Accounting for the Relational Shift and Context Sensitivity in the Development of Generalization

Paul H. Thibodeau (paul.thibodeau@oberlin.edu)
Erin M. Tesny (erin.tesny@oberlin.edu)
Oberlin College; Department of Psychology

Stephen J. Flusberg (stephen.flusberg@purchase.edu)
Purchase College, State University of New York; Department of Psychology

Abstract

Similarity-based generalization is fundamental to human cognition, and the ability to draw analogies based on relational similarities between superficially different domains is crucial for reasoning and inference. Learning to base generalization on shared relations rather than (or in the face of) shared perceptual features has been identified as an important developmental milestone. However, recent research has shown that children and adults can flexibly generalize based on perceptual or relational similarity, depending on what has been an effective strategy in the past in a given context. Here we demonstrate that this pattern of behavior naturally emerges over the course of development in a domain-general statistical learning model that employs distributed, sub-symbolic representations. We suggest that this model offers a parsimonious account of the development of context-sensitive, similarity-based generalization and may provide several key advantages over other popular structured or symbolic approaches to modeling analogical inference.

Keywords: Analogy; similarity; relational shift; distributed connectionist model; generalization; statistical learning

Introduction

Is a lemon more similar to a small yellow balloon or a green grape? The answer, it turns out, is not so straightforward. All three objects are small and round(ish), but the lemon and balloon are somewhat larger than the grape and both of them are yellow. On the other hand, the lemon and grape are filled with juice, grow on trees, and belong to the same basic category (fruit), while the balloon is man-made and filled with air. Your response, therefore, may depend on what type of similarity (you believe) the questioner has in mind; the lemon *looks* more similar to the yellow balloon but is *structurally* (and functionally) more similar to the grape.

Without any additional information, most adults would probably say that the lemon is more similar to the grape. The shared taxonomy and structural elements of the lemon and grape trump the superficial similarity of the lemon and balloon. However, this *relational match* requires relatively sophisticated knowledge of lemons and grapes; without it, the lemon will seem more similar to the balloon.

Indeed, experimental research has found that young children typically base similarity judgments on perceptual features before they have the relevant domain knowledge to make relational matches (Gentner & Ratterman, 1998). In other words, until young children gain sufficient knowledge of fruit, they are likely to say that a lemon is more similar to a yellow balloon than a grape. This developmental change in similarity matching – from an early reliance on surface-level, perceptual features to a later reliance on structural or relational properties – is known as the perceptual-to-relational shift (Gentner, 1988; Goswami, 1996; Piaget, 1952; Ratterman & Gentner, 1998).

Computational models have been instrumental in helping us understand the mechanistic underpinnings of relational reasoning, though they have focused primarily on adult-level competence (e.g., Falkenhainer, Forbus, & Gentner, 1989; Hummel & Holyoak, 1997). Recently, however, more attention has been given to the *development* of relational reasoning (Doumas, Hummel, & Sandhofer, 2008; Gentner, Rattermann, Markman, & Kotovsky, 1995; Leech, Mareschal, & Cooper, 2008; Morrison, Doumas, Richland, 2011; Thibodeau, Flusberg, Glick, & Sternberg, 2013). Notably, proponents of two modeling approaches that have been at the forefront of the field (SME, proposed by Falkenhainer, Forbus, & Gentner, 1989; and LISA, proposed by Hummel & Holyoak, 1997) have offered somewhat different (though arguably complementary) accounts of the emergence of relational reasoning. These two accounts highlight different aspects of cognitive development to explain the developmental trajectory of similarity-based generalization.

Gentner et al. (1995) used SME to show how conceptual change and knowledge accretion could give rise to the relational shift. That is, they argue that relational reasoning emerges as domain-specific knowledge increases (Gentner & Rattermann, 1991; Gentner, 1988; but see, e.g., Goswami, 1995 for a different perspective). In SME, concepts are hand-coded in a predicate calculus that represents both objects and their relations in a structured, symbolic fashion. Knowledge accretion is achieved in the model by manually re-coding representations (and not, e.g. through experiential learning). While this model can accurately capture the perceptual-to-relational shift in this fashion (i.e., by using "object-centered" representations to model the performance of younger children and "relation-centered" representations to model the performance of older children and adults), it leaves open the question of how conceptual re-representation emerges as people acquire domain knowledge through everyday experience (for an extended discussion of related issues see Thibodeau et al., 2013).

