design (e.g., Dunbar & Blanchette, 2001; Gentner, Brem, Ferguson, Wolff, Markman, & Forbus, 1997; Linsey, Wood, & Markman, 2008; Chan & Schunn, 2015). However, other psychological frameworks do not conceive of transfer in such a way, and instead focus on how prior (relational) knowledge "serves as a *lens* for the construal of new content rather than being the direct *focus* of cognition itself" (Day & Goldstone,

[5] It is worth noting that these researchers argue that even these "higher-level" abstract reasoning skills are grounded in perceptual processes (e.g., Landy & Goldstone, 2007).

2012, p. 165). That is, relational knowledge directly constrains how we perceive, interpret, and understand the world. Transfer is then inherent to how a student would make sense of any new problem or expository text: in terms of their prior knowledge. The remaining sections of this article concern these interpretative processes, the challenges these processes raise for education, and suggested solutions.

The Cohering Nature of Relational Knowledge

Thirty years before this article, Murphy and Medin (1985) criticized the feature-based categorization enterprise as an incomplete approach to human concept learning and representation. They argued that focusing on the learning of different arbitrarily defined artificial categories with unstructured sets of features neglects how people represent categories as coherent in the real world. That is, the subjects' prior background knowledge helps to tie together (or cohere) the relevant features. Following this critique, several years of research affirmed how cohering background knowledge can make categories easier to learn (e.g., Medin, Wattenmaker, & Hampson, 1987; Murphy & Allopenna, 1994), and how it drives the induction and generation of explanations (e.g., Lassaline & Murphy, 1996; Patalano & Ross, 2007).

Convincing evidence for the importance of background knowledge comes from Wisniewski and Medin (1994). They showed that subjects' naïve theories affect the interpretation, perception, and learning of features themselves, not just how categories are formed on the basis of features. Traditional feature-based categorization models could simulate many findings concerning the role of background knowledge (e.g., Murphy & Allopenna, 1994) by simply seeding the models with knowledge of a few category exemplars before simulating the experiments (Heit, 2001). However, because these models form categories from features but do not learn the features themselves, they cannot explain Wisniewski and Medin's results (see how category learning changes perception and feature interpretation, Goldstone, 1998; Schyns, Goldstone, & Thibaut, 1998).

Important for the current article's argument, there is considerable evidence that this cohering knowledge is at its essence relational: whether it is relations among features intrinsic to a category member, or between category members and other objects or knowledge structures. For example, research has demonstrated that features simply are not represented independently of each other, but that humans bind them together (Love & Markman, 2003) often by their predictive relations (e.g., Billman & Knutson, 1996; Mathy, 2010), which can enable forming abstractions (Delosh, 1997), in manners beyond what typical feature-based models can explain (e.g., Wattenmaker, Dewey, Murphy, & Medin, 1986).

A particular strand of this line of research has focused specifically on how causal relations cohere featural representations (see Rehder, 2010, for review; also see Rehder & Hastie, 2001; Sloman, Love, & Ahn, 1998). For example, the representation of the category *guitar* or *dog* are not just lists of features, but the features have particular causal relations among them: a guitar can make music because it has strings and a hollow wooden body, and dogs have particular features because of their DNA. Further, people reason that guitars exist because of peoples' intention to make music (Bloom, 1996), or that DNA mutations underlying particular dog features are the result of evolutionary processes (e.g., Gelman & Legare, 2011). Going beyond categories whose members share intrinsic features, Rehder and Ross (2001) discuss "abstract coherent categories." These categories comprise many superordinate categories; for example, *mammals* or *furniture*, but also relational concepts like *divorce* or *revolutions*. The exemplars of these categories can be classified as such because each exemplar's set of unique features can be cohered with the same relations. Going further, researchers hypothesized that cohering relations among disparate features are responsible for how conceptions are dynamically adapted to or even change in different contexts (Gabora, Rosch, & Aerts, 2008).

In addition to focusing on how relations cohere features within a category member, Jones and Love (2007) analyzed the normed feature listings for the large set of categories from McRae, Cree, Seidenberg, and McNorgan (2005). Jones and Love showed that information about extrinsic relations was prevalent in terms of how category members relate to other categories (see also Barr & Caplan, 1987; Goldwater et al., 2011). For example, part of our knowledge of dogs is that they are typically owned by humans as pets, and can assist humans in hunting other animals. In sum, concepts are coherent because of between-feature relations, and between-category relations (more on the latter below). These cohering relations affect how we classify novel disparate exemplars, generate explanations, and even alter the perception of features themselves.

Developing Expert Relational Knowledge Part 1: Top-Down Influences on Perception

If relational knowledge coheres concepts, drives interpretation of novel exemplars, and directly affects our perception, then successfully learning scientific concepts will change perception (or at least the focus of attention, e.g., Goldstone, Landy, & Son, 2010; Chi & VanLehn, 2012; Kellman & Massey, 2013). In other words, there is a top-down effect of prior knowledge, in form of relational knowledge on perception of features, their co-occurrences, and their interactions. Thus, prior knowledge determines how relations are constructed or recognized in novel (and familiar) contexts.

According to analogical learning theory, the process of aligning examples of scientific concepts will often trigger relational rerepresentation (Gentner, 2010; see Klenk & Forbus, 2013, for computational implementations of this process). Rerepresentation can result in larger, more abstract relational chunks tying together superficially different situations. Goldwater and Gentner (2015) had students structurally align examples of disparate phenomena with a common underlying causal structure. Subsequently, the students were able to sort novel phenomena into causal system categories, ignoring domain differences (similar to the science experts of Rottman et al., 2012). Goldwater and Gentner theorized that the analogical training did not just make the training items easier to retrieve and apply to the novel items, but gave the learners new representational vocabulary with which to interpret the novel exemplars. Schwartz, Bransford, and Sears (2005) argue that rerepresentation does not only support transfer, but transfer items themselves can lead to the rerepresentation of previously learned exemplars.

According to this view, scientific expertise is not just about easily retrieving the right kind of prior knowledge. Rather, expertise is reflected in efficient relational processing that entails pro-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

cessing larger relational chunks, and/or the more rapid and flexible coordination of relational chunks (e.g., see Smith, diSessa, & Roschelle, 1994, and below). These relational chunks allow experts to "just see" the structure in novel exemplars often without requiring an explicit reminder of the right knowledge to be mapped (and it may not be a conscious process at all, see Day & Gentner, 2007 Day & Goldstone, 2011). For example, Rikers, Schmidt, and Boshuizen (2002) discuss "knowledge encapsulation" in medical experts. In a doctor's area of expertise (e.g., cardiology for a cardiologist), diagnoses are fast and accurate because they have rerepresented their biological knowledge in service of efficiently mapping symptoms directly onto disorders bypassing several explanatory steps. That is, expert explanations for diagnoses actually contain *fewer* biological details than subexpert explanations. On the other hand, when doctors are given cases outside of their area of expertise (e.g., cardiology for a neurologist), they have to reason step-by-step through the biological mechanisms to get from symptoms to diagnoses (and this is a more error-prone process).

In education research, Sweller and colleagues' work on the interaction between element-interactivity and the effectiveness of different instructional tools (Chen et al., 2015; Leahy et al., 2015, see above) has revealed a similar relational chunking process across development in primary school students. Accordingly, the proper instructional tool depends on the relational complexity of a problem, but the relational complexity is not just determined by the intrinsic structure of the problem, but by the prior knowledge of the learner. That is, with high prior knowledge, a previously high element-interactivity item becomes processed as a low element-interactivity problem. Thus, the kind of problem that carried a heavy cognitive load and needed a worked-example to be learned properly becomes a problem for which a more generative and active learning tool is more effective. The more advanced learner with larger relational chunks has freed up cognitive capacity to engage in these more active processes (see Kalyuga, Ayres, Chandler, & Sweller, 2003, for a systematic review of such "expertise reversal effects").

The work on expertise development highlights why the ability to just see the structure is so important for fostering transfer: there are severe cognitive bottlenecks in both relational encoding and explicit retrieval (e.g., Gick & Holyoak, 1980; Waltz et al., 2000). Considering the rerepresentational capabilities of the analogical learning models (e.g., Gentner, 2010), together with the perspective of the perceptual learning theories (e.g., Kellman & Massey, 2013; Goldstone et al., 2010) suggests that science expertise is the codevelopment of both highly abstract relational representations, and highly efficient perceptual extraction abilities. An important line of future research is to investigate these two aspects of expertise development together. For example, using relational category learning paradigms, (e.g., in Corral & Jones, 2014) one could examine how learning higher-order relational structures change the perception of the lower-level features of the exemplars, and how experience in interpreting the features help build the higher-level structures.

Relational Concepts Comprise Knowledge Systems

Relational knowledge does not just cohere individual concepts (e.g., Rehder & Ross, 2001), but relational concepts form coherent systems of interconnected concepts. Relational schemas govern multiple role concepts, each defined by their place in such knowledge structures. There can be no reagents without catalysts, zeniths without nadirs, guests without hosts, or parents without children. Goldwater and colleagues (2011) supported this framework of relational category representation by demonstrating that instantiating a new relational structure licensed forming new role-governed categories.

Before the more recent focus on relational categories, for decades there has been research on systems of category knowledge (e.g., Collins & Quillian, 1969). However, the focus on feature-based taxonomies (e.g., of biological kinds) has limited consideration of the kinds of structures that exist among categories, and the kinds of reasoning that a conceptual system affords. Work on taxonomic categories have focused on which feature dimensions define the similarity space within a domain (e.g., Smith et al., 1974), how feature-based representations are distorted to improve discrimination among interrelated categories (i.e., categories are polarized, leaving contrasting categories such as Democrats and Republicans, seeming more dissimilar than they really are; Davis & Love, 2010; Goldstone, 1996; Heit & Nicholson, 2010), or how hierarchical class-inclusion relations affect feature-based inferences from one category to another (e.g., if the category of bird has some feature, then this feature is handed down to all subcategories, see Collins & Quillian, 1969; cf. O'Connor, Cree, & McRae, 2009). That is, when conceptual systems are comprised of solely feature-based categories, the categorical interrelations are generally constrained to what features are shared by which categories, and how that affects classification and induction.

Going beyond representing common and distinct features, coherent systems of relational concepts are powerful tools for learners that create new understanding. Loewenstein and Gentner (2005) demonstrated that interrelated relational language (*top*, *middle*, and *bottom*) aided 4 year-olds' spatial reasoning more so than relational language with less systematic interrelations (*in*, *on*, and *under;* which in turn was more helpful than no spatial language). Gentner, Ozyürek, Gürcanli, and Goldin-Meadow, (2013) show that deaf children to hearing parents who lack systematic spatial language input are years behind normatively developing children in their spatial reasoning ability (assessed by the task developed in Loewenstein & Gentner, 2005).

