Differences in the Development of Analogy Across Cultures: A Computational Account

Leonidas A.A. Doumas (leonidas@hawaii.edu)
University of Hawaii at Manoa, Department of Psychology
2430 Campus Rd. Honolulu, HI 96822

Robert G. Morrison (rmorrison@luc.edu)
Loyola University Chicago, Psychology Department,
6525 North Sheridan Road Chicago, IL 60626 USA

Lindsey E. Richland (lerich@uci.edu)
University of California, Irvine, Department of Education
2001 Berkeley Place, Irvine, CA 92697-5500

Abstract

Theories of the development of analogical reasoning emphasize either the centrality of relational knowledge accretion or changes in information processing. Recent cross-cultural data collected from children in the United States and China (Richland, Chan, Morrison, & Au, 2010) provides a unique way to test these theories. Here we use simulations in LISA/DORA (Doumas, Hummel, & Sandhofer, 2008; Hummel & Holyoak, 1997, 2003), a neurally-plausible computer model of relational learning and analogical reasoning, to argue that the development of analogical reasoning in children may best be conceptualized as an equilibrium between knowledge accretion and progressive improvement in information processing capability. Thus, improvements in inhibitory control in working memory as children mature enable them to process more relationally complex analogies. At the same time, however, children produce more complex and more accurate analogies in domains in which they have learned richer and more refined representations of relational concepts.

Relational thinking—i.e., thinking based on the roles that objects play rather than the literal features of those things—is a cornerstone of human cognition. It underlies, among many other things, our ability to make analogies, or to appreciate correspondences between domains (e.g., Holyoak & Thagard, 1995).

As with many cognitive processes, our ability to make analogies changes with development. While there is considerable agreement that analogy is a very important process in cognitive development (e.g., Gentner, 2003), there is considerable disagreement as to how the ability to reason analogically develops.

Theories of the Development of Analogical Reasoning

Three primary hypotheses have been put forward to explain age-related differences in analogical reasoning: changes in domain knowledge, a relational shift from object similarity to relational similarity, and increased processing or working memory (WM) capacity.

Goswami and colleagues (Goswami, 1992, 2001; Goswami & Brown, 1989) proposed that the ability to make analogies is present even in early infancy. However, children can only evidence this ability with age and increased knowledge. In other words, the change in children's ability to make analogies is not a function of a developing mechanism, but rather knowledge accretion.

Alternately, Gentner and Rattermann (1991; Rattermann & Gentner, 1998) argued that a domain-specific "relational shift" is responsible for changes in children's analogical abilities. Gentner and Rattermann suggest that as children build knowledge in a particular domain they progress from reasoning about that domain in terms of the perceptual features of objects, to the relations between those objects. For example, 3 year-old children will categorize objects based on overall featural similarity (e.g., they will match apples to red balls rather than bananas), however by age 4 or 5, children will categorize objects based on relational similarity (e.g., matching apples to bananas even in the presence of featural distracters like red balls; Gentner & Namy, 1999). The ability to make analogies based on relational commonalities between domains, therefore, progresses on a domain-by-domain basis with more complex analogies produced in domains in which knowledge is richer.

In contrast to accounts of analogy development based on increases in knowledge, the relational complexity hypothesis of Halford (1993; Andrews & Halford, 2002; Halford et al., 2002) holds that limits in children's WM capacity affects their ability to process relations simultaneously, and therefore their ability to make analogies. According to Halford and colleagues, children can process only specific levels of relational complexity, defined as the number of sources of variation that are related and must be processed together. The simplest level of relational complexity is a binary relation, where only two arguments are sources of variation. The relation, *chase* (dog, cat), for instance, specifies a single relation (*chase*) between two objects (dog, cat). To reason about this relation, a one must keep only the two objects and their relation in mind. A ternary relation (e.g.,

gives-to (boy, girl, book) is more complex, requiring a reasoner to consider three objects and their respective roles. The more complex the relation, the more WM resources are required to process it. As children mature, neural developments leading to increased WM capacities (see, e.g., Diamond, 2002) allow processing situations with greater relational complexity, and, by extension, children are capable of drawing richer and more complex analogies.