Morrison et al. (2011) used LISA to show how the development of inhibitory control mechanisms could support a shift in attention from perceptual to relational structure during generalization. On this account, the development of flexible cognitive control resources is crucial for being able to inhibit the allure of a superficial perceptual match. Importantly, and in contrast to SME, the basic principles of LISA have been extended in an attempt to explain how explicitly structured conceptual representations might be learned from experience (Doumas et al., 2008; although see Thibodeau et al., 2013 and Leech et al., 2008 for concerns with this approach).

There are clear advantages to both of these modeling approaches, especially since SME and LISA have been used to simulate such a wide range of findings relating to knowledge representation and reasoning (Gentner & Forbus, 2011; Hummel & Holyoak, 2005). Using these models to explain the developmental trajectory of relational reasoning, therefore, represents a parsimonious extension of each approach that helps explain several key pieces of data.

However, recent research has called into question the idea that similarity-based generalization follows a universal, across-the-board, perceptual-to-relational shift (Bulloch & Opfer, 2009; Opfer & Bulloch, 2007). According to the *predictive validity* view, children do not *necessarily* proceed from generalization by perceptual features to generalization by relational structure. Instead, they generalize flexibly over different types of similarity depending on the context of their judgment (Bulloch & Opfer, 2009; Opfer & Bulloch, 2007). In certain domains, children (and adults) will have *learned* that inferences based on relational similarity are more reliably predictive of success, while in other domains inferences based on perceptual similarity may actually be more successful.

Data supporting the predictive validity view come from studies in which children and adults are asked to make inferences about a novel object in different contexts (Bulloch & Opfer, 2009; Opfer & Bulloch, 2007). Consider the triad of insects in Figure 1. In each of the three insect triplets, there are two adults and one juvenile. The triads were designed such that the insects on the top row (the "samples": AA, *a*; BB, *b*) represent potential matches for the insects on the bottom (the "target": TT, *t*). In every case, the target juvenile looked similar to the juvenile from one of the samples (in this case both *b* and *t* are light whereas *a* is dark) and the target adults looked similar to the adults in the other sample (in this case both AA and TT are light whereas BB is dark).

Bulloch and Opfer (2009) designed two different conditions to examine whether they could influence how people would generalize about the target juvenile: one in which the relational information was relevant (the juvenile is the *offspring* of the associated adults) and another in which the relational information was irrelevant (the juvenile is the *prey* of the associated adults). They then had participants make inferences about the target juvenile, asking about category membership (is *t* the same kind as *a* or *b*?), an unobservable property (does *t* have "gogli" inside its blood similar to *a* or *b*?), and future appearance (will *t* look like *a* or *b* in the future?).

According to the predictive validity perspective, in the condition where the relation was relevant (i.e., when the participant was told that the juveniles were the offspring of the associated adults), participants should choose the sample in which the adults look like the target adults (i.e., AA). That is, they should make an inference based on relational similarity. In the context where the relation was irrelevant (i.e., when the participant was told that the juveniles were the prey of the associated adults), participants should choose the sample in which the juvenile looks like the target juvenile (i.e., *b*). That is, they should make an inference based on the perceptual similarity of the juveniles.

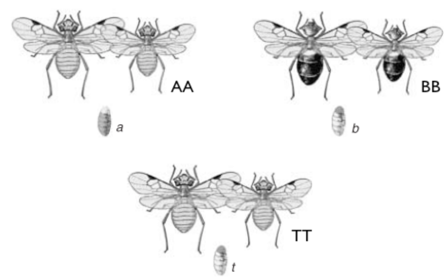

Figure 1. An example trial from Bulloch and Opfer (2009). The target juvenile (*t*) is perceptually more similar to (*b*) but is sometimes presented in a relational context that makes it more similar to (*a*).

As expected, Bulloch and Opfer (2009) found that adults based their inferences about the target juvenile on perceptual properties of the juveniles in the prey context and relational properties (i.e., the similarity of the adults) in the offspring context. Patterns of results from three-, four-, and five-year-old children, looked increasing like those of the adults. The proportion of relational matches in the offspring context increased from 61% among three-year-olds to 72% among four-year-olds and 79% among five-year-olds (adults chose the relational match 81% of the time in the offspring context). In contrast, the proportion of relational matches in the prey context decreased with age, from 56% among three-year-olds to 55% among four-year-olds and 45% among five-year-olds (adults chose the relational match 7% of the time in the prey context). This supports the view that there is not a universal trend from generalizing by perceptual features to generalizing by relational structure. Instead, these findings suggest that children and adults flexibly generalize using features or relations when contextually appropriate, based on their prior knowledge.[1]

[1] Nevertheless, we would argue that the nature of Bulloch & Opfer (2009)'s task does not provide strong evidence against the primacy of perceptual information. As the authors acknowledge, "children came to our

The Present Study

The data provided by Bulloch & Opfer complicate the traditional picture of the emergence of relational reasoning over the course of development. While popular models like SME and LISA can likely accommodate these findings, to do so might require ad-hoc changes to existing processing algorithms in order to account for the role of context and predictive validity.