Moran, Bereby-Meyer, and Bazerman (2008) demonstrated the generative power of conceptual systems. They discuss that in negotiation, there are a number of strategies that all work together toward the common goal of maximizing value creation for both parties such as

> (a) Logrolling, or trading off concessions on low-priority issues for gains on higher priority issues; (b) trading differential time preferences, or allocating more initial outcomes to the more impatient party and greater profits over a longer period to the more patient party; (c) compatibility, or identifying issues for which parties do not have a conflict of interest; (d) adding issues, or supplementing the agreement with issues not inherent in the initial negotiation framework; and (e) contingent contracts, or bets based on different expectations regarding a future event (p. 100).

Subjects in Moran and colleagues' study compared and contrasted example cases that illustrated two different strategies. Furthermore, they were instructed to try to describe the more general

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

principle behind both strategies, that is, considering the interrelations of parties' interests enables the creation of higher values for both sides. Subjects who received this training did not only transfer their knowledge of the trained strategies to novel negotiations, but showed use of value creation strategies that were not trained. In contrast, subjects who were only trained on a single strategy showed transfer on that trained strategy only, but showed no additional innovation. These latter subjects had no opportunity to find the higher-order organizing principle that coheres each specific relational concept into a conceptual system, and thus could not creatively generate novel solutions.

Developing Expert Relational Knowledge Part 2: Increased Coherence

The last section on conceptual systems prompts an update to our account of expertise. Developing expertise is not just about building larger relational chunks, but also about learning and understanding highly interconnected systems of relational knowledge. That is, knowledge systems become more coherent with increased domain expertise (e.g., Bransford, Brown, & Cocking, 1999; Thagard, 1997, 2007).

Lachner and colleagues (Lachner, Gurlitt, & Nückles, 2012; Lachner & Nückles, 2015) compared the knowledge organization of expert cardiologists and advanced cardiology students to develop a formal account of Boshuizen and colleagues' account of knowledge encapsulation (e.g., Rikers et al., 2002). Therefore, they created graph representations of the explanations and diagnoses of a test-case cardiology patient consisting of concept nodes and their interrelations. There were two central findings. The expert knowledge graphs had fewer nodes between their most distinct concepts: that is, their knowledge was comprised of larger relational chunks, subsuming (or perhaps bypassing, see above) the concepts used by students. In addition to the more abstract nature of expert knowledge organization, the graph structures showed that the medical students' explanations contained many more concepts that were isolated from each other in comparison to the expert explanations. Thus, expert knowledge is more coherent.

Lachner and Nückles (2015) then examined the effectiveness of the expert versus the advanced student explanation as instructional tools for novice students. They considered that while the expert coherence should be beneficial, perhaps their more abstract conceptions would offer a challenge. When the novice students attempted to learn from these explanations, the increased coherence of expert explanations triggered higher-quality learning strategies. For example, the novices learning from expert explanations were more likely to use self-explanations and make deeper inferences, while novices learning from advanced student explanations were more likely to merely paraphrase these explanations. These higher quality learning activities more than compensated for any potential challenge inherent to understanding the more abstract concepts utilized in the expert explanations. That is, the novices who learned from the expert explanations outperformed the novices who learned from the advanced student explanations on several assessments of comprehension and transfer.

In another example of research into the coherent nature of expert knowledge, Koponen and Pehkonen (2010) asked physics teachers and students to represent their own conceptual knowledge as network structures, and to propose experiments operationalizing the target concepts. Similar to Lachner and Nückles's (2015) findings, the teachers' network structure was more highly interconnected. Further, the specific experiments proposed reflected the individual teacher's knowledge organization. Koponen and Pehkonen took this as evidence that there is interplay and mutual feedback between expert knowledge organization and how that knowledge is used. Scientific conceptual structures are formed via the results of experimentation, and also determine how such experiments are designed.

In summary, developing a coherent system of relational concepts gives reasoners powerful tools. Nevertheless, more research is needed to investigate how sets of concepts succeed or fail to cohere during learning, and how these new concepts support reasoning. Hopefully, more studies across domains will replicate Moran and colleagues' effects from the domain of negotiation on the generative nature of conceptual systems (Moran et al., 2008) and give key insights into how to teach powerful conceptual systems. However, often the biggest challenge to teaching a conceptual system is that students do not come to the classroom without any conceptual system. Rather, they come to the learning task with a preexisting conceptual system (even if relatively incoherent) that conflicts with the material to be learned.

Conceptual Change and Problems of Conceptual Change

Education is a process of building a coherent and abstract knowledge system that correctly represents the interrelation of a large number of concepts. In the education literature, the development of such complex knowledge systems is generally referred to as conceptual change (also belief revision or theory change; see, e.g., Carey, 2000, 2009; Ohlsson, 2009, 2011, 2013; Smith et al., 1994; Vosniadou, 2010).

More specifically, the conceptual change literature deals with the problem of how learners develop and change their conceptual knowledge networks over time. Key questions of this literature focus on how prior knowledge (referred to as misconceptions, preconceptions, naïve concepts/beliefs/theories, everyday conceptions, intuitive theories, and alternative theories/beliefs) influences acquisition of scientific concepts (e.g., Schneider, Grabner, & Paetsch, 2009), whether people develop integrated (e.g., Vosniadou & Brewer, 1992) or fragmented knowledge bases (e.g., Tytler, 1998; diSessa, Gillespie, & Esterly, 2004), and whether—and to what extent—learning scientific concepts actually suppresses competing explanations based on prior knowledge (Shtulman & Valcarcel, 2012). Even if one is convinced by Heckler's (2011) argument that lower-level learning mechanisms are more theoretically useful than analyzing student prior knowledge, one still needs to recognize that a network of conceptual knowledge develops with education, and those mechanisms demand explanation.

The Challenge of Changing One's Conceptual Knowledge

While theoretical interpretations of conceptual change vary, there is a consistent pattern of how some prior/misconceptions are persistent and stubborn in the face of instruction.[6] Along with the challenge of relational complexity and knowledge transfer, overcoming prior conceptions is the biggest challenge students face (and these three challenges are interrelated). For example, while students have no formal knowledge of physics before entering the classroom, they have experienced physical forces such as gravity, and have folk conceptions of how such forces cause objects to move and bring these preconceptions to the classroom (e.g., Clement, 1982). Muller, Sharma, and Reimann (2008) demonstrated that when students are learning Newtonian mechanics and preconceptions or naïve theories of forces are not explicitly addressed, they often do not learn from explicit instruction. Instead of learning the target principles (i.e., Newton's Laws), they often interpret new information in terms of their preconceptions. Thus, when asked later to recall the instructional material, they only recall the information they started with as if their preconceptions were the content of the instruction.

It is worth illustrating at least one contemporary line of conceptual change research in STEM education in some detail. The study of complex systems has been an area of focus because principles of complexity underlie phenomena across the sciences, and are particularly hard for students to understand. Chi and colleagues (2012) provided a detailed analysis of the conceptual transitions students must undergo (building on the work of, e.g., Jacobson, 2001; Wilensky & Resnick, 1999). They explained that people's naïve theories of causal processes assume a sequential structure of subevents. Each subevent is contingent upon other subevents in direct and predictable ways. Small groups or individual agents cause effects in linear manners through their intentions to achieve goals, and these individual intentional actions of agents correspond to the overall macropattern of the process. For example, the way wolves hunt is similar to this. An alpha wolf leads the hunt, while supporting wolves' actions are contingent on the alpha wolf, by surrounding the prey, and so forth Properties of the macroprocess (such as the path taken and time lapsed) can be directly related to the actions of the individual agents.

Chi and colleagues (2012) contrast sequential processes with emergent processes. Emergent processes are central to complex systems. They describe how the random interaction of many agents governed by a few common rules lead to macrolevel patterns in a nonlinear manner with no direct correspondence to any subset of agents, and without a clear discrete sequence of subevents. How ants forage for food is an example of an emergent process and is contrasted with the sequential process of a wolf hunt in that unlike the wolves, the ants seem to move about randomly when searching for food. They simple emit pheromones when they find food, and follow the pheromones of other ants. At the microlevel, there is primarily randomness, but at the macrolevel there is a clear organized structure to the ants' foraging.

Chi and colleagues (2012), and Jacobson and colleagues (2011) argue that understanding emergent phenomena requires an *ontological shift* on the part of the learner. This shift requires students to recognize that emergent phenomena are governed by an entirely different kind of causality in comparison to how sequential processes are interpreted in naïve causal theories. Many processes that students learn in science class can be broken down into sequences of causally contingent subevents (even if there are no intentional agents like in the processes of mitosis or photosynthesis), while many other processes are emergent (such as molecular diffusion or natural selection). Chi and colleagues show that phenomena governed by emergent processes are more challenging to learn than processes governed by sequential processes because students try to interpret them both with their sequential causal schemas. In addition, understanding emergent processes taxes cognitive capacities more than understanding sequential processes, similar to the example of how challenging it is to interpret statistical interactions. While the causal explanations of novice students are typically wrong in both cases, it is easier for them to acquire the right explanations when the structure of the process fits the ontological structure of the students' naïve theory.

Further distinguishing simple learning of new information from conceptual change, Johnson and Carey (1998) analyzed the biological knowledge of adults with Williams Syndrome. These adults knew many facts about biology. However, their severe mental retardation prevented them from changing their childlike domain theories of biology that governed their organization of knowledge of individual facts; for example, that the concept of life is equivalent to the concept of animism. The lesson from this work is that memorizing new facts is easy, but drastically changing your conceptual networks is hard.

The challenge of conceptual transformation is further reflected in how gradual it is. Conceptual systems cannot simply be replaced with new ones. Repeatedly, the work of Vosniadou (1994, also see Vosniadou & Brewer, 1992; Vosniadou, Vamvakoussi, & Skopeliti, 2008) has shown that in between naïve mental models and expert mental models there exist synthetic models that have elements of both. Evidence suggests that this synthetic model stage cannot be skipped.[7]

To meet these challenges, recent work suggests that one of the most effective general methods for learning novel conceptual systems is through the rigorous analysis of the knowledge components that define a domain. These cognitive task analyses help to understand how students' prior conceptions relate to the domain knowledge they should acquire (Koedinger & Roll, 2012). A meta-analysis by Clark, Feldon, van Merriënboer, Yates, and Early (2007) showed that curriculum redesigns based on cognitive task analyses have a strong average effect (Cohen's $d = 1.7$) for improving learning outcomes over the curriculum they are designed to replace. Given these promising results, cognitive task analyses should become a more frequently used source to allow for rigorous experimentation into the mechanisms behind learning conceptual systems.

The Challenge of Researching Conceptual Change

Despite the advances in understanding learners' conceptual development via methods such as cognitive task analysis, conceptual

[6] Of course, as we will discuss designing instruction for conceptual change by reflecting on misconception and so forth does improve effectiveness over instruction that ignores student's prior conceptions, for example, Ebenezer and Gaskell (1995); Ebenezer, Chacko, Kaya, Koya, and Ebenezer (2010), among many others, see below.