Likewise, Richland, Morrison, and Holyoak (2006) proposed that inhibitory control might help to explain the relationship between maturation and the impact of relational complexity on analogical reasoning in young children. While inhibitory control has been a major topic in models of cognitive development (Bjorklund & Harnishbeger, 1990; Diamond, 2002) it has not previously been applied to understanding the development of analogy; however, the hypothesis that inhibitory control is important for the development of analogy is consistent with results from other cognitive tasks (e.g., Diamond, Kirkham & Amso, 2002; Lorsbach & Reimer, 1997; Zelazo et al., 2003).

Multiple Sources in Analogical Development

Richland, Morrison and Holyoak (2006) developed a set of scene analogy problems to investigate relational complexity and featural distraction within a single analogical reasoning task. They found that children from age 3 to 14 steadily improved in their ability to solve more relationally complex problems and resist distraction.

In a follow-up study Richland, Chan, Morrison, and Au (2010) used these same problems with Cantonese speaking 3-4 year old children from Hong Kong. While US children of this age showed main effects of both relational complexity and featural distraction, Chinese children only showed an effect of featural distraction (see Figure 5).

There are several reasons to believe that the Chinese children would score differently on analogical reasoning problems than U.S. children based on their knowledge base and experience with reasoning about relations. Adult studies have shown cultural differences in normative patterns for drawing relational inferences (see Nisbett 2003) such that Chinese and Japanese reasoners may be more attuned to relational correspondences than U.S. participants. These differences also appear cross-culturally in children's socialization and linguistic routines. For example, Asian caregivers use more action oriented language and referential verbs than relatively object-focused U.S. caregivers (e.g., Chinese: (Mandarin) Tardif, 1996; Tardif, Gelman & Xi, 1999; Tardif, Shatz, Naigles, 1997; (Cantonese) Leung, 1998). Chinese children themselves may additionally show a higher relative rate of verb usage in Mandarin (Tardif, 1996; 2006; Tardif, Shatz, & Naigles, 1997; Tardif, Gelman, & Xu, 1999) than U.S. children of comparable ages who show a more pronounced noun bias. In contrast, there is no theoretical reason to expect differences in information processing capacity between the US and Hong Kong (Hedden, et al., 2000).

Accordingly, Richland et al. (2010) reasoned that the US and HK 3-4 year old children each had decreased inhibitory control relative to older children resulting in their distractibility, but that HK children had more sophisticated relational representations than US children resulting in their superior ability to solve more relationally complex problems.

A Computational Account of the Multiple-Source Theory of Analogical Development

Previous Work

Traditionally, researchers have attempted to model the effects of knowledge accretion and increased working memory capacity on analogical development in isolation. For example, Gentner and colleages (e.g., Gentner et al., 1995) used SME (Falkenhainer, Forbus, & Gentner, 1989) to model the relational shift data of Gentner and Rattermann (1991). Gentner et al. captured the differences in analogical reasoning in 4 and 5 year-old children by providing the model with more relational representations at age 5 than at age 4. That is, with limited knowledge of relations, the model behaved like the younger children in Gentner and Rattermann's experiments, making analogies based on over-all perceptual similarity. However, with increased relational knowledge, the model behaved more like the older children, making analogies based on shared relations. Importantly the representations provided to the model had to be hand-coded by the modeler.

More recently, Morrison, Doumas, and Richland (2006), used the LISA model (Hummel & Holyoak, 1997, 2003) in an attempt explain changes in children's analogy making in terms of changes in capacity limits. LISA is a model of analogy-making that relies on time as a signal to bind distributed (i.e., connectionist) representations of objects and relational roles into structured (i.e., symbolic) representations. LISA is powerful, in part, because it benefits from both the flexibility of connectionist approaches and the structure-sensitivity of symbolic approaches (an important property for demonstrating human like relational reasoning; see, e.g., Doumas & Hummel, 2005; Holyoak & Hummel, 2000; Penn, Holyoak, & Povinelli, 2008). In addition, as a consequence of using time to carry binding information, LISA suffers from capacity limitations that mirror those of human WM (Hummel & Holyoak, 2003; Morrison, Doumas, & Richland, 2006; Morrison et al., 2005). LISA relies on lateral inhibition between units to establish the temporal patterns that carry binding information. By decreasing lateral inhibition, LISA's WM is effectively reduced. Morrison et al. (2006), used this property of to capture the pattern of results from Richland et al. (2006).

