A Connectionist Account of Analogical Development

Robert Leech (r.leech@psychology.bbk.ac.uk)
School of Psychology, Birkbeck, University of London, Malet Street London
WC1E 7HX, United Kingdom

Denis Mareschal (d.mareschal@bbk.ac.uk)
School of Psychology, Birkbeck, University of London, Malet Street London
WC1E 7HX, United Kingdom

Richard Cooper (r.cooper@bbk.ac.uk)
School of Psychology, Birkbeck, University of London, Malet Street London
WC1E 7HX, United Kingdom

Abstract

We present a connectionist model that provides a mechanistic account of the development of simple relational analogy completion. Drawing analogies arises as a bi-product of pattern completion in a network that learns input/output pairings representing relational information. Analogy is achieved by an initial example of a relation priming the network such that the subsequent presentation of an input produces the correct analogical response. The results show that the model successfully solves simple A:B::C:D analogies and that its developmental trajectory closely parallels that of children. Finally, the model makes two strong empirical predictions.

Introduction

"Analogy lies at the core of human cognition" (Holyoak, Gentner & Kokinov, 2001). Analogies underlie creative thought and problem solving, and as such are implicated in virtually all aspects of human life, including cognitive development. Children spontaneously use analogy to extend their knowledge about the biological and physical world and to solve problems (Goswami, 1996). Research by Paen and Wilkening (1997) has suggested that analogies occur spontaneously with young children to explain physical problems. This is consistent with the studies of Ingaki and Hatano (1987) that suggested children spontaneously use analogies involving humans to explain the behaviour of other animals but not stones. There is also some evidence that 10-month-olds infants are able to reason by analogy to solve problems (Chen, Sanchez & Campbell, 1997).

Given its importance to cognition, it is not surprising that analogy has been the subject of many highly specified theories and computational models going back to the early 1960s (French, 2002). However, although there has been a considerable amount of research into the analogical abilities of children, there have been no specifically developmental models of the underlying mechanisms. The existing developmental accounts have either been adaptations of existing adult models (e.g. Gentner, 1989) or else under specified verbal theories (e.g. Goswami, 1996).

Our work is an attempt to provide a mechanistic account of the emerging ability to draw analogies. This account suggests that analogical abilities develop through the interactions of a simple memory system and the gradual accumulation of factual knowledge. No special mechanisms are required to deal with analogy.

The rest of this paper unfolds as follows. First, we consider current theories of analogical development, and highlight the key phenomena that a successful developmental model should capture. Next we present, the Goswami and Brown (1989) experiment in greater detail and offer it as a target modelling task. We then present the model architecture, training, and how it forms analogies. The results show how the model performs when presented with analogies and non-analogies. The developmental profile of the model is described, as are the errors the model makes. Finally, we discuss two novel empirical predictions derived from the model.

Theories of Analogical Development

Some of the earliest research into the development of analogical reasoning was carried out by Piaget (Piaget, Montangero & Billeter, 1977). He found that children younger than 7-years-old made many errors on formal analogical problems. He suggested that it was not until the onset of adolescence that children consistently and successfully completed analogies. Piaget saw the age-related increase in competence as mirroring the development of other cognitive reasoning skills. The performance of the older children corresponded to his postulated formal operational stage.

One more recent domain general account suggests that changes in analogical reasoning reflect increased working memory capacity with age. Halford, Wilson & Phillips (1998) argue that one of the most fundamental constraints acting on cognitive development is the relational complexity that can be processed in parallel in working memory. They further suggest that the ability to process binary relations

(relations with two arguments) allows for the ability to form a:b::c:d analogies.

Gentner offers a different account of the development of analogical reasoning by applying her structure mapping theory of adult analogical reasoning (Gentner, 1983) to cognitive development. This framework argues that analogy is the mapping of systems of relations from a base domain to a target. She suggests that there is a relational shift that occurs in children's analogies through development (Gentner & Toupin, 1986; Gentner, 1988). Children's analogical reasoning changes from being initially based on the surface similarity of object attributes to gradually including relational information between objects and then later on involving systems of relations.

