What moves in a mysterious way? A domain-general account of learning about animacy and causality

Gary Lupyan (glupyan@cnbc.cmu.edu)
Department of Psychology, Carnegie Mellon University
Center for the Neural Basis of Cognition
Pittsburgh, PA 15213 USA

David H. Rakison (rakison@andrew.cmu.edu)
Department of Psychology, Carnegie Mellon University
Pittsburgh, PA 15213 USA

Abstract

A fundamental debate within cognitive science concerns how infants, children, and adults learn about the motion properties of animates and inanimates. In this paper, we show that an associative-learning mechanism implemented as a neural network can use animacy relations to predict and discriminate between familiar and novel causal and noncausal events. This is possible because (1) animate objects are more similar to each other than to inanimate objects, (2) causal events are more likely to have animate agents and inanimate recipients than noncausal events, (3) noncausal events are more likely to have animate (i.e., self-propelled) "recipients", and (4) causal and non-causal events correlate with different types of motions. The results suggest that the emergence of concepts for animacy and causality in infancy and beyond can be explained without theory-driven top-down processing or specialized modules.

Introduction

Over the past twenty years, infants' perception and understanding of causality has received increasing attention in the developmental literature. This burgeoning interest reflects a theoretical consensus that an appreciation of cause and effect relations is a cornerstone of the ability to understand the way the world works. It is now well established that infants in the first year of life are sensitive to aspects of causality including agency and recipiency. Leslie and Keeble (1987) showed 6½- to 7-month-old infants a series of simple launching events based on those developed by Michotte (1963). In the direct launching condition, infants were habituated to a green brick-shaped object that moved from left to right across a screen and contacted a red brick-shaped object that then moved in the same direction until off the screen. In the delayed launching condition, infants were habituated to similar events except that there was a short delay between impact and reaction. During the test phase for both conditions, infants were presented with the same basic event seen during the habituation phase except that it was reversed.

The rationale for this design was that the reversal of the direct launching event switched the agent-recipient relationship and the spatiotemporal properties from those seen during habituation but the reversal of the delayed launching event affected the spatiotemporal properties alone. Infants at 7 months of age who were habituated to the direct launching event recovered visual attention to the reversal more than the infants who were habituated to the delayed launching event. Leslie and Keeble (1987) interpreted these results to mean that infants in the direct launching condition were sensitive to the causality in the event. These results were extended by Oakes and colleagues (Oakes & Cohen, 1990; Cohen & Oakes, 1993) who found that infants at 6 months are unable to discriminate causal from non-causal events.

An important theoretical question that has been hotly debated concerns when and how infants learn that animates tend to act as agents and that inanimates tend to act as recipients of a causal action. According to one perspective, the complexity of the learning space implies that perceptually-based associative processes that allow infants to represent causality in simple Michotte-like events are insufficient to account for how they acquire concepts that include motion characteristics of different object kinds. More generally, it has been suggested that associative learning alone cannot act as the foundation for early representations because there are so many correlations in the world to which one could attend that it is impossible to know which ones are important for category membership and which are not (Keil, 1981).

As a solution to these problems, a number of theoretical frameworks proposed that knowledge about the motion properties of animates and inanimates is acquired via innate specialized processes or modules (e.g., Leslie, 1995; Mandler, 1992; Gelman, 1990; Premack, 1990). Leslie (1994, 1995), for example, theorized that infants possess three innately derived modules that, in combination, allow infants rapidly to develop an understanding of the physical (theory of body), psychological (theory of mind), and cognitive properties of animates and inanimates. According to Mandler (1992, 2000, 2003), infants possess an innate specialized process called perceptual analysis that recodes the perceptual display into an abstract and accessible construct. This process generates image-schemas, or conceptual primitives, that summarize crucial characteristics of objects' spatial structure and movement such as agency and recipiency.