Approaches using SME and LISA both suffer from limitations, though. First, each approach is based on a

single explanatory variable. As a result, the knowledge accretion approach seems insufficient to explain the results of the Scene Analogy task (see Richland et al., 2006), while the simply changing capacity limits cannot explain the cross-cultural findings of Richland et al. (2010). In addition, both approaches rely on hand-coded relational representations that must be added by the modeler. Neither model makes any claims where these representations, which both models require in order to reason relationally—and that provide the explanatory mechanism in the knowledge accretion case—come from in the first place.

Doumas, Hummel, and Sandhofer (2008) have developed an extension of the LISA model, called DORA (*Discovery of Relations by Analogy*) that learns structured representations of relations from unstructured (i.e., flat feature vector) representations of object properties. DORA provides a means by which the representations used by LISA are learned from examples, and, consequently, provides an opportunity to understand the interplay between the dual sources of knowledge accretion and increasing capacity limits as effectors of the changes in children's analogy making.

The LISA/DORA model

LISA (Hummel & Holyoak, 1997, 2003) is a symbolic-connectionist model of analogy and relational reasoning. DORA (Doumas et al., 2008) is an extension of LISA that learns structured (i.e., symbolic) representations of relations from unstructured inputs. That is, DORA provides an account of how the structured relational representations LISA uses to perform relational reasoning can be learned from examples.

DORA accounts for over 20 phenomena from the literature on relational learning, as well as its development (e.g., Doumas & Hummel, 2010; Doumas et al., 2008). In addition, as DORA learns relational representations, it comes to take LISA as a special case, and can simulate the additional 30+ phenomena in relational thinking simulated by LISA. The description of LISA/DORA that follows is a brief overview due to space constraints. For full details of the models and their operations see Doumas et al. (2008) and Hummel and Holyoak (1997, 2003).

LISAese Representations In LISA (and DORA *after* it has gone through learning) relational structures are represented by a hierarchy of distributed and localist codes (see Figure 1). At the bottom, "semantic" units represent the features of objects and roles in a distributed fashion. At the next level, these distributed representations are connected to localist units (POs) representing individual predicates (or roles) and objects. Localist role-binding units (RBs) link object and predicate units into role-filler binding pairs. At the top of the hierarchy, localist P units link RBs into whole relational propositions.

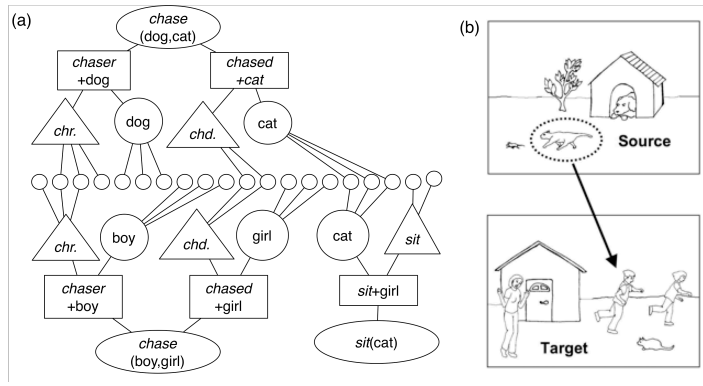

Figure 1. (a) Representation of a LISA/DORA representation like that used to simulate a Scene analogy problem like that in (b). The P (oval), RB (rectangle), and predicate (triangle) units were learned during Simulation Part One. Objects (circles) desribed the objects involved in the Scene Analogy problem. (b) Example of a scene analogy problem from Richland et al., 2006.

Propositions are divided into two mutually exclusive sets: a driver and one or more recipients. In LISA/DORA, the sequence of firing events is controlled by the driver. Specifically, one (or at most three) proposition(s) in the driver becomes active (i.e., enter working memory). When a proposition in the driver becomes active, role-filler bindings must be represented dynamically on the units that maintain role-filler independence (i.e., POs and semantic units; see Hummel & Holyoak, 1997). In LISA binding information is carried by synchrony of firing (with roles firing simultaneously with their fillers). In DORA, binding information is carried by systematic asynchrony of firing, with bound role-filler pairs firing in direct sequence (see Doumas et al., 2008 for details).[1] Activation flows from the driver units to their semantic units. Units in the driver and recipient share the same pool of semantic units. Thus, units in the recipient become active in response to the pattern of activation imposed on the semantic units by the active driver proposition.