In addition, analogical competence varies from domain to domain. Thus, some researchers such as Goswami (1996) and Gentner (1989) suggest that the crucial constraint on analogical development is the knowledge that the child has about the relevant relations. Indeed, Goswami and Brown (1989) showed that if sufficiently well known relations were used 4- and 5-year-olds could complete analogies of the kind Piaget had suggested they would fail. They also found that successful analogy performance correlated with an independent test of the child's knowledge of a domain. Similar conclusions were reached by Rattermann and Gentner (1998).

Key Phenomena

From the developmental literature, it follows that any successful theory of the development of analogical reasoning has to capture the following key components:

(1) The relational shift - where there is a move in preference from judging similarity on the basis of attributes to relations (Gentner, 1988; Gentner & Rattermann, 1998; Gentner & Toupin, 1986).

(2) The strong role of relational knowledge underlying successful analogical reasoning (Goswami & Brown, 1989; Gentner & Rattermann, 1998).

(3) The domain specific nature of this knowledge and consequently the domain specific change over development in the ability to form analogies (Goswami & Brown, 1989).

(4) Spontaneous analogical ability within a domain, without the need for specific teaching of analogies (Paen & Wilkening, 1997; Ingaki & Hatano 1987).

(5) Finally, knowledge acquisition *over time* - something that children do, but which is almost entirely absent from developmental work on analogy so far.

The Goswami and Brown Task

These key phenomena are all, by and large, present in Goswami and Brown's (1989) experiment, making it a good starting point for modelling analogical development. This experiment is a variant of the classic IQ item analogy test of the form "A is to B as C is to D". For example, "dog is to puppy as cat is to…?", with "kitten" as the correct D term. Goswami and Brown (1989) wanted to make the task easy enough for young children by showing children pictures of

familiar objects (such as lemons, bread, televisions) in different familiar states (cut lemon, switched on television). For example, the bread could be either cut into slices or uncut in a loaf. Children were shown the A, B and C terms of an analogy (Playdoh, cut Playdoh and apple) and were asked to choose from four pictures. These included the correct answer and three distractors. The distractors were: (1) a different object with the same analogical causal change (cut bread), (2) the same object non-analogically causally changed (bruised apple), (3) a semantically similar object (a banana), and (4) an object with a similar appearance (a ball).

A Theory and a Model

There are two central questions which the current account attempts to answer. The first question is what mechanisms drive analogy formation in children? In the present work, analogy completion is driven by priming. For an A:B::C…? analogy, the A and B terms prime a relation which then biases the C term to produce the analogically appropriate D term. This idea is consistent with Schunn and Dunbar (1996) who found evidence of priming with earlier analogous tasks improving subjects' performance in complex tasks whilst they were unaware of the analogy.

The second question is what mechanisms underlie analogical development? The development of the analogical abilities is achieved through the gradual acquisition of knowledge about relations. The model learns causal relational information from which analogies can be formed. As in the Goswami and Brown tasks, these are simple familiar relations; for example, *cutting* or *melting*.

A causal relation is assumed to be embodied in the transformation of state A to state B (e.g. the transformation of "apple" to "cut apple"; see also Jani & Levine, 2000). In other words, there is no explicit coding of the relation itself. The similarity between two different examples of a relation (e.g. the similarity between apple:cut-apple and bread:cut-bread) lies in how the perceptual features of the objects involved are transformed.
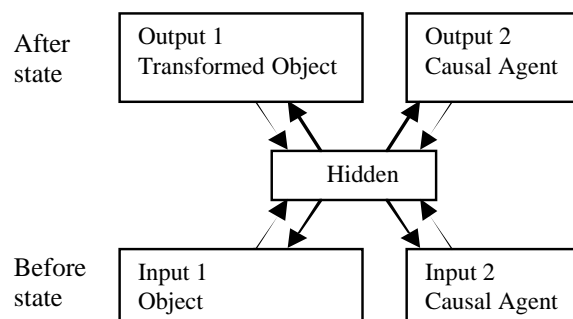

Figure 1. Schema of the model architecture. The input layer codes objects before a transformation, and the output layer codes objects after a transformation has occurred.

Figure 1 shows the architecture of the connectionist network used to model both the acquisition of relational information

and the completion of analogies. All network weights are bi-directional and symmetrical. The input layer is split into two banks of units, representing the presentation of two different objects in a "pre"-transformation state. Similarly, the output layer is split into two banks, representing the same two objects in their "post"-transformation state. At both input and output, objects are encoded in terms of perceptual features only (e.g. shape, size, color).