There are a number of problems with these accounts. First, there is little, if any, direct evidence that infants possess innate modules or specialized mechanisms that allow them to encode distinct kinds of information (and in particular, motion characteristics). It also remains opaque how an appropriate module or mechanism is "triggered" by certain kinds of input but not others. Second, the notion that infants possess two separate mechanisms for concept formation – one for perceptual information and one for conceptual information—is unparsimonious and creates a heavy biological burden (Quinn & Eimas, 2000; Quinn, Johnson, Mareschal, Rakison, & Younger, 2000; Rakison & Hahn, 2004). Third, in the absence of empirical evidence it is unclear how a specialized process—for example, perceptual analysis—abstracts dynamic, motion related information into a simpler, more available form or whether such a process is different from perceptual categorization of movement patterns (Quinn & Eimas, 2000; Rakison, 2003). With these problems in mind, we propose that there are sufficient regularities in the world for an associative mechanism to account for learning in this domain. The goal of the current work was to test this proposal.

The use of Michotte launching tasks to study causality has been modeled before by Chaput & Cohen, 2001. The current paper differs in significant ways from this model, the most notable of which is the use of representations of real-world objects, allowing us to use their distributional statistics to test the claim that an associative-learning mechanism with no theory of causality or animacy can use the emergent animacy distinction to discriminate causal events from noncausal events and generalize its learning to novel scenes.

Aims

Through analysis of semantic representation corpora we aim to show that animacy emerges as a highly salient distinction. With no theory of animacy on the part of the learner, animates may nevertheless cluster in the conceptual space with other animates, and inanimates with other inanimates. We proceed to show that an associative-learning mechanism (implemented as a neural network) can use this property of clustering among animates and inanimates to predict whether novel events show a causal or noncausal relationship, while also predicting what types of motion the objects in the events should engage in. The model additionally examines whether this mechanism is sufficient to demonstrate the kinds of dishabituation patterns reported in the empirical literature, for instance, dishabituating to noncausal events when trained on causal events.

The structure of animacy

Our first step was to determine whether animacy is indeed a salient feature in concept representations. We coded the animacy status of the 539 objects in the McRae et al. corpus (in press) and then performed a principal-components analysis on the between-concept cosine matrix for all the concepts. For instance, the entry for "alligator" had a value of 0 for "airplane" (i.e., alligators and airplanes shared no features produced by the raters), 0.13 for apple, 0.02 for axe, and so on. To quantify the degree to which the animacy distinction is represented in the correlation matrix, we then performed a K-means cluster analysis on the first 50 principal components of the matrix. Of the 138 animate concepts, 134 (97.1%) were correctly classified. Of the 401 inanimate concepts, 399 (99.5%) were correctly classified, χ^2=502.80, p<.0005 (Figure 1). The animate concepts that were incorrectly classified by the algorithm were python, snail, worm, and caterpillar, and the 2 misclassified inanimates were airplane, and jet. Notice that all the incorrectly classified animates lack legs. Of the incorrectly classified inanimates, airplane and jet arguably have movement as a prominent property—evidence that motion properties are critical to animacy. Additionally, we found that animates were more tightly grouped, having on average significantly smaller distances from the cluster centroid than inanimates: two-sample t-test not assuming equal variance, $t(352)$ = 14.11, p<.0005.

One objection to drawing inferences concerning the salience of animacy in concepts using the method just described is that some features provided by the human raters were directly related to animacy. For instance, animal concepts had the feature "is animal"; therefore all animals would correlate with each other based on at least that one feature. If features related to animacy are the most consistent (as in the case of animals possessing the "is animal" feature), a separation of animate and inanimate concepts based on the concept correlation matrix would be merely a confirmation that features related to animacy are ones most consistently provided by adults, and that these features are sufficient to cluster animates and inanimates, which although interesting does not address the question of how infants would know which objects are animals. After all, if animacy the distinction comes from features such as "is animal," infants would first have to be innately sensitive to these features, a position we reject.

To provide an independent confirmation of the idea that animacy is a salient factor in item concepts, we applied the same clustering procedure to a completely different semantic corpus—the correlated occurrence analogue to lexical semantics (COALS: Rohde, Gonnerman, & Plaut, 2006; http://dlt4.mit.edu/~dr/COALS/). Rather than being based on features provided by human raters, the semantic representations in this corpus are automatically generated from large text corpora (the method is similar to HAL, developed by Lund and Burgess, 1996). The algorithm produces vectors with similarity relations based on the similarity of the contexts in which the words are used. Relationships between semantic vectors generated by the COALS method have been shown to correlate well with similarity ratings provided by human raters. Other than not coding for specific features, an advantage of COALS-generated representations for present purposes is that this method has been optimized to produce binary vectors, which are suitable for use in training neural networks.