Here, we present a series of artificial neural network simulations to investigate the development of context-sensitive, similarity-based generalization. The model architecture and simulated environment build on previous work that has explored the capacity of certain connectionist networks to capture and explain the development of semantic knowledge (Rogers & McClelland, 2004) and relational reasoning (e.g., Flusberg et al., 2011; Kollias & McClelland, 2013; Leech, et al., 2008; Thibodeau et al., 2013). This research has shown how and why higher-level cognitive abilities like analogical reasoning could spontaneously emerge over the course of development based on domain-general principles of statistical learning and distributed representation. The present simulations advance this work by focusing specifically on the relational shift and the mechanisms that support context-sensitive inferences.

Notably, this approach helps address some of the limitations of classical structured and symbolic models like SME and LISA (while retaining important insights from the empirical literature; e.g. the causal role that language seems to play in driving the development of relational reasoning; see Flusberg et al., 2011; Gentner & Ratterman, 1991; Thibodeau et al., 2013). In particular, our model is naturally context-sensitive (a well-known strength of connectionist networks; Rogers & McClelland, 2004) and embodies the key principles underlying the predictive validity account of similarity-based reasoning.

Methods

The environment and structure of our model was designed to replicate some of the essential features of Bulloch and Opfer's (2009) study. As input, the model takes a distributed representation of a juvenile insect and a relational context. As output, the model learns to complete the inputs with the appropriate adult, category, or property (see Table 1). That is, the model learns that a given juvenile is *born to* a pair of adults, is *eaten by* pair of adults, *will look like* a pair of adults, *is* a particular type of bug, and *has* specific properties.

Importantly, there is *coherent covariation* (Rogers & McClelland, 2004) between the *born to, will look like,* and *has* relations. Juveniles *will look like*, belong to the same category as, and *have* the same property as the adults that they are *born to*. In contrast, knowing that a given juvenile is *eaten by* a particular pair of adults does not license inferences about future appearance, category membership, or internal properties.

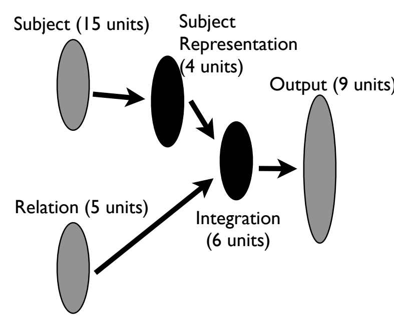

Figure 2. The network architecture for the feedforward connectionist model, an adaptation of the Rumelhart network (Rumelhart, 1990).

During training, the model learns about six juveniles in each of the five relational contexts (see Figure 1 for an illustration of the network and Table 2 for simulation parameters). These juveniles are presented to the model as distributed patterns over 15 input units. The patterns that represent the juveniles were designed to be equally different from one another, with slightly negative pairwise correlations (r = -0.2).

juv	born	eaten	look	is	has
1	$adults_1$	$adults_2$	$adults_1$	$type_1$	$prop_1$
2	$adults_1$	$adults_3$	$adults_1$	$type_1$	$prop_1$
3	$adults_2$	$adults_1$	$adults_2$	$type_2$	$prop_2$
4	$adults_2$	$adults_3$	$adults_2$	$type_2$	$prop_2$
5	$adults_3$	$adults_1$	$adults_3$	$type_3$	$prop_3$
6	$adults_3$	$adults_2$	$adults_3$	$type_3$	$prop_3$
7	$adults_2$		$adults_2$	$type_2$	$prop_2$
7		$adults_2$	$adults_1$	$type_1$	$prop_1$

Table 1. Training and Test Patterns. The top six rows represent training patterns and the bottom two represent test patterns. In training, the network learns about six juvenile bugs in each of five relational contexts for a total of 30 training patterns. At test, the model is given partial information about a novel juvenile and is asked to make inferences about the future appearance, category membership, and internal properties of that juvenile.

task knowing the value of the parent-offspring relation" (p. 120), which suggests that their participants may, at least in the offspring context, experience a perceptual-to-relational shift before they turn three.