Input 1, Output 1 and the hidden layer have 40 units each. Input 2 and Output 2 have 10 units each. The activation of any unit varies according to a sigmoidal activation function from 0 to 1. The initial weights are uniformly randomised between ±0.5.

Because of the bi-directional connections, input activation can cycle through the network before settling into a stable attractor state. During training, contrastive Hebbian learning was used to change the connection weights so the attractors on the output units coincided with target output states of the network (O'Reilly, 1996). The learning rate was set to 0.1, and activations were updated for 50 cycles between each weight update.

The networks were trained on input patterns produced on-the-fly by adding Gaussian noise (μ=0.0, σ=0.2) to prototypes. The prototypes were randomly generated input vectors within the range [0, 1]. There were 20 different possible Input 1 (object) prototype patterns, and 4 different Input 2 (causal agent) prototype patterns. This was meant to capture the fact that although two instances of cutting an apple with a knife are similar, they are not identical.

Four transformation vectors were also randomly generated. The transformation vectors encode the relation between the pre- and post transformation states of the object. In fact, the transformed state of the object (Output 1) is obtained by adding a transformation vector to an object (Input 1). For example, Input 1 (e.g., apple): [0.5 0.2 0.8 0.2 0.4], might be transformed by the vector (e.g., cut): [-0.4 0.0 0.0 0.0.7 0.0] resulting in Output 1 (cut apple): [0.1 0.2 0.8 0.9 0.4]. Note that while the transformation vector is used to generate the output corresponding to any particular input, the transformation vector itself is never presented to the network. Different objects (e.g. "bread" or "apple") transformed by the same relation (e.g. *cutting*) are transformed by the same vector. Thus, the network can learn about a particular transformation by generalising across sets of Input/Output pairings.

Input 2 represents a causal agent (e.g. "a knife") which when it occurs alongside certain objects represented at Input 1 (e.g. "an apple") leads to a transformation of that object at Output 1 (e.g. "a cut apple"). Consequently, the target pattern for Output 1 (the post-transformation state of the object) depends on Input 2 patterns. Output 2 is always the same as Input 2.

Training consists in randomly selecting an object and an agent, computing the output states, and updating the weights such that the actual output state produced by the network moves closer to the target output state. One can think of the input and target output states as two temporally contiguous states of the world. Because both the input and target can be obtained by direct observation of the world, learning of relational information does not require an external teacher, and constitutes a form of self-supervised learning (Japkowicz, 2001). Here, each of the 20 Input 1 objects can be affected by 2 of the 4 causal agents and so 2 of the 4 transformations. Thus, there are 360 potential analogies for the network to be tested on. However, when an object is presented in conjunction with one of the remaining 2 causal agents, the target Output 1 is the same as the untransformed Input 1. Thus, the presence of the causal agent alone is not a predictor of whether a transformation will occur.

The testing of analogy completion proceeds in a very different way from learning of relation information. Priming underlies the network's ability to complete a:b::c:d analogies. Priming occurs because the bi-directional connections allow the hidden and output units to maintain activity resulting from an initial event (e.g., an a:b event). The activity that is maintained in the network impacts on how new input is then subsequently processed (e.g., a c:? input).

Consider the following example. First, the input and output units are clamped with the "apple" at Input 1 and "cut apple" at Output 1 while Input 2 and Output 2 are initially set to 0.5, the resting value. This corresponds to being presented with the information apple:cut apple (i.e., the first half of an a:b::c:d analogy). The causal agent is not presented to the network at any point during testing. After 50 activation cycles the network settles into the attractor state it was trained on by filling in Input 2 and Output 2 and arriving at hidden unit activations consistent with the transformation *cutting*. Following this, the Input 1 and Output 1 units are unclamped and a second pattern, corresponding to "bread", is presented to Input 1 and nothing at Output 1. Input 2 and Output 2 are initially presented with resting activation patterns and then unclamped. This corresponds to being presented with the information bread:? (i.e., the second probe half of the a:b::c:d analogy). By unclamping the original inputs and outputs and by presenting a different input pattern, the network is no longer in equilibrium and settles into a new attractor. During training, the network has encoded in the connections to and from the hidden layer the similarities in the transformations corresponding to relations such as *cutting*. Consequently, the prior priming of the "apple" and "cut apple" transformation biases the network to settle into the attractor state corresponding to "bread" clamped at Input 1 and the transformation *cutting*, which gives "cut bread" at Output 1. Now the network has produced the appropriate response at Output 1 to complete the analogy (i.e., apple:cut apple::bread:cut bread).