Relational Learning Very simply, DORA uses comparison to isolate shared properties of objects and to represent them as explicit structures. DORA starts with simple feature-vector representations of objects (i.e., a node connected to set of features describing that object). When DORA compares one object to another, corresponding features of the two representations fire simultaneously. Any semantic features common to both objects receive twice as much input and thus become roughly twice as active as features connected to one but not the other. By recruiting a new PO unit and learning connections between that unit and active semantics via Hebbian learning (wherein the strength of connections is a function of the units' activation), DORA learns stronger connections between the new PO unit and more active

[1] Asynchrony-based binding allows role and filler to be coded by the same pool of semantic units, which allows DORA to learn representations of relations from representations of objects (Doumas et al., 2008).

semantic units. The new PO thus becomes an explicit representation of the featural overlap of the compared objects. Applied iteratively this process results in explicit and structured representations of object properties and relational roles (see Doumas et al., 2008). Comparison also allows DORA to learn representations of multi-place relations by linking sets of constituent role-filler pairs into relational structures (i.e., to learn the *chases* relation by linking together representations of the roles *chaser* and *chased*; see Doumas et al., 2008 for details).

Mapping For the purposes of analogical mapping, LISA/DORA learns *mapping connections* between units of the same type (e.g., PO, RB, etc.) in the driver and recipient (e.g., between PO units in the driver and PO units in the recipient). These connections grow whenever corresponding units in the driver and recipient are active simultaneously. They permit LISA to learn the correspondences (i.e., mappings) between corresponding structures in separate analogs. They also permit correspondences learned early in mapping to influence the correspondences learned later.

Simulations

Methods

We tested the hypothesis that differences in performance between U.S. and Chinese children were due to differences in relational knowledge. Specifically, we hypothesized that the relational representations of children from Hong Kong were more developed than those of children from the U.S. We used LISA/DORA to test this hypothesis by simulating the results of Richland et al. (2010). The simulation consisted of two complementary parts. In the first part we used DORA to develop representations of relational concepts from examples. We simulated the difference in U.S. and Chinese children by allowing DORA increased learning trials in order to simulate the Chinese children, reflecting the assumption that the experience of children in Hong Kong differs from children in the U.S. We then used the representations that DORA had learned during the first part of the simulation to simulate the Richland et al. (2010) task.

Simulation Part One We used DORA's relational learning algorithm (see Doumas et al., 2008 for details) to develop relational representations from unstructured examples. We started DORA with representations of 100 objects attached to random sets of features (chosen from a pool of 100). We then defined 4 relations (*chase*, *reach-for*, *angry-with*, and *hang-from*). Each relation consisted of two roles, each with three semantic features (e.g., for the *chase* relation, both the roles *chaser* and *chased* were each defined by three specific semantic units). Each of the 100 objects was attached to the features of between 1 and 3 relational roles chosen at random. For example, object1 might be attached to the features for *chaser* (one role of *chases*) and *reaching* (one role of *reach-for*). On each iteration we presented DORA with sets of objects from similar relations, and allowed it to compare the objects and learn from the results (as per DORA's relation learning algorithm). As DORA learned new representations it would also use these representations to make subsequent comparisons. For instance, if DORA learned an explicit representation of the property *chases* (*x*, *y*) by comparing sets objects attached to the roles of *chase* (i.e., *chaser* and *chased*), it could use this new representation for future comparisons. On each trial we selected between 2 and 4 representations and let DORA compare them and learn from the results (i.e., perform predication, and relation learning routines).

We ran 25 sessions each consisting of 800 learning trials During each session, the inhibition parameter was set to a value sampled from a random distribution with a mean of 0.7, and a standard distribution of 0.1. The value of the parameter reflected the reduced WM capacity evidenced in young children (see Morrison et al., 2006).. We measured the quality of the representations DORA had learned during the last 100 trials after each 100 trials. Quality was calculated as the mean of connection weights to relevant features (i.e., those defining a specific transformation or role of a transformation) divided by the mean of all other connection weights + 1 (1 was added to the mean of all other connection weights to normalize the quality measure to between 0 and 1). A higher quality denoted stronger connections to the semantics defining a specific relation relative to all other connections (i.e., a more pristine representation of the relation). Figure 2 shows the quality of the representations DORA learned for each set of 100 comparisons from 100 to 800. As expected, the quality of the representations DORA learns increase as a function of experience (see Doumas et al., 2008 for more details)