[7] Taylor and Ahn (2012) showed a similar pattern using much simpler causal learning paradigms. Their study showed that when learners first encounter data that suggest a causal explanation, and then later encounter data that re-explained the original data with a new causal factor, learners' causal representations contained both the initial false causal interpretation, and the later presented true causal relations.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

change research faces several challenges. For example, most of the scientific evidence in the conceptual change literature comes from analyses of interview data. That is, because the scientific concepts studied are complex, it is very complicated to develop materials (e.g., a sufficiently large number of stimuli) that provide reliable and valid quantitative measures of the complex knowledge structures associated with such highly interrelated relational concepts. However, as Schneider and Hardy (2013) pointed out, no standards have been established on how interview data should be analyzed. They argue that this lack of standards might have caused controversies in the conceptual change research community; for example, with regard to the question of whether and to what extent knowledge structures are integrated or fragmented.

To tackle this question, Schneider and Hardy (2013) provided one of the first explicit quantitative statistical models to describe developmental pathways of conceptual knowledge in a scientific domain (for a similar approach see Edelsbrunner, Schalk, Schumacher, & Stern, 2015). In the domain of floating and sinking, they demonstrated that there is a limited number of pathways of knowledge development. The development depends on systematic knowledge-construction processes, indicating strong interindividual differences with some children showing increases and other children showing decreases in knowledge fragmentation (and some showing a stable pattern). Despite these recent theoretical and methodological advances in the conceptual change research, we propound that using the methodology of category learning research developed by cognitive psychologists might be beneficial for addressing these major challenges. For example, in a study conducted by Rütsche and Schalk (in preparation), participants learned different artificial relational categories to induce prior knowledge to evaluate how well this knowledge can be combined with novel relational knowledge in the next learning session. This approach allows for tight control and systematic variation of the prior knowledge that learners bring to the second session. Furthermore, presenting a series of clearly delineated classes of transfer items allows assessing the degree of how learners rely on featural and relational similarity; that is, how they construct categories.

Ohlsson (2013) put forward another challenge for conceptual change research. He claims that the conceptual change literature has been dominated by a normative view. More specifically, researchers often studied why learners do not change their conceptual knowledge even though they are provided with falsified evidence, and with a scientific concept that might be—and the educator typically promotes it as such—more appropriate and powerful in explaining phenomena. Ohlsson argues that most people do not revise their conceptual knowledge or beliefs based on evidence, but that they follow a more pragmatic strategy of evaluating the utility of the knowledge to explain phenomena (a comparable argument was put forward by Smith et al., 1994, and Clement, 1993). Further, (at least) children's performance is driven not just by a static set of concepts (naïve, scientific, or otherwise), but by the interdependence between the task environment and mental constructs (Kloos, Fisher, & Van Orden, 2010).

Again, we assume that controlled category learning experiments might be fruitful in shedding some light on how people actually evaluate the cognitive utility of scientific concepts and naïve conceptions and how they use or refuse to use them accordingly. For example, Markman and Ross (2003) review how category use affects category learning, and that different uses such as inferential predictions, communication, and problem solving each lead to different learning patterns than typical inductive classification. Likewise, when forming analogies people focus on the particular relational correspondences in line with their current goals (Spellman & Holyoak, 1996). Having subjects learn potentially "contradicting" relational categories in different manners or with different goals in mind (thus, each having separate utilities), would potentially help to better understand how even science experts seem to retain naïve conceptions along with their expert knowledge (e.g., Goldberg & Thompson-Schill, 2009).

Taken together, the conceptual change literature is dealing with the acquisition and development of highly complex interrelated conceptual knowledge systems. The complexity leads to specific challenges. The challenges have been clearly identified and led to intensive discussions among educational psychologists, but have proved to be difficult to answer when continuing to work solely with the methodology mainly used in the conceptual change literature. We propose that combining the conceptual change research with relational category learning research makes it possible to design controlled experiments that tackle exactly these challenges.

Indeed, some of the techniques we have been arguing for in this article have already been shown to successfully effect conceptual change. Demonstrating the utility of highlighting alignable differences (i.e., contrasts), Gadgil, Nokes-Malach, and Chi (2012) showed that contrasting naïve and expert conceptions about how the human heart and the circulatory system work fosters large learning gains (also see Loibl & Rummel, 2014). Similarly, Chi and colleagues (2012) indicated how directly contrasting emergent phenomena with sequential processes aids learning the complexities of emergence. In a similar vein, Jacobson and colleagues (2011) demonstrated how comparing two examples of complex system simulations fostered overcoming naïve causal theories and to learn the target causal concepts. Goldstone and Son (2005) developed concreteness fading methods that enable learners to encode high quality perceptual representations of complex system simulations fostering mastery and transfer of target concepts.

Other successful methods for transfer discussed above, for example, testing with relational categorization (McDaniel et al., 2013) and practicing retrieval have not yet been investigated with regard to whether they benefit conceptual change specifically. It is possible that there have been such benefits, but it is unknown because analyzing prior conceptions is not typically a part of these paradigms. In light of how novel instructional materials are often remembered in terms of prior misconceptions (Muller et al., 2008), could retrieval practice (in some situations) have the negative consequence of actually reinforcing naïve theories? We hope not, and believe that careful research should be able to incorporate the strengths of methods such as retrieval practice with successful conceptual change techniques.

Building on these successes of applying analogical and perceptual learning to conceptual change in STEM (and several we lack the room to discuss), we hope to persuade readers that applying category learning paradigms will be an effective approach to study conceptual change and develop novel instructional techniques.

Knowledge Restructuring

Another recent line of cognition research, that we believe is most ready to be integrated with conceptual change research, is

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

research on knowledge restructuring in feature-based category learning conducted by Lewandowsky and colleagues. This research relates to conceptual change in that it examines the processes underlying learners' ability to shift their classification strategies (e.g., the mappings of features to categories) after an initial learning session. The processes of shifting classification strategies have been examined with both familiar (e.g., Little, Lewandowsky, & Heit, 2006, with ad hoc categories for words) and artificial categories (e.g., Kalish, Lewandowsky, & Davies, 2005, with category boundaries defined by arbitrary functions).

Knowledge restructuring is revealed through behavioral patterns similar to behavioral patterns described in conceptual change theories. First, just as learners struggle to change their conceptual knowledge depending on instructional input, category knowledge is resistant to restructuring. For example, Kalish and colleagues (2005) show that when subjects adopt a suboptimal categorization strategy (e.g., a particular partitioning of the feature-space), subjects require *both* information about an alternative strategy, and error-based feedback of the suboptimality of their initial strategy to induce a strategy change.

Another parallel finding between knowledge restructuring and conceptual change is reflected in the role of working memory for these learning processes. Sewell and Lewandowsky (2012, Experiment 2) demonstrate that knowledge restructuring requires ample working memory resources. Likewise, Muller and colleagues (2008) showed that overcoming naïve theories of forces to learn Newtonian mechanics required increases in cognitive load. Thus, if learners experience high cognitive load when processing and interpreting information based on their existing knowledge or concepts, they might start searching for a way to decrease the load. This search consequently might increase the likelihood that they restructure their knowledge or change their concepts in accordance with the presented (instructional) information.

Further, Sewell and Lewandowsky's (2011, 2012) work parallels theoretical and empirical work in the conceptual change literature (e.g., diSessa et al., 2004; Ohlsson, 2011; Schneider & Hardy, 2013; Smith et al., 1994; Vosniadou & Brewer, 1992). First, both literatures focus on how knowledge is composed of multiple rules or principles. Second, reasoning and problem solving entails online coordination of these knowledge components. Third, conceptual change and knowledge restructuring often require developing new ways of coordinating knowledge components. We now explain each parallel in turn.

Sewell and Lewandowsky (2011, 2012) used an inductive classification task wherein the set of exemplars from two categories compose a two-dimensional feature space (e.g., rectangles of varying heights and widths). To properly classify exemplars, subjects must have learned the rules that divide the space into the categories (e.g., rectangles wider than some value are category A, narrower are category B). However, in their task the entire feature space could not be divided between two categories with a single classification rule (like the one just described), but multiple partial rules must be coordinated and applied to different areas of the feature space (e.g., one rule applied to short rectangles, another to tall rectangles). Subjects were trained in one of two ways to coordinate the rules that led to equivalent classifications within the regions of space experienced during training, but yield different classifications of novel regions of feature space at testing. That is, the two coordination strategies foster different patterns of transfer. After training with the initial rule coordination strategy, subjects hear about the alternative strategy and then must classify transfer items accordingly.

As learners must recoordinate multiple rules to classify transfer items in Sewell and Lewandowsky (2011, 2012); Smith and colleagues (1994) focus on how novice learners must recoordinate multiple pieces of causal knowledge to solve physics problems and explain physical phenomena in line with physical principles. For example, when asked to explain the role of spokes on a bicycle wheel in supporting a bicycle, novices attempt to relate multiple (naïve) causal rules to the various surface features of the bicycle and wheel. However, their explanations deviate from experts' explanations. Smith and colleagues (1994) argue that as learners accrue knowledge, knowledge is still composed of multiple fragments, but knowing the proper causal principles allows the learner to recoordinate their causal knowledge. If successful, the recoordination enables a learner to focus on the right relationships between principles and surface features. It is not that conceptual networks develop a holistic coherent structure with increased expertise (as suggested by the work above), but that expert *coordination of* knowledge pieces becomes more coherent, efficient, and explanatory. This account of knowledge restructuring allows us to advance the main argument of this article in multiple ways.

First, Smith and colleagues (1994) rejected one possible interpretation of the relational shift in education (such as shown by Chi et al., 1981). According to this interpretation, novices show shallow reasoning that only considers the surface features, while experts can essentially zoom right past the surface features and think in purely abstract terms. Smith and colleagues propose instead that novices and experts both infer deep structures from surface features (see also Chi & VanLehn, 2012), and seek to explain the surface features with causal principles. Thus, in Smith and colleagues' view, the novice-expert shift is about how recoordination of causal knowledge enables (a) inferring the proper causal relations from disparate surface features and (b) solving novel problems using similar strategies as the ones learners were using as novices (that granted, are augmented with more experience).[8] Likewise, the empirical analyses of Schneider and Hardy (2013) suggest that there are multiple (but not unlimited) ways depending on the starting point (i.e., the prior knowledge) and the input (i.e., the instruction) of how novices change their conceptual knowledge over time. One open question is whether the recoordination of causal relations is a tightly coupled process with how relational structures are rerepresented as more abstract and interconnected with increased experience in a domain (Gentner, 2010; Lachner & Nückles, 2015; Pillay, 1999; Quinn, Pegg, & Panizzon, 2009), or whether the more extreme proposal (approximating the view of Smith and colleagues) is true: that conceptual networks do not truly develop a coherent holistic structure in long-term memory.