Results

The network is only ever trained to predict the causal consequences of an object and an agent. During training, the output SSE is reduced from 4.31 initially to 0.88 by 10,000 epochs. After 20,000 epochs the network is able to reliably

produce at output all the correct patterns with a mean SSE < 0.65.

An example of analogy and non-analogy

Figure 2 shows the changing activations of a fully trained network as it settles into a stable attractor state in response to new input. The different lines represent the sum of squared difference (SSD) between the actual output produced and four possible distractor patterns. The distractors are comparable to those used in the Goswami and Brown task. The pattern with the lowest SSD can be thought of as the current choice of the network. In Figure 2a the network is presented with the analogy problem "bread:cut bread::apple:...?" and produces the analogically appropriate response "cut apple".

Moreover, if the network is completing analogies appropriately, it should not apply analogies where these are inappropriate. Figure 2b shows the network's response when presented with the non-analogy "bread:bread::apple...?". Here, the network correctly produces the untransformed "apple" at Output 2. Figures 2a and 2b both show how the network responds when presented with the "apple" input pattern. However, the network's output is different. These cases differ only in what preceded the test input "apple", and whether this primed a transformation.
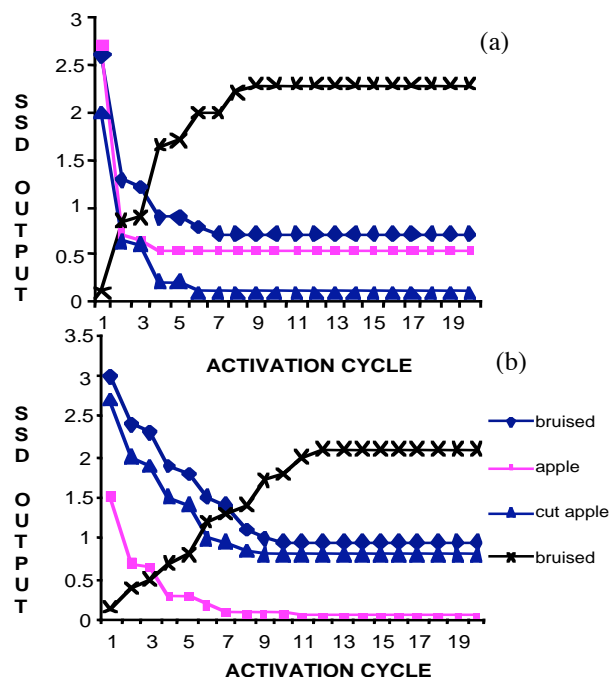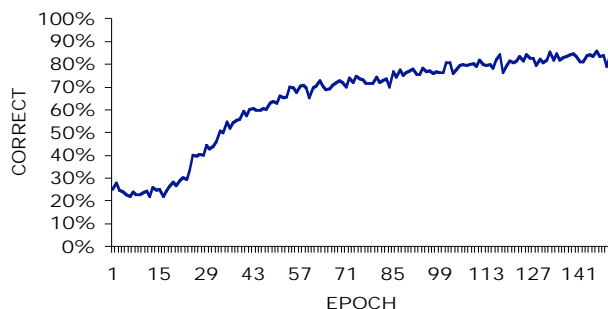

Figure 2a and 2b: The sum of squared difference between the actual output activation and four training patterns after 20,000 epochs of training. 2a is after presentation of an analogy. 2b is after presentation of the non-analogy.

Developmental profile

Figure 3 shows the network's performance when tested on all 360 analogies across time. If the analogically appropriate response has the lowest SSD then the network is assumed to have successfully completed the analogy. After 100 epochs of training, only 29% of analogies are completed successfully. However, by 20,000 epochs of training, the network produces the analogically appropriate response for almost 85% of analogies.

The network completes different analogies at different points in its training. The percent of analogies correct has an inverse relationship to the average sum of squared error

(SSE) at output. So, the percent of analogies correct is related to the network's knowledge of the transformations.
Figure 3. Percentage of analogically appropriate responses at output over training. These results are the average of 20 replications.