For consistency, we performed this analysis on the same concepts that were coded in the McCrae et al. (in press) corpus. Because corpus-based semantic representations cannot differentiate between homonyms, we removed words with obvious dual meanings (e.g., "bat"), leaving 513 concepts. Applying K-means cluster analysis to the first binary 50 binary dimensions of the representations revealed that animacy was saliently represented. Of the 379 inanimate concepts, 316 (83%) were correctly classified. Of the 134 animate concepts, 119 (89%) were correctly classified, χ^2=225.3, p<.0005. The animates were again found to be more tightly clustered than the inanimates, $t(176)$=6.71, p<.0005. This lower, but still impressively high clustering results from analyzing a highly compressed dataset: 50-dimensional binary vectors from—the full corpus contains 1,500 real-valued dimensions for each concept. The clustering is possible with no reliance on explicit features generated by humans.

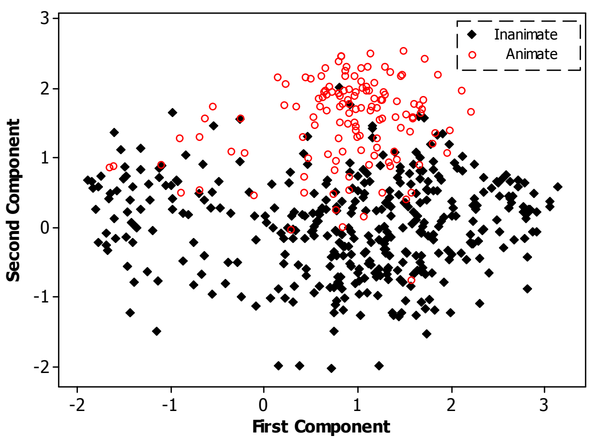

Figure 1: A score plot of the first 2 principal components of the 50-dim concept vectors, grouped by animacy.

From animacy to causality

The clustering of animates and inanimates leads to the prediction that after observing animates and inanimates interact in "causal" and "noncausal" ways, an associative-learning mechanism should be able to generalize the learned patterns to novel scenes. For instance, after observing that a cat caused a ball to roll through physical contact, an infant could infer that a dog would need to make contact with a pencil to make it move. To test this hypothesis in the current model, we trained a connectionist network on "scenes" constructed from the concepts listed in the McRae et al. (in press) corpus. Each "scene" consisted of two objects and a motion. The objects were randomly chosen from the corpus consistent with the animacy relations shown in Table 1. The motions depended on the intended "causality" of the action. Causal actions involved the direct motion while noncausal, actions (indicating self-propulsion) had a gap, delay, or gap+delay motion. The probabilities used to generate the relations are meant to correspond roughly to the real world, but the results we report do not depend on the precise values shown in Table 1.

The hypothesis that an associative learning mechanism should generalize "causality" to novel scenes was tested in three ways. First, we examined whether after being trained on a subset of causal and noncausal scenes, the network could predict the correct causal or noncausal motion when presented with new scenes. Second, we explored whether the network showed an increase in error when the exemplars typically associated with a causal scene (e.g., animate agent, inanimate recipient) are mismatched with a noncausal motion (or vice-versa). An increase in error indicates a violation of expectations. And because looking is associated with a violation of expectations (Gilmore & Thomas, 2002), error has been used as a proxy for looking time (Sirois & Mareschal, 2001). A reliable increase in error in a network is therefore comparable to dishabituation. Third, we tested whether, after training, the hidden representations of the network clustered the causal and noncausal events into separate groups when tested on novel scenes.

Method

Architecture

The model was implemented as a simple recurrent network (SRN; Elman, 1990) trained using a variant of standard backpropagation with momentum (Rohde, 1999). The input consisted of patterns of activity across three groups of units, corresponding to the first object, the second object, and the motion in the scene (see *Materials* below). The input layers were fully connected to hidden units which in turn projected to a "context group" and the output groups (Figure 2). The context group projected back to the hidden layer using copy-back connections, providing the network with a simple form of memory, necessary to learn the temporal sequence of the motions involved in the training scenes.