Second, the proposition that relational knowledge changes perception can be contrasted with the work of Sewell and Lewandowsky (2012) to highlight that unlike "flat" featural representations, relational knowledge is hierarchical. In hierarchical relational knowledge, features exist as instantiations of the deeper structure that explain why the features are the way they are. Further, direct perception of the deep

[8] Building on the "knowledge in pieces" account, Wagner (2006) argues that transfer does not require the increased abstraction of relational knowledge per se, but a refinement of knowledge that specifically accounts for contextual variation.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

structure is impossible (e.g., seeing the forces that enable bicycle spokes to support the wheel without seeing the material spokes), it can only be inferred by different kinds of relations among the surface features (Chi & VanLehn, 2012). Even idealized representations like formalized rules and symbolic equations have surface features to instantiate abstract relations (and are themselves objects of perception, Landy & Goldstone, 2007), idealized representations simply try to do so with as few surface features as possible (Belenky & Schalk, 2014).[9]

Thus, given the hierarchical relationship between relations and features, one way of how the knowledge restructuring paradigm could further inform conceptual change research would be to examine the rearrangement of a system of relations that a set of surface features instantiate (and further, how the perception of those features themselves then change), perhaps through mechanisms like schema-refinement and elaboration (Corral & Jones, 2014). This research would expand upon the success of utilizing structural alignment to foster conceptual change (Chi et al., 2012; Gadgil et al., 2012; Jacobson et al., 2011; Opfer & Doumas, 2008) by offering a more precise characterization of the step-by-step process of how relational knowledge is reorganized.

Advancing Theories of Cognition With Conceptual Change Research

In the penultimate section of the article, we argue how considering the challenges experienced by educational psychologists should inform the work of cognitive psychologists. More specifically, to properly characterize the acquisition of structured domain-knowledge, cognitive research needs to expand its focus on learning and reasoning with coherent sets of relational categories. A strength of the cognitive approach to categorization is the use of fully specified implementable formal models of category learning. However, it is unclear to what degree current cognitive models of concept learning (even ones aimed at relational categories, e.g., Doumas et al., 2008; Goodman, Tenenbaum, Griffiths, & Feldman, 2008; Kemp, 2012; Tomlinson & Love, 2006) can explain transitions from naïve to expert networks of conceptions that characterize the process of learning a (scientific) domain.

Rehder and Murphy (2003) developed a connectionist model to simulate prior-knowledge effects in categorization tasks. They analyzed both cases when prior knowledge is advantageous, and when it conflicts with new information. While successful in accounting for many findings, their model does not represent knowledge as explicit relational structures, and so cannot account for many critical conceptual change phenomena. Kemp, Tenenbaum, Niyogi, and Griffiths (2010) have developed a modeling framework to explicitly capture the relations among categories and features that comprise domain theories (e.g., Carey, 1985; and Blouw, Solodkin, Thagard, & Eliasmith, 2015, for a neurally inspired modeling framework that has the power to represent theories as recursively bound collections of rules). However, this work has not yet been applied to the problem of how such domain theories are restructured by educational input. Goldstone and Rogosky (2002) have developed a model to simulate translating across conceptual systems, and Friedman and colleagues (Friedman, 2012; Friedman, Barbella, & Forbus, 2012) have used the Structure-Mapping Engine to simulate the role of analogy and self-explanation in conceptual change in child development and science learning. While these lines of research show promise, both the empirical and theoretical work is far from complete. If cognitive psychologists take on the challenge of designing categorization tasks to investigate the acquisition and transformation of (relational) conceptual systems, then this will enable the design of computational models that allow scrutinizing underlying cognitive processes and their interplay.[10]

Examining how conceptual knowledge is transformed has been a major topic of cognitive research in young children (e.g., Carey, 1985; Vosniadou & Brewer, 1992). For example, researchers (e.g., Hatano & Inagaki, 1994; Medin & Atran, 2004) have extensively investigated how children's conceptions of life and nature (i.e., their "naïve biology") change with development and differ across cultures. However, once children get older, the topic of conceptual change has primarily been studied by educational researchers who tend to use different research methods (such as student interviews). To complement these methods, we think the use of traditional cognitive psychology paradigms will be fruitful. However, cognitive psychologists will have to adapt their methods to meet the affordances of studying learning of (complex systems of) relational categories and conceptual change.

Summary, Recommendations, and Conclusion

The primary argument of this article that without a focus on relational categories, the applicability of category learning research to education will be limited. Relational representations are formally distinct from feature-based representations, and this formal distinction has critical implications for the challenges facing students and educators. First, processing complex relational structures is more working memory intensive than processing feature-based structures (Forbus et al., 1995; Hummel & Holyoak, 1997; Halford et al., 2012). Second, categorizing a problem or phenomenon by its relational structure may entail an abstract representation that ignores superficial contextual differences. However, first learned relations are often encoded in a context-specific manner, and therefore, extensive experience might be necessary to form the abstract schemas needed for knowledge transfer (Chi et al., 1981; Doumas et al., 2008; Gentner, 2010). Third, relational categories are defined by how they interrelate, and interrelated relational categories compose the coherent conceptual systems that define domain knowledge. Students bring naïve conceptual systems to the learning task that must be transformed to expert conceptions, but naïve conceptions are often stubborn to transformation.

Further, we argued that it is not enough to just examine relational reasoning, but relational *categorization* needs to be a spe-

[9] To envision the difference between hierarchical and flat knowledge representations, contrast the inability to directly display a relational abstraction with a feature-based category prototypical exemplar that truly can itself be the central tendency "abstraction" of the category (e.g., the prototype dot pattern from Posner & Keele, 1968).

[10] In a way, we are updating the 30-year old critique of the field of categorization by Murphy and Medin (1985). They claimed that cognitive theories have to account for how prior knowledge affects category learning. However, with the benefit of 30 more years of research, we are arguing for the specific focus on how prior relational knowledge conflicts with educational content.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

cific focus.[11] Evidence suggests that the biggest obstacle to problem solving is not learning general problem-solving procedures, but having the requisite knowledge to properly categorize problems to know what solution procedure to apply (e.g., Goldwater & Gentner, 2015; Rohrer & Pashler, 2010; Taconis et al., 2001).

Taking these last two paragraphs together, our overarching recommendation that we have for both cognitive and educational psychologists is to use category-learning paradigms to examine the learning and transformation of systems of relational concepts. For example, experiments could utilize relational category learning paradigms with multiple phases that force a process of conceptual change. Further, by "category learning paradigms" we mean more than just the standard classification tasks, but also problem-solving, communication, and inference-based learning tasks often used for feature-based categories (see Markman & Ross, 2003, for review). Advancing such a research program will benefit the field of cognitive psychology by challenging the scope of formal models of categorization to include these data. At the same time, it will benefit educational psychologists by providing a new source of rich data to test theories of conceptual change, and by providing new potential methods for educational interventions.

The research program to examine relational conceptual change will be aided by systematically combing the education literature to identify different kinds of relational category structures, and different ways of how surface features may be correlated with, or completely orthogonal to relational structures (e.g., the cognitive task analyses of Clark et al., 2007). Just as the feature-based category literature has long had a focus on how different kinds of feature-based category structures vary in their difficulty to learn (e.g., Shepard et al., 1961), many insights into relational category learning could be gained by examining different forms of structures (see Corral & Jones, 2014, and Tenenbaum, Kemp, Griffiths, & Goodman, 2011).

The search for relevant category structures will benefit from working with domain-experts. For example, Shipley and colleagues (Resnick & Shipley, 2013; also see Shipley, Tikoff, Manduca, & Ormand, 2013; Jee et al., 2013) collaborated with practicing geoscientists and this collaboration has revealed perceptual and spatial skills vital to geoscience expertise that vision scientists have left unexamined. We cannot predict what kinds of skills or knowledge will be found when doing analogous work in the conceptual domain, but this kind of interdisciplinary collaborative research is highly promising. Even further, interdisciplinary research may be necessary. Unlike the conceptual change research with children wherein as psychologists we know the target concept, we probably do not have comprehensive knowledge of the kinds of conceptual structures students learn while developing expertise in other fields. To design learning materials and develop reliable assessment techniques, comprehensive expert knowledge of a domain is essential.

While we believe these lines of research would greatly advance both literatures, we hope that this article inspires readers to identify even more viable paths of research that have the potential to mutually inform the cognitive and the education literatures on relational category learning and conceptual change. Nevertheless, we need to highlight a remaining challenge that these lines of research alone will not surmount. To fully integrate the research by cognitive and educational psychologists, there must be additional research into the role of the learning environment. Many in the educational community doubt the utility of laboratory research findings because of the environment in which it was done (at least until it can be corroborated in the classroom). We, however, agree with Stern and Schneider (2010) who pointed out that it is highly productive to zoom in and out (i.e., going from the lab to the classroom and back) and using different methods (from interviews to highly controlled lab studies, computational models, or neuropsychological measures) when doing research on complex learning processes like conceptual change.

The analogical reasoning literature is a good example of this complementary approach. It has been successful in developing precise theoretical and computational cognitive models (e.g., French, 2002; Klenk & Forbus, 2013), in designing learning intervention for real classrooms (especially, in form of comparisons and contrasts; e.g., Schwartz et al., 2011; Rittle-Johnson & Star, 2011; Ziegler & Stern, 2014), and to serve as conceptual framework to study the effectiveness of specific classroom practices across countries (e.g., Richland, Zur, & Holyoak, 2007). Therefore, when the knowledge being learned in laboratory settings (even when completely artificial) is of a hierarchical relational nature and specifically designed to mimic the complex conceptual change phenomena observed by educational psychologists, there is great opportunity for application.

It is broadly assumed that the approaches of cognitive and educational psychology to category and concept learning are complementary. The cognitive psychologists reverse-engineer the learning mechanisms and representational forms, while the educational psychologists design interventions to facilitate learning of particular content informed by the research of cognitive psychologists. We are arguing that it is possible for the fields to strive not only for complementarity, but for synergy. By focusing on the learning and transformation of relational conceptual systems, there can be a unified research approach. When investigating the same kinds of concepts, to foster learning and transfer, experiments revealing the underlying mechanisms and representations do not need to be different from experiments examining instructional methods and learning environments.

[11] In this regard, we are adding to the recent review by Dumas, Alexander, and Grossnickle (2013) that discusses a variety of forms of relational reasoning in education (i.e., analogy, anomaly, antinomy, and antithesis).

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48,* 87–113. http://dx.doi.org/10.1080/00461520.2013.775712

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56,* 149–178. http://dx.doi.org/10.1146/annurev.psych.56.091103.070217

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9,* 161–185. http://dx.doi.org/10.1007/s11409-013-9107-6

Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition, 15,* 397–418. http://dx.doi.org/10.3758/BF03197730

Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology, 35,* 99–134. http://dx.doi.org/10.1006/cogp.1998.0675

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery-approach goals in preparation for future learning. *Journal of the Learning Sciences, 21,* 399–432. http://dx.doi.org/10.1080/10508406.2011.651232

Belenky, D. M., & Schalk, L. (2014). The effects of idealized and grounded materials on learning, transfer, and interest: An organizing framework for categorizing external knowledge representations. *Educational Psychology Review, 26,* 27–50. http://dx.doi.org/10.1007/s10648-014-9251-9

Billman, D., & Knutson, J. (1996). Unsupervised concept learning and value systematicity: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 458–475. http://dx.doi.org/10.1037/0278-7393.22.2.458

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64,* 417–444. http://dx.doi.org/10.1146/annurev-psych-113011-143823

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition, 60,* 1–29. http://dx.doi.org/10.1016/0010-0277(95)00699-0

Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2015). Concepts as semantic pointers: A framework and computational model. [early view]. *Cognitive Science*. http://dx.doi.org/10.1111/cogs.12265

Braithwaite, D. W., & Goldstone, R. L. (2014). Benefits of variation increase with preparation. *Proceedings of the Thirty-Sixth Annual Conference of the Cognitive Science Society* (pp. 1940–1945). Quebec City, Canada: Cognitive Science Society.