This profile parallels the evidence from the developmental literature suggesting that domain specific changes in the abilities of children correlate with their domain knowledge of the relations involved (e.g., Rattermann and Gentner, 1998).

Analogical errors

It is important to compare the kind of errors that the network makes with those made by children. The network made two types of errors: appearance errors (where the network responded at Output 1 with the same object as at Input 1) and transformational errors (where the network produces the correct changeable object at Output 1 but transformed by a non-primed relation). These are the same kinds of errors predominantly made by children (Rattermann and Gentner, 1998).

	Children		Networks	
Response	4-year-olds	5-year-olds	2200 epochs	3500 epochs
Correct analogy	35%	67%	36.4%	65.8%
Wrong transformation	35%	28%	34.7%	23.5%
Similarity based error	22%	3%	28.9%	10.7%

Table 1. Distribution of responses made by children and the network

In Ratterman and Gentner (1998), children were not tested on all available relations, only those where the transformations were familiar enough for them to successfully complete analogies. Indeed, the 6-year-olds performed at ceiling, successfully completing nearly 100% of analogies used in this study[1]. However, in the results above, the networks are tested on every possibly analogy. So, to enable a comparison of the networks performance with the children's performance, we selected a subset of 83% of analogies such that the networks would have 100% correct performance on those analogies after 20,000 epochs of training.

Table 1 compares the responses of children in the Rattermann and Gentner (1998) study with those of the networks. The network simulates the children's responses closely. The network shows a shift over training with a considerable decrease in the proportion of appearance responses (from 28.9% to 10.7%). This is matched with an increase in transformational responses (either correct or transformational error) from 71.1% to 89.3%. This is consistent with the relational shift phenomenon in which children give more transformation-based analogies as they get older.

Developmental asymmetry

One interesting phenomenon of the networks is a developmental asymmetry in analogy completion when base and target are reversed. Figure 4 shows the SSE of the output during training. It also includes arrows indicating when four randomly chosen analogies and their reverses were first successfully completed. For each analogy, there is a period when the model can complete an analogy one way but it fails to complete its reverse. For example, "apple:cut apple::bread::cut bread", is completed hundreds of epochs later than the reverse analogy: "bread:cut bread::apple:cut apple". This asymmetry has so far not been tested and so is a strong prediction about children's behaviour.
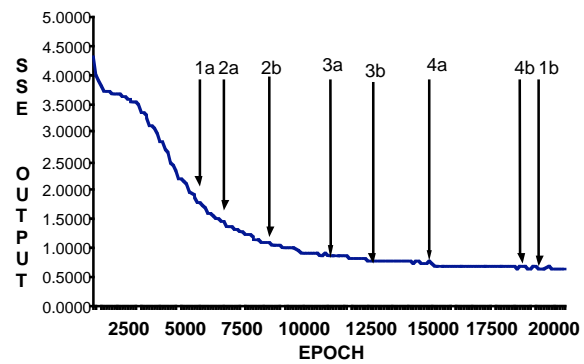
Figure 4: The average SSE at output for every input pattern over training. The arrows indicate when four analogies (numbered 1a-4a) and their reverse analogies (numbered 1b-4b) are first successfully and consistently completed.

Transformational similarity

Figure 5 highlights another interesting prediction of the network. In the network, the number of transformational errors is largely determined by the degree of similarity between the transformational vectors encoding different relations. The Euclidean distance between the
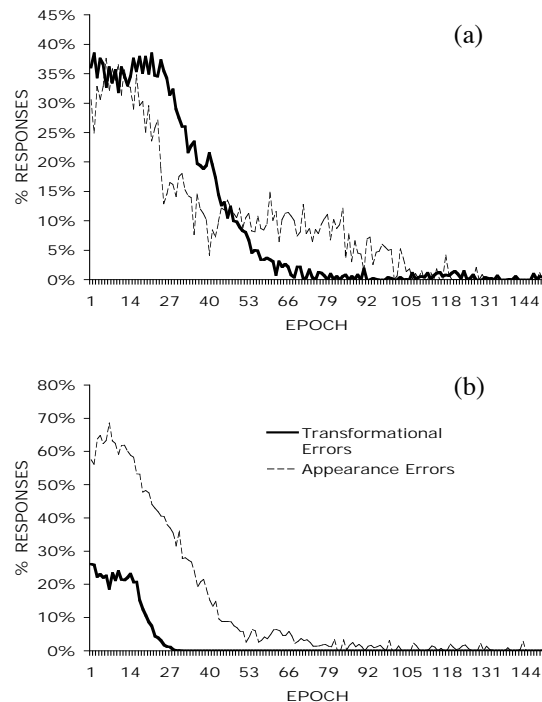
Figure 5: The percentage of responses which are one of the two types of errors when the transformations are (a) more or (b) less similar. The results are the average of 20 replications.