In addition, the design of the displays does not present a clear contrast between purely perceptual and purely relational options. Notice that in Figure 1 the target juvenile (t) is a better perceptual match to the juvenile on the right (b) but the target adults (TT) are a better perceptual match to the adults on the left (AA). Since the inference questions focused on the target juvenile, it was argued that attending to the perceptual similarity of the adults represented a relational inference. However, it is unclear if children who chose the relational option did so because of the relational condition or because of the salient perceptual similarity between the sample and target adults. This latter possibility seems especially likely since there was an overall preference for the "relational" option (even five-year-olds in the prey condition chose the relational match over 45% of the time).

Further, these results offer no account for numerous other studies that find evidence of the primacy of perceptual features (e.g., Gentner, 1988; Gentner & Rattermann; Ratterman & Gentner, 1998).

To test the network's ability to generalize, it is given partial information about a novel juvenile after it has learned about the six training juveniles. The pattern that represents this "test juvenile" was designed to be perceptually similar to one pair of juveniles that the network learned about in training and relationally similar to another. Perceptual similarity is operationalized as overlap in the distributed input representations (r = 0.4 between the novel juvenile and each of the perceptually similar juveniles and r = -0.2 between each of the other juveniles). For instance, *juvenile*$_7$ might be perceptually similar to *juvenile*$_1$ and *juvenile*$_2$ (i.e., in terms of its distributed representation) but relationally similar to *juvenile*$_3$ and *juvenile*$_4$ in the sense that it might be *born to* the same adults as *juvenile*$_3$ and *juvenile*$_4$ (see the bottom two rows of Table 1).

Epochs of Training	30,000
Learning rate	0.005
Noise	0
Initial weight range	-0.1/0.1
Error measure	Cross-entropy error
Activation function	Sigmoid
Momentum	0

Table 2. Simulation parameters.

We presented the network with two kinds of inference conditions after it had finished learning about the six training juveniles. In one, the network was given the novel juvenile and information about whom that juvenile was *born to*. In the other, the network was given the novel juvenile and information about whom that juvenile was *eaten by*. In neither case was the network told what the novel juvenile *will look like*, *is*, or *has*. These were inferences that the network was asked to make.

We presented the novel information (a single pattern) to the network until it had fully learned whom the juvenile was *born to* or *eaten by* and monitored the trajectory of its inferences. Simulations were run ten times in each condition to ensure that results were not the product of an idiosyncratic result and to allow for statistical tests.

Our prediction was that the network would initially make inferences about the novel juvenile that were consistent with the perceptually similar juveniles (i.e., that the network would infer *juvenile*$_7$ will look like, is of the same type as, and has the same properties as *juvenile*$_1$ and *juvenile*$_2$). However, we expected that the network would change what it thought about the novel juvenile in the *born to* condition (i.e., to infer that *juvenile*$_7$ is actually more similar to *juvenile*$_3$ and *juvenile*$_4$ because it is also born to *adults*$_2$); we expected no such change in the *eaten by* condition. In other words, we expected the network to behave flexibly, learning to use the relational information when it was predictive (based on its own prior experiences during training) and to ignore it when it was not.

Results

As predicted, the network initially made perceptual matches in both contexts. Learning in the offspring condition, however, led to a shift in the inference patterns of the model, consistent with a perceptual-to-relational shift. Such a shift did not occur in the prey condition since there was no coherent covariation between the *eaten by* and inferential relational contexts (see Figure 3).

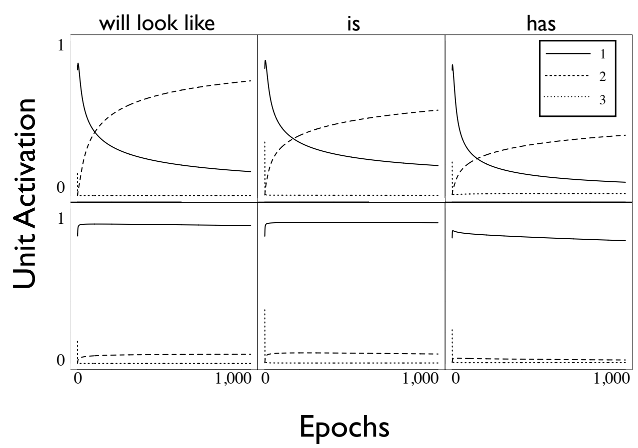

Figure 3. The network's response to the three inference questions by condition over time in one representative simulation. The top row shows results from the offspring condition while the bottom shows results from the prey condition with inferences about future appearance (left), category membership (middle), and internal property (right). In the offspring condition, there is a relational shift: initially the network infers that the novel juvenile will look like *adults*$_1$, is of *type*$_1$, and has *property*$_1$ (as indicated by the strong activation of the solid line at epoch 0); however, over time it infers that the novel juvenile will look like *adults*$_2$, is of *type*$_2$, and has *property*$_2$. In the prey condition the network makes the same initial inferences; however, importantly, it shows no relational shift.