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

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24,* 61–100.

Burgess, P. W., Gilbert, S. J., Okuda, J., & Simons, J. S. (2006). Rostral prefrontal brain regions (Area 10): A gateway between inner thought and the external world? In W. Prinz & N. Sebanz (Eds.), *Disorders of volition* (pp. 373–396). Cambridge, MA: MIT Press.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1118–1133. http://dx.doi.org/10.1037/a0019902

Carey, S. (1985). *Conceptual development in childhood*. Cambridge, MA: MIT Press.

Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology, 21,* 13–19. http://dx.doi.org/10.1016/S0193-3973(99)00046-5

Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195367638.001.0001

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. B. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education, 29,* 3–20. http://dx.doi.org/10.2307/749715

Carvalho, P. F., & Goldstone, R. L. (2014). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & Cognition, 42,* 481–495. http://dx.doi.org/10.3758/s13421-013-0371-0

Carvalho, P. F., & Goldstone, R. L. (2015). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic Bulletin & Review, 22,* 281–288. http://dx.doi.org/10.3758/s13423-014-0676-4

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87,* 5–17. http://dx.doi.org/10.1037/0022-0663.87.1.5

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127,* 355–376. http://dx.doi.org/10.1037/0096-3445.127.4.355

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1147–1156. http://dx.doi.org/10.1037/0278-7393.15.6.1147

Chan, J., & Schunn, C. (2015). The impact of analogies on creative concept generation: Lessons from an in vivo study in engineering design. *Cognitive Science, 39,* 126–155. http://dx.doi.org/10.1111/cogs.12127

Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology, 107,* 689–704. http://dx.doi.org/10.1037/edu0000018

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145–182. http://dx.doi.org/10.1207/s15516709cog1302_1

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152. http://dx.doi.org/10.1207/s15516709cog0502_2

Chi, M. T. H., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science, 36,* 1–61. http://dx.doi.org/10.1111/j.1551-6709.2011.01207.x

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist, 47,* 177–188. http://dx.doi.org/10.1080/00461520.2012.695709

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11,* 356–373. http://dx.doi.org/10.1080/15248371003700015

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage, 14,* 1136–1149. http://dx.doi.org/10.1006/nimg.2001.0922

Clark, R. E., Feldon, D., van Merriënboer, J. J. G., Yates, K., & Early, S. (2007). Cognitive task analysis. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communication and technology* (3rd ed., pp. 577–593). Mahwah, NJ: Erlbaum.

Clement, C. A., & Gentner, D. (1991). Systematicity as a selection constraint in analogical mapping. *Cognitive Science, 15,* 89–132. http://dx.doi.org/10.1207/s15516709cog1501_3

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50,* 66–71. http://dx.doi.org/10.1119/1.12989

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30,* 1241–1257. http://dx.doi.org/10.1002/tea.3660301007

Cohen, N. A. (2009). Relational memory and the hippocampus. *Oral presentation given at the University of Texas at Austin Department of Psychology*.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240–247. http://dx.doi.org/10.1016/S0022-5371(69)80069-1

Corral, D., & Jones, M. (2014). The effects of relational structure on analogical learning. *Cognition, 132,* 280–300. http://dx.doi.org/10.1016/j.cognition.2014.04.007

Craig, S., & Lewandowsky, S. (2012). Whichever way you choose to categorize, working memory helps you learn. *The Quarterly Journal of Experimental Psychology, 65,* 439–464. http://dx.doi.org/10.1080/17470218.2011.608854

Davis, T., & Love, B. C. (2010). Memory for category information is idealized through contrast with competing options. *Psychological Science, 21,* 234–242. http://dx.doi.org/10.1177/0956797609357712

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Davis, T., Love, B. C., & Preston, A. R. (2012). Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. *Cerebral Cortex, 22,* 260–273. http://dx.doi.org/10.1093/cercor/bhr036

Day, S. B., & Gentner, D. (2007). Nonintentional analogical inference in text comprehension. *Memory & Cognition, 35,* 39–49. http://dx.doi.org/10.3758/BF03195940

Day, S. B., & Goldstone, R. L. (2011). Analogical transfer from a simulated physical system. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 551–567. http://dx.doi.org/10.1037/a0022333

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist, 47,* 153–176. http://dx.doi.org/10.1080/00461520.2012.696438

De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., & Verschaffel, L. (2011). Abstract or concrete examples in learning mathematics: A replication and elaboration of Kaminski, Sloutsky, and Heckler‘s study. *Journal for Research in Mathematics Education, 42,* 109–126.

De Corte, E. (2003). Transfer as the productive use of acquired knowledge, skills, and motivations. *Current Directions in Psychological Science, 12,* 142–146. http://dx.doi.org/10.1111/1467-8721.01250

Delosh, E. L. (1997). *Effect of mnemonic variables on function and category learning*. Unpublished Dissertation. Purdue University, U.S.

diSessa, A., Gillespie, N. M., & Esterly, J. B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science, 28,* 843–900. http://dx.doi.org/10.1207/s15516709cog2806_1

Dominey, P. F., Lelekov, T., Ventre-Dominey, J., & Jeannerod, M. (1998). Dissociable processes for learning the surface structure and abstract structure of sensorimotor sequences. *Journal of Cognitive Neuroscience, 10,* 734–751. http://dx.doi.org/10.1162/089892998563130

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84,* 795–805. http://dx.doi.org/10.1037/0021-9010.84.5.795

Doumas, L. A., & Hummel, J. E. (2013). Comparison and mapping facilitate relation discovery and predication. *PLoS ONE, 8,* e63889. http://dx.doi.org/10.1371/journal.pone.0063889

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115,* 1–43. http://dx.doi.org/10.1037/0033-295X.115.1.1

Dumas, D., Alexander, P. A., & Grossnickle, E. M. (2013). Relational reasoning and its manifestations in the educational context: A systematic review of the literature. *Educational Psychology Review, 25,* 391–427. http://dx.doi.org/10.1007/s10648-013-9224-4

Dunbar, K., & Blanchette, I. (2001). The in vivo/in vitro approach to cognition: The case of analogy. *Trends in Cognitive Sciences, 5,* 334–339. http://dx.doi.org/10.1016/S1364-6613(00)01698-3

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences, 14,* 172–179. http://dx.doi.org/10.1016/j.tics.2010.01.004

Ebenezer, J. V., Chacko, S., Kaya, O. N., Koya, S. K., & Ebenezer, D. L. (2010). The effects of common knowledge construction model sequence of lessons on science achievement and relational conceptual change. *Journal of Research in Science Education, 47,* 25–46.

Ebenezer, J. V., & Gaskell, P. J. (1995). Relational conceptual change in solution chemistry. *Science Education, 79,* 1–17. http://dx.doi.org/10.1002/sce.3730790102

Edelsbrunner, P. A., Schalk, L., Schumacher, R., & Stern, E. (2015). Pathways of conceptual change: Investigating the influence of experimentation skills on conceptual knowledge development in early science education. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 620–625). Austin: TX: Cognitive Science Society.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140. http://dx.doi.org/10.1037/0096-3445.127.2.107

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41,* 1–63. http://dx.doi.org/10.1016/0004-3702(89)90077-5

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A Model of Similarity-Based Retrieval. *Cognitive Science, 19,* 141–205. http://dx.doi.org/10.1207/s15516709cog1902_1

Fox, E. J., & Sullivan, H. J. (2007). Comparing strategies for teaching abstract concepts in an online tutorial. *Journal of Educational Computing Research, 37,* 307–330. http://dx.doi.org/10.2190/EC.37.3.e

French, R. M. (2002). The computational modeling of analogy-making. *Trends in Cognitive Sciences, 6,* 200–205. http://dx.doi.org/10.1016/S1364-6613(02)01882-X

Friedman, S. E. (2012). *Computational conceptual change: An explanation based approach*. Doctoral dissertation, Northwestern University, Department of Electrical Engineering and Computer Science, Evanston, IL.

Friedman, S. E., Barbella, D., & Forbus, K. (2012). Revising domain knowledge with cross-domain analogy. *Advances in Cognitive Systems, 2,* 13–24.

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review, 26,* 9–25. http://dx.doi.org/10.1007/s10648-014-9249-3

Gabora, L., Rosch, E., & Aerts, D. (2008). Toward and ecological theory of concepts. *Ecological Psychology, 20,* 84–116. http://dx.doi.org/10.1080/10407410701766676

Gadgil, S., Nokes-Malach, T. J., & Chi, M. T. H. (2012). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction, 22,* 47–61. http://dx.doi.org/10.1016/j.learninstruc.2011.06.002

Gelman, S. A., & Legare, C. H. (2011). Concepts and folk theories. *Annual Review of Anthropology, 40,* 379–398. http://dx.doi.org/10.1146/annurev-anthro-081309-145822

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory, 4,* 161–178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

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

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.

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

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82,* 1173–1188. http://dx.doi.org/10.1111/j.1467-8624.2011.01599.x

Gentner, D., & Asmuth, J. (2008). Can relationality be distinguished from abstractness in noun mutability. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 863–868). Austin, TX: Cognitive Science Society.