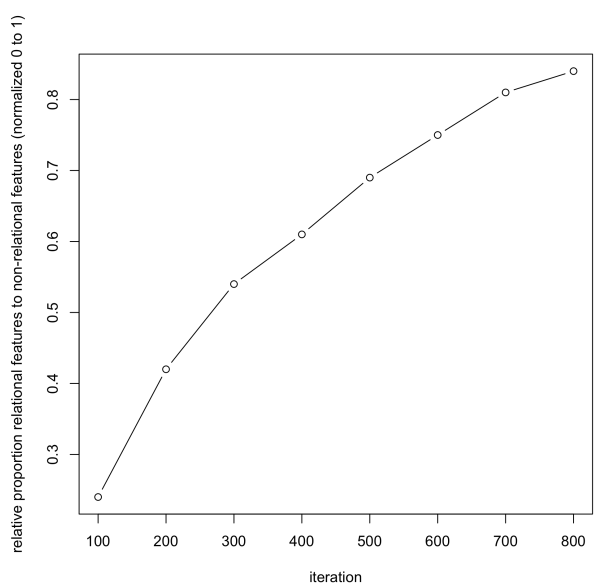

Figure 2. Quality of the representations DORA learned during Simulation Part One.

Simulation Part Two To model the Scene Analogy Problems we used representations of the four problem types (1-relation, no distracter; 1-relation, distracter; 2-relation, no distracter; 2-relation, distracter) composed from the representations DORA had learned during Simulation Part One. For example, to represent the problem from Figure 1, we used a representation of the *chase* relation DORA had learned during Simulation Part One (relational role, RB, and P units) along with object units (e.g, boy and girl) composed of 5 semantic features describing that object (see Figure 1). For 2-Relation problems both relations were represented in LISA's WM together (Hummel & Holyoak, 1997). Vitally, we simulated children from the U.S. by using the representations DORA had learned after only 400 comparisons, and those of the children from Hong Kong using the representations DORA had learned after 600 comparisons.

The laternal inhibition parameter was set exactly as in Simulation Part One. Each simulation run consisted of firing three phase sets in LISA/DORA's working memory, "randomly" assigned by LISA/DORA and allowing LISA/DORA to try to map the representation in the driver to the representation in the recipient. When LISA/DORA failed to determine a stable mapping after firing three phase sets, an answer was selected at random.

Results

The simulation results along with the experimental results from Richland et al. (2010) are presented in Figure 3. LISA/DORA's performance mirrored experimental results for the age groups from both the U.S. and China across conditions.

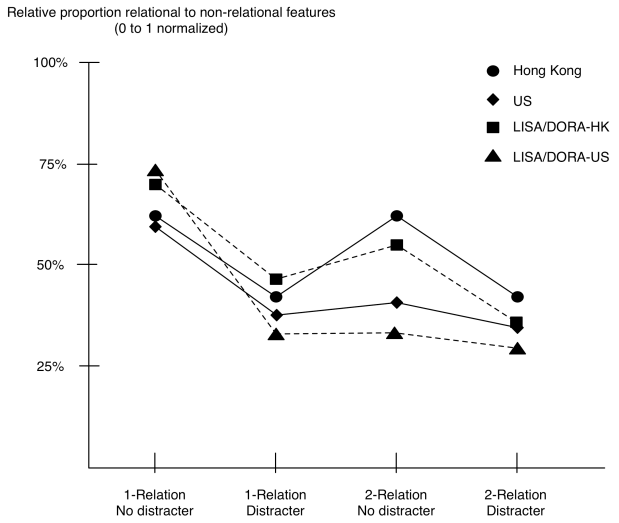

Figure 3: Experimental (Richland et al. 2010) and Simulation,.

General Discussion

In this paper we presented simulations in LISA/DORA that support the hypothesis that both maturation of inhibitory control in working memory and development of knowledge representations is critical for the development of adult-like analogical reasoning. Specifically, we demonstrated that simple changes in inhibition levels in LISA/DORA (i.e., inhibition between elements of competing relational representations in working memory) coupled with DORA's predicate learning routines could account for both relational complexity and featural distraction effects in young children's analogical reasoning performance across cultures. In contrast, approaches based on knowledge accretion and capcity changes in isolation seem unable to capture all of these results.