To statistically analyze the inferential tendencies of the model, we conducted three repeated measures ANOVAs. The first contrasted pre- and post-learning in the offspring condition and found a main effect of perceptual inferences, $F[1,35]$ = 12.61, p < .01 and a strong interaction between learning and inference type, $F[1,35]$ = 74.54, p < .001. Before learning, the model was strongly biased toward making perceptual inferences. After learning, however, the network showed a dramatic shift towards relational inferences (see the first and third pairs of bars in Figure 4). That is, the model initially treated the novel juvenile like the learned, perceptually similar juveniles. But this changed when it was told that the novel juvenile was *born to* a different set of parents. Over time, it re-conceptualized this juvenile to make inferences that were consistent with the juveniles that were born to the same adults.

The second ANOVA contrasted pre- and post-learning in the prey condition and found a strong main effect of perceptual inferences, $F[1,35]$ = 60.00, p < .001 and a slight interaction between learning and inference type, $F[1,35]$ = 7.15, p < .05. As in the offspring condition, the model first made perceptual inferences. Unlike the offspring condition, we did not see a crossover after learning, although it did become slightly more likely to make a relational inference (see the first two pairs of bars in Figure 4).

Finally, the third ANOVA contrasted the post-learning inferences across the two conditions and found a significant interaction, $F[1,35] = 14.50$, $p < .001$. Whereas the network made more perceptual matches in the prey condition, it made more relational matches in the offspring condition (see the second and third pairs of bars in Figure 4).

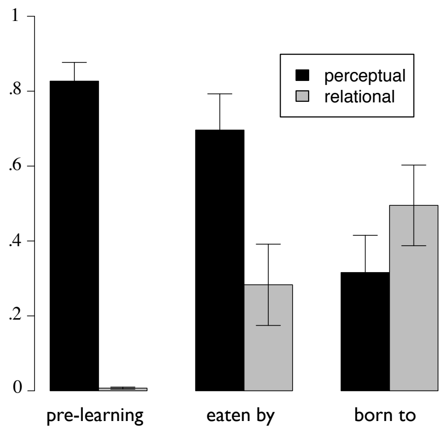

Figure 4. Average activations of units that reflect perceptual and relational inferences before learning (left), after learning in the *eaten by* condition (middle), and after learning in the *born* to condition (right). Error bars reflect standard error of the mean.

General Discussion

The results of our simulations support the view that many important phenomena in the development of similarity-based generalization can be explained by a general-purpose model of semantic learning (Rogers & McClelland, 2004). Specifically, our model captures the documented primacy of perceptual information (Gentner, 1988) and the context-flexibility of relational and perceptual generalization (Opfer & Bulloch, 2009), all without positing analogy-specific machinery or structured, symbolic representations (as with SME, Falkenhainer, Forbus, & Gentner, 1989; and LISA, Hummel & Holyoak, 1997).

On this view, the primacy of perceptual information and context-flexibility emerge naturally from learned distributed representations of objects and relations. Of note, the model provides an account of how conceptual knowledge is re-organized through experience as it acquires domain-specific knowledge (Gentner et al., 1995) and how this re-representation gives rise to relational reasoning. Importantly, it does not require the concurrent development of working memory or inhibitory control (as was the case in Morrison et al., 2011; although see Kollias & McClelland, 2013 for a fully connectionist account that considers these important cognitive mechanisms).

With this said, it is important to be clear that we are not claiming that our model can account for all facets of human analogical reasoning. Many of the tasks that SME and LISA model so well rely on processes that we purposefully did not try to simulate for the sake of theoretical and practical simplicity (e.g., Bowdle & Gentner, 1997; Morrison et al., 2004). For instance, whereas we argue that the kinds of inferences that are made in Bulloch and Opfer (2009)'s task do not require highly developed mechanisms for inhibitory control, it is very likely that other kinds of analogy tasks do (e.g., Gick & Holyoak, 1980). Further, our model does not offer an account of analogical reasoning in which highly structured information is learned and leveraged for inference very quickly (e.g., Gentner & Markman, 1995. For an extended discussion of these issues, see Thibodeau et al., 2013).