Gentner, D., & Boroditsky, L. (2001). Individuation, relational relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cam-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

bridge, United Kingdom: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511620669.010

Gentner, D., Brem, S., Ferguson, R., & Wolff, P. (1997). Analogy and creativity in the works of Johannes Kepler. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative thought: An investigation of conceptual structures and processes* (pp. 403–459). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10227-016

Gentner, D., & France, I. M. (1988). The verb mutability effect: Studies of the combinatorial semantics of nouns and verbs. In S. L. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence* (pp. 343–382). San Mateo, CA: Morgan Kaufman. http://dx.doi.org/10.1016/B978-0-08-051013-2.50018-5

Gentner, D. & Smith, L. (2012). Analogical reasoning. In V. S. Ramachandran (Eds.), *Encyclopedia of Human Behavior* (2nd Ed.) (pp. 130–136). Oxford, UK: Elsevier.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 151–175). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/11156-009

Gentner, D., Levine, S. C., Ping, R., Isaia, A., Dhillon, S., Bradley, C., & Honke, G. (2015). Rapid learning in a children's museum via analogical comparison. [early view]. *Cognitive Science*. http://dx.doi.org/10.1111/cogs.12248

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8,* 285–307. http://dx.doi.org/10.1080/15248370701446434

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95,* 393–408. http://dx.doi.org/10.1037/0022-0663.95.2.393

Gentner, D., Loewenstein, J., Thompson, L., & Forbus, K. D. (2009). Reviving inert knowledge: Analogical abstraction supports relational retrieval of past events. *Cognitive Science, 33,* 1343–1382. http://dx.doi.org/10.1111/j.1551-6709.2009.01070.x

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52,* 45–56. http://dx.doi.org/10.1037/0003-066X.52.1.45

Gentner, D., Ozyürek, A., Gürcanli, O., & Goldin-Meadow, S. (2013). Spatial language facilitates spatial cognition: Evidence from children who lack language input. *Cognition, 127,* 318–330. http://dx.doi.org/10.1016/j.cognition.2013.01.003

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225–277). London: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511983689.008

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. http://dx.doi.org/10.1006/cogp.1993.1013

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306–355. http://dx.doi.org/10.1016/0010-0285(80)90013-4

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38. http://dx.doi.org/10.1016/0010-0285(83)90002-6

Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental "roots" in mature biological knowledge. *Psychological Science, 20,* 480–487. http://dx.doi.org/10.1111/j.1467-9280.2009.02320.x

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52,* 125–157. http://dx.doi.org/10.1016/0010-0277(94)90065-5

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory & Cognition, 24,* 608–628. http://dx.doi.org/10.3758/BF03201087

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612. http://dx.doi.org/10.1146/annurev.psych.49.1.585

Goldstone, R. L., & Day, S. B. (2013). Similarity. In H. Pashler (Ed.), *The encyclopedia of mind* (pp. 696–699). Thousand Oaks, CA: SAGE Reference. http://dx.doi.org/10.4135/9781452257044.n257

Goldstone, R. L., Landy, D. H., & Son, J. Y. (2010). The education of perception. *Topics in Cognitive Science, 2,* 265–284. http://dx.doi.org/10.1111/j.1756-8765.2009.01055.x

Goldstone, R. L., & Medin, D. L. (1994). Time course of comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 29–50. http://dx.doi.org/10.1037/0278-7393.20.1.29

Goldstone, R. L., & Rogosky, B. J. (2002). Using relations within conceptual systems to translate across conceptual systems. *Cognition, 84,* 295–320. http://dx.doi.org/10.1016/S0010-0277(02)00053-7

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

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

Goldwater, M. B., & Gentner, D. (2015). On the acquisition of abstract knowledge: Structural alignment and explication in learning causal system categories. *Cognition, 137,* 137–153. http://dx.doi.org/10.1016/j.cognition.2014.12.001

Goldwater, M. B., & Markman, A. B. (2011). Categorizing entities by common role. *Psychonomic Bulletin & Review, 18,* 406–413. http://dx.doi.org/10.3758/s13423-011-0058-0

Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The empirical case for role-governed categories. *Cognition, 118,* 359–376. http://dx.doi.org/10.1016/j.cognition.2010.10.009

Goldwater, M. G., Sibley, D., Gentner, D., LaDue, N., & Libarkin, J. (under revision). Analogical retrieval and the development of science expertise.

Goldwater, M. B., Smith, L., & Hinze, S. (in preparation). Failure of retrieval practice to support spontaneous analogical transfer.

Goodman, N. D., Tenenbaum, J. B., Griffiths, T. L., & Feldman, J. (2008). Compositionality in rational analysis: Grammar-based induction for concept learning. In M. Oaksford & N. Chater (Eds.), *The probabilistic mind: Prospects for rational models of cognition*. Oxford: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199216093.003.0017

Green, A. E., Fugelsang, J. A., Kraemer, D. J., Shamosh, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research, 1096,* 125–137. http://dx.doi.org/10.1016/j.brainres.2006.04.024

Gurlitt, J., Dummel, S., Schuster, S., & Nückles, M. (2012). Differently structured advance organizers lead to different initial schemata and learning outcomes. *Instructional Science, 40,* 351–369. http://dx.doi.org/10.1007/s11251-011-9180-7

Halford, G. S., Andrews, G., Wilson, W. H., & Phillips, S. (2012). Computational models of relational processes in cognitive development. *Cognitive Development, 27,* 481–499. http://dx.doi.org/10.1016/j.cogdev.2012.08.003

Halford, G. S., Baker, R., McCredden, J. E., & Bain, J. D. (2005). How many variables can humans process? *Psychological Science, 16,* 70–76. http://dx.doi.org/10.1111/j.0956-7976.2005.00782.x

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21,* 803–831. http://dx.doi.org/10.1017/S0140525X98001769

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Halford, G. S., Wilson, W. H., & Phillips, S. (2010). Relational knowledge: The foundation of higher cognition. *Trends in Cognitive Sciences, 14,* 497–505. http://dx.doi.org/10.1016/j.tics.2010.08.005

Hambrick, D. Z. (2003). Why are some people more knowledgeable than others? A longitudinal study of knowledge acquisition. *Memory & Cognition, 31,* 902–917. http://dx.doi.org/10.3758/BF03196444

Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., . . . Ladue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *Journal of Experimental Psychology: General, 141,* 397–403. http://dx.doi.org/10.1037/a0025927

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-specific experience and Knowledge in skilled performance. *Current Directions in Psychological Science, 20,* 275–279. http://dx.doi.org/10.1177/0963721411422061

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition, 50,* 171–188. http://dx.doi.org/10.1016/0010-0277(94)90027-2

Heckler, A. F. (2011). The ubiquitous patterns of incorrect answers to science questions: The role of automatic, bottom-up processes. *Psychology of Learning and Motivation-Advances in Research and Theory, 55,* 227–263. http://dx.doi.org/10.1016/B978-0-12-387691-1.00008-9

Heckler, A. F., & Scaife, T. M. (2015). Patterns of response times and response choices to science questions: The influence of relative processing time. *Cognitive Science, 39,* 496–537. http://dx.doi.org/10.1111/cogs.12166

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87,* 18–32. http://dx.doi.org/10.1037/0022-0663.87.1.18

Heit, E. (2001). Background knowledge and models of categorization. In U. Hahn & M. Ramscar (Eds.), *Similarity and categorization* (pp. 155–178). Oxford, United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198506287.003.0009

Heit, E., & Nicholson, S. P. (2010). The opposite of republican: Polarization and political categorization. *Cognitive Science, 34,* 1503–1516. http://dx.doi.org/10.1111/j.1551-6709.2010.01138.x

Higgins, E. J., & Ross, B. H. (2011). Comparisons in category learning: How best to compare for what. In L. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1388–1393). Austin, TX: Cognitive Science Society, Austin.

Holyoak, K. J., & Koh, K. (1987). Surface and structural similarity in analogical transfer. *Memory & Cognition, 15,* 332–340. http://dx.doi.org/10.3758/BF03197035

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13,* 295–355. http://dx.doi.org/10.1207/s15516709cog1303_1

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52,* 35–44. http://dx.doi.org/10.1037/0003-066X.52.1.35

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427–466. http://dx.doi.org/10.1037/0033-295X.104.3.427

Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: Differences between experts and novices. *Complexity, 6,* 41–49. http://dx.doi.org/10.1002/cplx.1027

Jacobson, M. J., Kapur, M., So, H. J., & Lee, J. (2011). The ontologies of complexity and learning about complex systems. *Instructional Science, 39,* 763–783. http://dx.doi.org/10.1007/s11251-010-9147-0

Jacoby, L. L., Wahlheim, C. N., & Coane, J. H. (2010). Test-enhanced learning of natural concepts: Effects on recognition memory, classification, and metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1441–1451. http://dx.doi.org/10.1037/a0020636

Jamrozik, A., Sagi, E., Goldwater, M., & Gentner, D. (2013). Relational words have high metaphoric potential. *Meta4NLP, 2013,* 21.

Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., & Sageman, B. (2013). Finding faults: Analogical comparison supports spatial concept learning in geoscience. *Cognitive Processing, 14,* 175–187. http://dx.doi.org/10.1007/s10339-013-0551-7

Johnson, S. C., & Carey, S. (1998). Knowledge enrichment and conceptual change in folkbiology: Evidence from Williams syndrome. *Cognitive Psychology, 37,* 156–200. http://dx.doi.org/10.1006/cogp.1998.0695

Jones, M., & Love, B. C. (2007). Beyond common features: The role of roles in determining similarity. *Cognitive Psychology, 55,* 196–231. http://dx.doi.org/10.1016/j.cogpsych.2006.09.004

Jung, W., & Hummel, J. E. (2011). Progressive alignment facilitates learning of deterministic but not probabilistic relational categories. In L. Carlson, C. Hölscher, & T. F. Shipley, *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2643–5648). Austin, TX: Cognitive Science Society.

Kalish, M. L., Lewandowsky, S., & Davies, M. (2005). Error-driven knowledge restructuring in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 846–861. http://dx.doi.org/10.1037/0278-7393.31.5.846

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38,* 23–31. http://dx.doi.org/10.1207/S15326985EP3801_4

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320,* 454–455. http://dx.doi.org/10.1126/science.1154659

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26,* 379–424. http://dx.doi.org/10.1080/07370000802212669

Karpicke, J. D., & Aue, W. R. (2015). The testing effect is alive and well with complex materials. *Educational Psychology Review, 27,* 317–326. http://dx.doi.org/10.1007/s10648-015-9309-3

Kellman, P. J., & Massey, C. M. (2013). Perceptual learning, cognition, and expertise. *Psychology of Learning and Motivation, 58,* 117–165. http://dx.doi.org/10.1016/B978-0-12-407237-4.00004-9

Kemp, C. (2012). Exploring the conceptual universe. *Psychological Review, 119,* 685–722. http://dx.doi.org/10.1037/a0029347

Kemp, C., Tenenbaum, J. B., Niyogi, S., & Griffiths, T. L. (2010). A probabilistic model of theory formation. *Cognition, 114,* 165–196. http://dx.doi.org/10.1016/j.cognition.2009.09.003

Kittur, A., Hummel, J. E., & Holyoak, K. J. (2004). Feature- vs. relation-defined categories: Probab(alistic)ly not the same. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-sixth Annual Conference of the Cognitive Science Society* (pp. 696–701). Mahwah, NJ: Erlbaum.

Klenk, M., & Forbus, K. (2013). Exploiting persistent mappings in cross-domain analogical learning of physical domains. *Artificial Intelligence, 195,* 398–417. http://dx.doi.org/10.1016/j.artint.2012.11.002

Kloos, H., Fisher, A., & Van Orden, G. C. (2010). Situated naïve physics: Task constraints decide what children know about density. *Journal of Experimental Psychology: General, 139,* 625–637. http://dx.doi.org/10.1037/a0020977

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36,* 757–798. http://dx.doi.org/10.1111/j.1551-6709.2012.01245.x

Koedinger, K. R., & Roll, I. (2012). Learning to think: Cognitive mechanisms of knowledge transfer. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 789–806). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199734689.013.0040

Kok, E. M., de Bruin, A. B. H., Robben, S. G. F., & van Merriënboer, J. J. G. (2013). Learning radiological appearances of diseases: Does

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

comparison help? *Learning and Instruction, 23,* 90–97. http://dx.doi.org/10.1016/j.learninstruc.2012.07.004

Koponen, I. T., & Pehkonen, M. (2010). Coherent knowledge structures of physics represented as concept networks in teacher education. *Science & Education, 19,* 259–282. http://dx.doi.org/10.1007/s11191-009-9200-z

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19,* 585–592. http://dx.doi.org/10.1111/j.1467-9280.2008.02127.x

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44. http://dx.doi.org/10.1037/0033-295X.99.1.22

Kurtz, K. J., & Loewenstein, J. (2007). Converging on a new role for analogy in problem solving and retrieval: When two problems are better than one. *Memory & Cognition, 35,* 334–341. http://dx.doi.org/10.3758/BF03193454

Kurtz, K. J. (2013). Learning toward analogical transfer: The category status hypothesis. *Oral presentation at the Third International Conference on Analogy at the University of Burgundy, Dijon, France*.