---

[1] There are many other analogies that six-year-olds would perform very poorly on (e.g. Piaget et al, 1977).

transformational vectors determines the number of transformational errors. The transformation vectors for the simulation results depicted in 5a have a maximum Euclidean distance of 10 from each other, whereas distance between those for the results depicted in 5b is greater than 15. There are many fewer transformational errors in 5b than in 5a. The implication of this is that analogies using more distinct transformations should produce fewer transformational errors with children.

Discussion

This work presents a connectionist account of the development of simple analogy completion. Analogy is seen as a form of relational priming. A relation, at least for the simple cases considered here, is the transition from one state to another causally related state. These assumptions allow the network to account for several key developmental phenomena observed with children.

When we consider the network's errors, it demonstrates a relational shift paralleling that in children. Furthermore, this transition from drawing analogies on appearance to relations occurs only through increasing training. This again conforms to the developmental evidence suggesting that knowledge accretion underpins the development of analogical reasoning. A second consequence of analogy depending on training is that analogies involving different relations and different objects are completed at different times. This fits with the domain specific nature of analogical reasoning highlighted in much of the developmental literature.

It is important to note that the network is never directly trained on analogies. Its ability to complete analogies emerges from the way relational information is represented and tested. This is consistent with developmental studies that suggest untrained analogy use in young children in some circumstances.

The model makes two novel predictions. The developmental asymmetry prediction differentiates between the present model and structure-mapping accounts, which do not naturally suggest such an asymmetry. The transformational error prediction is interesting because until now most empirical work has focused on the structural or perceptual similarity of the objects involved in an analogy. Instead, this prediction derives from the transformations of objects and their impact on analogical reasoning. We are currently testing these two predictions.

The analogical abilities of the network are achieved without an explicit structure-mapping system or any built-in syntactic structure. Of course, the current model is unable to account for many more complex child and adult analogies, (e.g. Rutherford's solar system - atom analogy). However, the present work suggests that, at least for simple cases, simple memory-based mechanisms may suffice to explain behavior.

References

Chen, Z., Sanchez, R.P., & Campbell, T. (1997). From beyond to within their grasp. *Developmental Psychology, 33*, 790-801.

French, R. M. (2002). The Computational Modeling of Analogy-Making. *Trends in Cognitive Sciences, 6*, 200-205.

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, 59*, 47-59.

Gentner, D., Holyoak, K. J., & Kokinov, B. (Eds.). (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: MIT Press.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277-300.

Goswami, U. (1996). Analogical reasoning and cognitive development. Advances in *Child Development and Behaviour, 26*, 91-138.

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

Halford, G.S., Wilson, W.H. & Phillips, W. (1998) Processing capacity defined by relational complexity: *Behavioral Brain Sciences, 21,* 803-831.

Ingaki, K., & Hatano, G. (1987). Young children's spontaneous personifications as analogy. *Child Development, 58*, 1013-1020.

Japkowicz, N. (2001) Supervised and unsupervised binary learning by feedforward neural networks. *Machine Learning, 42*, 97-122.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences. *Neural Computation, 8*, 895-938.

Pauen, S. & Wilkening, F. (1996). Children's analogical reasoning about natural phenomena. *Journal of Experimental Child Psychology, 67*, 90-114.

Piaget. J., Montangero, J. & Billeter, J. (1977) La formation des correlats. In J. Piaget, Ed., Recherces sur L'Abstraction Reflechissante I. Paris: Presses Unversitaires 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*, 453-478.

Schunn, C.D. & Dunbar, K. (1996). Priming, Analogy, & Awareness in complex reasoning. *Memory and Cognition, 24*, 271-284.