Conclusion

Similarity-based generalization is fundamental to human cognition, and the ability to draw analogies based on abstract relational connections between superficially different domains is crucial for reasoning and inference (Gentner, 1983, 2010; Hofstadter, 2001; Penn, Holyoak, & Povinelli, 2008). Learning to base generalization on shared relations rather than (or in the face of) shared perceptual features has been identified as an important developmental milestone (Piaget, 1952; Gentner, 1988; Leech et al., 2008; Ratterman & Gentner, 1998). Unlike many other approaches to analogical reasoning that use symbolic representations and analogy-specific mapping mechanisms, we have shown that context-sensitive perceptual and relational reasoning can emerge over the course of development in a domain-general learning model that employs distributed, sub-symbolic representations.

References

Bowdle, B., & Gentner, D. (1997). Informativity and asymmetry in comparisons. *Cognitive Psychology, 34*, 244–286.

Bulloch, M.J., & Opfer, J.E. (2009). What makes relational reasoning smart? Revisiting the perceptual-to-relational shift in the development of generalization. *Developmental Science, 12,* 114-122.

Doumas, L., Hummel, J., & Sandhofer, C. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115*, 1–43.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41,* 1–63.

Flusberg, S. J., Thibodeau, P. H., Sternberg, D. A., & Glick, J. J. (2010). A connectionist approach to embodied conceptual metaphor. *Frontiers in Psychology*, 1:12.

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

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child development*, pages 47–59.

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.

Gentner, D., & Forbus, K. D. (2011). Computational models of analogy. *WIREs Cognitive Science, 2,* 266–276.

Gentner, D., & Markman, A. B. (1995). Analogy-based reasoning in connectionism. In M. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 91–93). Cambridge, MA: MIT Press.

Gentner, D., & Ratterman, M.J. (1991). Language and the career of similarity. In S.A. Gelman & J.P. Byrnes (Eds.), *Perspectives on thought and lanauge: Interrelations in development* (pp. 225-277). London: Cambridge University Press.

Gentner, D., Rattermann, M. J., Markman, A. B., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263-313). Hillsdale, NJ: LEA.

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

Goswami, U. (1995). Transitive relational mappings in three and four year olds: The analogy of Goldilocks and the three bears. *Child Development, 66,* 877–892.

Goswami, U. (1996). Analogical reasoning in cognitive development. In H. Reese (Ed.), *Advances in child development and behavior* (pp. 92-135). San Diego, CA: Academic Press.

Hofstadter, D. (2001). Analogy as the core of cognition. In Gentner, D., Holyoak, K., and Kokinov, B., editors, *The analogical mind: Perspectives from cognitive science*, pages 499–538. MIT Press: Cambridge, MA.

Hummel, J. E. (2010). Symbolic versus associative learning. *Cognitive Science, 34,* 958–965.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104,* 427–466.

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture. *Current Directions in Psychological Science, 14,* 153–157.

Keil, F.C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Kollias, P. & McClelland, J. L. (2013). Context, cortex, and associations: A connectionist developmental approach to verbal analogies. *Frontiers in Psychology, 4,* 857.

Leech, R., Mareschal, D., & Cooper, R. (2008). Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. *Behavioral and Brain Sciences*, *31,* 357–378.

Morrison, R.G., Doumas, L.A.A., & Richland, L.E. (2011). A computational account of children's analogical reasoning: blanacing inhibitory control in working memory and relational representation. *Developmental Science, 14,* 516-529.

Morrison, R., Krawczyk, D., Holyoak, K., Hummel, J., Chow, T., Miller, B., & Knowlton, B. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *Journal of Cognitive Neuroscience, 16,* 260–271.

Opfer, J. E., & Bulloch, M. J. (2007). Causal relations drive young children's induction, naming, and categorization. *Cognition*, *105,* 207-217.

Penn, D., Holyoak, K., & Povinelli, D. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, *31,* 109–130.

Piaget, J. (1952). *The child's concept of number*. New York: Norton.

Rattermann, M. & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, *13,* 453–478.

Rogers, T. T. and McClelland, J. L. (2004). *Semantic Cognition*. MIT Press, Cambridge, MA.

Thibodeau, P.H., Flusberg, S.J., Glick, J.J., & Sternberg, D.A. (2013). An emergent approach to analogical inference. *Connection* Science, *25,* 27-53.