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 1303–1310. http://dx.doi.org/10.1037/a0031847

Kurtz, K. J., & Gentner, D. (2013). Detecting anomalous features in complex stimuli: The role of structured comparison. *Journal of Experimental Psychology: Applied, 19,* 219–232. http://dx.doi.org/10.1037/a0034395

Kurtz, K. J. (2007). The divergent autoencoder (DIVA) model of category learning. *Psychonomic Bulletin & Review, 14,* 560–576. http://dx.doi.org/10.3758/BF03196806

Lachner, A., Gurlitt, J., & Nückles, M. (2012). A graph-oriented approach to measuring expertise-detecting structural differences between experts and intermediates. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 653–658).

Lachner, A., & Nückles, M. (2015). Bothered by abstractness or engaged by cohesion? Experts' explanations enhance novices' deep-learning. *Journal of Experimental Psychology: Applied, 21,* 101–115. http://dx.doi.org/10.1037/xap0000038

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 720–733. http://dx.doi.org/10.1037/0278-7393.33.4.720

Lassaline, M. E., & Murphy, G. L. (1996). Induction and category coherence. *Psychonomic Bulletin & Review, 3,* 95–99. http://dx.doi.org/10.3758/BF03210747

Leahy, W., Hanham, J., & Sweller, J. (2015). High element interactivity information during problem solving may lead to failure to obtain the testing effect. *Educational Psychology Review, 27,* 291–304. http://dx.doi.org/10.1007/s10648-015-9296-4

Libarkin, J. C. (2001). Development of an assessment of student conception of the nature of science. *Journal of Geoscience Education, 49,* 435–442.

Libarkin, J. C. (2008). *Concept inventories in higher education science*. Prepared for the National Research Council Promising Practices in Undergraduate STEM Education Workshop 2 (Washington, DC, Oct. 13–14, 2008).

Linsey, J. S., Wood, K. L., & Markman, A. B. (2008). Modality and representation in analogy. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 22,* 85–100. http://dx.doi.org/10.1017/S0890060408000061

Little, D. R., Lewandowsky, S., & Heit, E. (2006). Ad hoc category restructuring. *Memory & Cognition, 34,* 1398–1413. http://dx.doi.org/10.3758/BF03195905

Little, J. L., & McDaniel, M. A. (2015). Individual differences in category learning: Memorization versus rule abstraction. *Memory & Cognition, 43,* 283–297. http://dx.doi.org/10.3758/s13421-014-0475-1

Loewenstein, J. (2010). How one's hook is baited matters for catching an analogy. *Psychology of Learning and Motivation, 53,* 149–182. http://dx.doi.org/10.1016/S0079-7421(10)53004-4

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development, 2,* 189–219. http://dx.doi.org/10.1207/S15327647JCD0202_4

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50,* 315–353. http://dx.doi.org/10.1016/j.cogpsych.2004.09.004

Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction, 34,* 74–85.

Love, B. C. (2003). The multifaceted nature of unsupervised category learning. *Psychonomic Bulletin & Review, 10,* 190–197. http://dx.doi.org/10.3758/BF03196484

Love, B. C., & Markman, A. B. (2003). The nonindependence of stimulus properties in human category learning. *Memory & Cognition, 31,* 790–799. http://dx.doi.org/10.3758/BF03196117

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review, 111,* 309–332. http://dx.doi.org/10.1037/0033-295X.111.2.309

Markman, A. B. (1997). Constraints on analogical inference. *Cognitive Science, 21,* 373–418. http://dx.doi.org/10.1207/s15516709cog2104_1

Markman, A. B. (1999). *Knowledge representation*. Mahwah, NJ: Erlbaum.

Markman, A. B., & Gentner, D. (1993a). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431–467. http://dx.doi.org/10.1006/cogp.1993.1011

Markman, A. B., & Gentner, D. (1993b). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32,* 517–535. http://dx.doi.org/10.1006/jmla.1993.1027

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129,* 592–613. http://dx.doi.org/10.1037/0033-2909.129.4.592

Markman, A. B., & Stilwell, C. H. (2001). Role-governed categories. *Journal of Experimental & Theoretical Artificial Intelligence, 13,* 329–358. http://dx.doi.org/10.1080/09528130110100252

Mathy, F. (2010). The long-term effect of relational information in classification learning. *European Journal of Cognitive Psychology, 22,* 360–390. http://dx.doi.org/10.1080/09541440902903603

Mayer, R. E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories, and templates. *Instructional Science, 10,* 135–175. http://dx.doi.org/10.1007/BF00132515

Mayer, R. E. (1982). Memory for algebra story problems. *Journal of Educational Psychology, 74,* 199–216. http://dx.doi.org/10.1037/0022-0663.74.2.199

McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: Stable tendencies for learning exemplars versus abstracting rules. *Journal of Experimental Psychology: General, 143,* 668–693. http://dx.doi.org/10.1037/a0032963

McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in middle-school science: Successful transfer performance on classroom exams. *Applied Cognitive Psychology, 27,* 360–372. http://dx.doi.org/10.1002/acp.2914

McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, 37,* 547–559. http://dx.doi.org/10.3758/BF03192726

Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review, 111,* 960–983. http://dx.doi.org/10.1037/0033-295X.111.4.960

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cog-*

nitive Psychology, 19, 242–279. http://dx.doi.org/10.1016/0010-0285(87)90012-0

Mestre, J. P., Docktor, J. L., Strand, N. E., & Ross, B. H. (2011). Conceptual problem solving in physics. *Cognition and education: The psychology of learning and motivation: Advances in research and theory, 55,* 269–298.

Ming, N. C. (2009). Analogies vs. contrasts: A comparison of their learning benefits. In B. Kokinov, D. Gentner, & K. Holyoak (Eds.), *New frontiers in analogy research: Proceedings of the second international conference on analogy* (pp. 338–347). Sofia, Bulgaria: New Bulgarian University.

Moran, S., Bereby-Meyer, Y., & Bazerman, M. (2008). Stretching the effectiveness of analogical training in negotiations: Teaching diverse principles for creating value. *Negotiation and Conflict Management Research, 1,* 99–134. http://dx.doi.org/10.1111/j.1750-4716.2007.00006.x

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education, 92,* 278–296. http://dx.doi.org/10.1002/sce.20244

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 904–919. http://dx.doi.org/10.1037/0278-7393.20.4.904

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289–316. http://dx.doi.org/10.1037/0033-295X.92.3.289

Nokes, T. J. (2009). Mechanisms of knowledge transfer. *Thinking & Reasoning, 15,* 1–36. http://dx.doi.org/10.1080/13546780802490186

Nokes-Malach, T. J., & Mestre, J. P. (2013). Toward a model of transfer as sense-making. *Educational Psychologist, 48,* 184–207. http://dx.doi.org/10.1080/00461520.2013.807556

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 101,* 53–79. http://dx.doi.org/10.1037/0033-295X.101.1.53

Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 510–520. http://dx.doi.org/10.1037/0278-7393.14.3.510

O'Connor, C. M., Cree, G. S., & McRae, K. (2009). Conceptual hierarchies in a flat attractor network: Dynamics of learning and computations. *Cognitive Science, 33,* 665–708. http://dx.doi.org/10.1111/j.1551-6709.2009.01024.x

Ohlsson, S. (2009). Resubsumption: A possible mechanism for conceptual change and belief revision. *Educational Psychologist, 44,* 20–40. http://dx.doi.org/10.1080/00461520802616267

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511780295

Ohlsson, S. (2013). Beyond evidence-based belief formation: How normative ideas have constrained conceptual change research. *Frontline Learning Research, 2,* 70–85.

Opfer, J. E., & Doumas, L. A. A. (2008). Analogy and conceptual change in childhood. *Behavioral and Brain Sciences, 31,* 723. http://dx.doi.org/10.1017/S0140525X08005992

Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching, 49,* 744–777. http://dx.doi.org/10.1002/tea.21028

Parkin, B. L., Hellyer, P. J., Leech, R., & Hampshire, A. (2015). Dynamic network mechanisms of relational integration. *The Journal of Neuroscience, 35,* 7660–7673. http://dx.doi.org/10.1523/JNEUROSCI.4956-14.2015

Patalano, A. L., & Ross, B. H. (2007). The role of category coherence in experience-based prediction. *Psychonomic Bulletin & Review, 14,* 629–634. http://dx.doi.org/10.3758/BF03196812

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31,* 109–130. http://dx.doi.org/10.1017/S0140525X08003543

Pillay, H. K. (1999). An analysis of knowledge representation of students in electronic problem tasks. *European Journal of Psychology of Education, 14,* 325–338. http://dx.doi.org/10.1007/BF03173118

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications*. Upper Saddle River, NJ: Prentice Hall.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77,* 353–363.

Quilici, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology, 88,* 144–161. http://dx.doi.org/10.1037/0022-0663.88.1.144

Quinn, F., Pegg, J., & Panizzon, D. (2009). First-year biology students' understandings of meiosis: An investigation using a structural theoretical framework. *International Journal of Science Education, 31,* 1279–1305. http://dx.doi.org/10.1080/09500690801914965

Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2014). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review, 27,* 483–504.