We conclude that both relational knowledge acquisition and inhibitory control in working memory shape an individual's analogical reasoning performance. We suggest that the development of analogical reasoning in children can be conceptualized as an interaction between these two factors. As children age their knowledge about relations advances while their working-memory capacity as modulated by inhibitory control also improves. At a given time during development, the child is able to perform an analogical task based on both their level of relational knowledge and their working-memory resources. Specifically, the equilibrium operates such that greater relational knowledge can impose fewer processing demands while less knowledge imposes higher demands. Thus, Hong Kong children given the same working-memory resources can better solve relational complex problems. Thus, as relational knowledge increases in a domain, the demands on working memory decline, allowing for more complex reasoning at any given age. This pattern in cognitive development builds on an understanding of working-memory effects in expertise (e.g., Chase & Simon, 1973) where advanced relational knowledge can decrease processing demands and thereby allow experts to accomplish cognitive tasks which novices cannot.

We believe that to truly understand the development of relational reasoning in children, future experimental and computational studies must take into account both advances in relational knowledge and changes in inhibitory control in working memory, and importantly, studying how these two aspects of development interact.

References

Andrews, G. & Halford, G.S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, *45, 153-219.*

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10,* 48-71.

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

Diamond A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry (pp. 466-503). In D.T. Stuss & R.T. Knight (Eds). *Principles of frontal lobe function*. London: Oxford University Press.

Doumas, L. A. A. & Hummel, J. E. (2010). A computational account of the development of the generalization of shape information. *Cognitive Science, 34*, 698-712.

Doumas. L. A. A., Hummel, J. E., & Sandhofer, C. M. (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.

Gentner, D. (2003). Why we're so smart. In D. Gentner and S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought*, 95-235. Cambridge, MA: MIT Press.

Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487-513.

Gentner, D. , & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P.Byrnes (Eds). *Perspectives on thought and language: 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.

Goswami, U. (1992). *Analogical reasoning in children*. Hillsdale, NJ: Erlbaum.

Goswami, U. (2001). Analogical reasoning in children. In D. Gentner, K. J. Holyoak & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 437-470). Cambridge, MA: MIT Press.

Goswami, U., & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition, 35*, 69-95.

Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.

Halford, G. S., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the balance scale: The influence of relational complexity. *Journal of Experimental Child Psychology*, *81*, 383 – 416.

Hedden, T., Park, D. C., Nisbett, R., Ji, L., Jing, Q., & Jiao, S. (2002). Cultural variation in verbal versus spatial neuropsychological function across the life span. Neuropsychology, 16(1), 65-73.

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. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220-264.

Leung, V. W.-Y. (1998). The use of nouns versus verbs in Cantonese-speaking children's early vocabularies and their mothers'speech. PhD dissertation, University of Hong Kong.

Morrison, R.G. (2005). Thinking in working memory. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 457-473). Cambridge, UK: Cambridge University Press.

Nisbett, R. E. (2003). *The geography of thought: How Asians and Westerners think differently ... and why.* New York: The Free Press.

Piaget, J., Montangero, J., & Billeter, J. (1977). La formation des correlats. In J. Piaget (Ed.) *Recherches sur l'abstraction reflechissante I* (pp. 115-129). Paris: Presses Universitaires de France.

Rattermann, M.J., 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*, pp. 453-478

Richland L.E., Chan, T-K., Morrison, R.G., & Au, T.K-F. (2010). Young children's analogical reasoning across cultures: Similarities and differences. *Journal of Experimental Child Psychology, 105,* 146-153 .

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.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. Developmental Psychology, 32, pp. 492-504.

Tardif, T. (2006). But Are They Really Verbs? Chinese Words for Action. In. K. Hirsh-Pasek, K. Golinkoff, R. Michnick; *Action meets word: How children learn verbs.* New York, NY, US: Oxford University Press, pp. 477-498.

Tardiff, T. Gelman, S. A., Xu, F. (1999). Putting the 'noun bias' in context: A comparison of English and Mandarin. *Child Development, 70*, 620-635.

Tardif, T. Shatz, M. Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language,* 24, 535-565.