Reed, S. K. (1989). Constraints on the abstraction of solutions. *Journal of Educational Psychology, 81,* 532–540. http://dx.doi.org/10.1037/0022-0663.81.4.532

Rehder, B. (2010). Causal-based categorization: A review. *Psychology of Learning and Motivation, 52,* 39–116. http://dx.doi.org/10.1016/S0079-7421(10)52002-4

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General, 130,* 323–360. http://dx.doi.org/10.1037/0096-3445.130.3.323

Rehder, B., & Murphy, G. L. (2003). A knowledge-resonance (KRES) model of category learning. *Psychonomic Bulletin & Review, 10,* 759–784. http://dx.doi.org/10.3758/BF03196543

Rehder, B., & Ross, B. H. (2001). Abstract coherent categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1261–1275. http://dx.doi.org/10.1037/0278-7393.27.5.1261

Rein, J. R., & Markman, A. B. (2010). Assessing the concreteness of relational representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1452–1465. http://dx.doi.org/10.1037/a0021040

Resnick, I., & Shipley, T. F. (2013). Breaking new ground in the mind: An initial study of mental brittle transformation and mental rigid rotation in science experts. *Cognitive Processing, 14,* 143–152. http://dx.doi.org/10.1007/s10339-013-0548-2

Resnick, L. B. (2010). Nested learning systems for the thinking curriculum. *Educational Researcher, 39,* 183–197. http://dx.doi.org/10.3102/0013189X10364671

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94,* 249–273. http://dx.doi.org/10.1016/j.jecp.2006.02.002

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science, 6,* 177–192. http://dx.doi.org/10.1002/wcs.1336

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Mathematics. Cognitive supports for analogies in the mathematics classroom. *Science, 316,* 1128–1129. http://dx.doi.org/10.1126/science.1142103

Rikers, R. M., Schmidt, H. G., & Boshuizen, H. P. (2002). On the constraints of encapsulated knowledge: Clinical case representations by

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

medical experts and subexperts. *Cognition and Instruction, 20,* 27–45. http://dx.doi.org/10.1207/S1532690XCI2001_2

Rittle-Johnson, B., & Star, J. R. (2011). The power of comparison in learning and instruction: Learning outcomes supported by different types of comparisons. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 199–225). Waltham, MA: Elsevier. http://dx.doi.org/10.1016/B978-0-12-387691-1.00007-7

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101,* 836–852. http://dx.doi.org/10.1037/a0016026

Roediger, H. L., III, Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. *Psychology of Learning and Motivation-Advances in Research and Theory, 55,* 1–36. http://dx.doi.org/10.1016/B978-0-12-387691-1.00001-6

Rohrer, D. (2012). Interleaving helps students distinguish among similar concepts. *Educational Psychology Review, 24,* 355–367.

Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review, 21,* 1323–1330. http://dx.doi.org/10.3758/s13423-014-0588-3

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39,* 406–412. http://dx.doi.org/10.3102/0013189X10374770

Rohrer, D., Taylor, K., & Sholar, B. (2010). Tests enhance the transfer of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 233–239. http://dx.doi.org/10.1037/a0017678

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605. http://dx.doi.org/10.1016/0010-0285(75)90024-9

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439. http://dx.doi.org/10.1016/0010-0285(76)90013-X

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology, 16,* 371–416. http://dx.doi.org/10.1016/0010-0285(84)90014-8

Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 629–639. http://dx.doi.org/10.1037/0278-7393.13.4.629

Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 456–468. http://dx.doi.org/10.1037/0278-7393.15.3.456

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. http://dx.doi.org/10.1037/0278-7393.16.1.42

Rottman, B. M., Gentner, D., & Goldwater, M. B. (2012). Causal systems categories: Differences in novice and expert categorization of causal phenomena. *Cognitive Science, 36,* 919–932. http://dx.doi.org/10.1111/j.1551-6709.2012.01253.x

Rütsche, B., & Schalk, L. (in preparation). Constructing complex relational categories: The influence of prior knowledge.

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36,* 1019–1050. http://dx.doi.org/10.1111/j.1551-6709.2012.01250.x

Sagi, E., Kaufmann, S., & Clark, B. (2009). Semantic density analysis: Comparing word meaning across time and phonetic space. In R. Basili & M. Pennacchiotti, *Proceedings of the EACL 2009 Workshop on GEMS: Geometrical Models of Natural Language Semantics* (pp. 104–111). Athens, Greece.

Sakamoto, Y., & Love, B. C. (2010). Learning and retention through predictive inference and classification. *Journal of Experimental Psychology: Applied, 16,* 361–377. http://dx.doi.org/10.1037/a0021610

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in grades 5 and 6. *Journal of Educational Psychology, 101,* 359–372. http://dx.doi.org/10.1037/a0013840

Schneider, M., & Hardy, I. (2013). Profiles of inconsistent knowledge in children's pathways of conceptual change. *Developmental Psychology, 49,* 1639–1649. http://dx.doi.org/10.1037/a0030976

Schneider, W., Körkel, J., & Weinert, F. E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology, 81,* 306–312. http://dx.doi.org/10.1037/0022-0663.81.3.306

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16,* 475–522. http://dx.doi.org/10.1207/s1532690xci1604_4

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1–51). Norwich, CT: Information Age.

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. http://dx.doi.org/10.1037/a0025140

Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences, 21,* 1–17. http://dx.doi.org/10.1017/S0140525X98000107

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., & Van Hoesen, G. W. (2001). Prefrontal cortex in humans and apes: A comparative study of area 10. *American Journal of Physical Anthropology, 114,* 224–241. http://dx.doi.org/10.1002/1096-8644(200103)114:3<224::AID-AJPA1022>3.0.CO;2-I

Sewell, D. K., & Lewandowsky, S. (2011). Restructuring partitioned knowledge: The role of recoordination in category learning. *Cognitive Psychology, 62,* 81–122. http://dx.doi.org/10.1016/j.cogpsych.2010.09.003

Sewell, D. K., & Lewandowsky, S. (2012). Attention and working memory capacity: Insights from blocking, highlighting, and knowledge restructuring. *Journal of Experimental Psychology: General, 141,* 444–469. http://dx.doi.org/10.1037/a0026560

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs, 75* (13, Whole No. 517).

Shipley, T., Tikoff, B., Manduca, C., & Ormand, C. J. (2013). Structural geology practice and learning, from the perspective of cognitive science. *Journal of Structural Geology, 54,* 72–84. http://dx.doi.org/10.1016/j.jsg.2013.07.005

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124,* 209–215. http://dx.doi.org/10.1016/j.cognition.2012.04.005

Simms, N. K. (2013). *Overcoming the object bias in analogical development: Encoding time, allocation of attention, and relational knowledge* (Doctoral dissertation, Northwestern University).

Singley, M. K., & Anderson, J. R. (1989). *Transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Sloman, S. A., Love, B. C., & Ahn, W. K. (1998). Feature centrality and conceptual coherence. *Cognitive Science, 22,* 189–228. http://dx.doi.org/10.1207/s15516709cog2202_2

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241. http://dx.doi.org/10.1037/h0036351

Smith, J. P., III, diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *Jour-*

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

nal of the Learning Sciences, 3, 115–163. http://dx.doi.org/10.1207/s15327809jls0302_1

Spellman, B. A., & Holyoak, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology, 31,* 307–346. http://dx.doi.org/10.1006/cogp.1996.0019

Stains, M., & Talanquer, V. (2008). Classification of chemical reactions: Stages of expertise. *Journal of Research in Science Teaching, 45,* 771–793. http://dx.doi.org/10.1002/tea.20221

Stern, E., & Schneider, M. (2010). A digital road map analogy of the relationship between neuroscience and educational research. *ZDM - International Journal on Mathematics Education, 42,* 511–514.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22,* 123–138. http://dx.doi.org/10.1007/s10648-010-9128-5

Sweller, J. (2011). Cognitive load theory. *The Psychology of Learning and Motivation: Cognition in Education, 55,* 37–76. http://dx.doi.org/10.1016/B978-0-12-387691-1.00002-8

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2,* 59–89. http://dx.doi.org/10.1207/s1532690xci0201_3

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296. http://dx.doi.org/10.1023/A:1022193728205

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120,* 439–471. http://dx.doi.org/10.1037/a0033138

Taconis, R., Ferguson-Hessler, M. G. M., & Broekkamp, H. (2001). Teaching science problem solving: An overview of experimental work. *Journal of Research in Science Teaching, 38,* 442–468. http://dx.doi.org/10.1002/tea.1013

Talanquer, V. (2006). Common sense chemistry: A model for understanding students alternative conceptions. *Journal of Chemical Education, 83,* 811–816. http://dx.doi.org/10.1021/ed083p811

Taylor, E. G., & Ahn, W. K. (2012). Causal imprinting in causal structure learning. *Cognitive Psychology, 65,* 381–413. http://dx.doi.org/10.1016/j.cogpsych.2012.07.001

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331,* 1279–1285. http://dx.doi.org/10.1126/science.1192788

Thagard, P. (1997). *Creative thought: An investigation of conceptual structures and processes*. Washington, DC: American Psychological Association.

Thagard, P. (2007). Coherence, truth, and the development of scientific knowledge. *Philosophy of Science, 74,* 28–47. http://dx.doi.org/10.1086/520941

Tomlinson, M. T., & Love, B. C. (2006). From pigeons to humans: Grounding relational learning in concrete examples. In Y. Gil & R. J. Mooney (Eds.), *Proceedings of the Twenty-first National Conference on Artificial Intelligence* (pp. 199–204). AAAI Press.

Tomlinson, M. T., & Love, B. C. (2010). When learning to classify by relations is easier than by features. *Thinking & Reasoning, 16,* 372–401. http://dx.doi.org/10.1080/13546783.2010.530464

Tran, R., Rohrer, D., & Pashler, H. (2015). Retrieval practice: The lack of transfer to deductive inferences. *Psychonomic Bulletin & Review, 22,* 135–140.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352. http://dx.doi.org/10.1037/0033-295X.84.4.327

Tytler, R. (1998). Children's conceptions of air pressure: Exploring the nature of conceptual change. *International Journal of Science Education, 20,* 929–958. http://dx.doi.org/10.1080/0950069980200803

Vendetti, M. S., Matlen, B. J., Richland, L. E., & Bunge, S. A. (2015). Analogical reasoning in the classroom: Insights from cognitive science. *Mind, Brain, and Education, 9,* 100–106. http://dx.doi.org/10.1111/mbe.12080

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4,* 45–69. http://dx.doi.org/10.1016/0959-4752(94)90018-3

Vosniadou, S. (Ed.), (2010). *International handbook of research on conceptual change*. Mahwah, NY: Erlbaum.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24,* 535–585. http://dx.doi.org/10.1016/0010-0285(92)90018-W

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3–34). Mahwah, NJ: Erlbaum.

Wagner, J. F. (2006). Transfer in pieces. *Cognition and Instruction, 24,* 1–71. http://dx.doi.org/10.1207/s1532690xci2401_1

Walker, C. H. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition and Instruction, 4,* 25–42. http://dx.doi.org/10.1207/s1532690xci0401_2

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition, 28,* 1205–1212. http://dx.doi.org/10.3758/BF03211821

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18,* 158–194. http://dx.doi.org/10.1016/0010-0285(86)90011-3

Weinert, F. E., & Schneider, W. (Eds.). (1999). *Individual development from 3–12: Findings from the Munich Longitudinal Study*. Cambridge, United Kingdom: Cambridge University Press.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology, 8,* 3–19. http://dx.doi.org/10.1023/A:1009421303064

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science, 18,* 221–281. http://dx.doi.org/10.1207/s15516709cog1802_2

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8,* 247–261. http://dx.doi.org/10.1037/h0074898

Ziegler, E., & Stern, E. (2014). Delayed benefits of learning elementary algebraic transformations through contrasted comparisons. *Learning and Instruction, 33,* 131–146. http://dx.doi.org/10.1016/j.learninstruc.2014.04.006

Zimmerman, B. J. (2000). Self-efficacy: An essential move to learn. *Contemporary Educational Psychology, 25,* 82–91. http://dx.doi.org/10.1006/ceps.1999.1016

Received November 18, 2014
Revision received December 7, 2015
Accepted December 30, 2015 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.