Materials

We used the binary semantic vectors generated using the COALS method (Rohde et al., 2006). Fifty dimensions were used. Motion was represented using an 8-unit layer. On each timestep a single bit in the motion layer was active according to the current position of the first or second object. The networks' task was to map the object layer to itself (auto-association), and predict the correct active bit of the motion layer occurring at the next time-step. There were four types of motion: *direct*—a smooth linear motion corresponding to one object contacting the other, and the second object starting to move after the contact; *gap*—identical to *direct* except for a spatial gap in the middle of the event, corresponding to smooth movement, but no contact; *delay*—identical to *gap* except for a temporal rather than a spatial interruption; *gap+delay*—a combination of both a spatial and a temporal gap. The *direct* motion was associated with the causal events, while the *gap*/*delay*/*gap+delay* motions were randomly paired with the noncausal events.

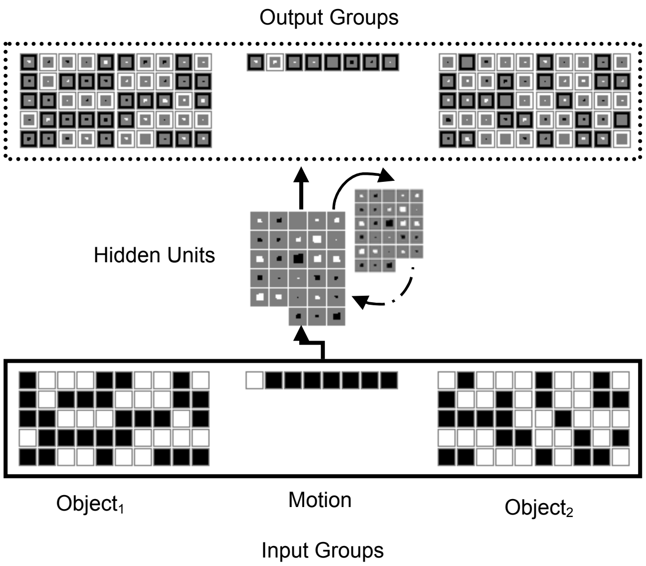

Figure 2: A simplified schematic of the simple-recurrent network architecture. All input and output groups are symmetrically connected to the hidden units. See text for details.

The training corpus consisted of 60 scenes each composed of two objects and a motion. Animate and inanimate concept vectors were randomly selected from the corpus and loaded into the $Object_1$ and $Object_2$ layers, respectively in accordance with probabilities in Table 1. For instance, 75% of the causal scenes consisted a randomly chosen animate first object, and a randomly-chosen inanimate second object. All causal scenes involved a direct launching event and one of the three noncausal events (gap, delay, gap+delay) was randomly chosen for each of the noncausal scenes. An additional 60 scenes were then generated for testing the network.

Procedure and training

Ten networks were trained for 150 epochs on 60 scenes randomly generated in accordance with the relations in Table 1. As with previous simulations, the motion layer was made more salient by increasing its output error and derivatives by a factor of 15 (a factor of 15 was used throughout the simulations in this paper whenever the global motion was the only moving part of the scene).

Table 1: Causal structure of the scenes used in training.

	1st Object	2nd Object	Motion	Prob.
Causal				
	Animate	Inanimate	direct	.75
	Animate	Animate	direct	.25
Non-causal				
	Inanimate	Animate	Gap/Delay/ Gap+Delay	.50
	Animate	Animate	Gap/Delay/ Gap+Delay	.50

Testing

The first test of the hypothesis consisted of comparing which motions were predicted by the network when it was presented with a novel scene. The outputs of the motion layer were compared against the four types of motion: direct, gap, delay, and gap+delay. The motion pattern with the smallest Euclidean distance from the network's output was taken as the response of the network for that particular scene.

One alternative to our contention that the network has learned to predict the causal and noncausal motions based on the animacy relations of the objects in the scene is the possibility that the network is predicting the motion independently of the objects involved. Because novel noncausal scenes have corresponding noncausal motions, the network may be abstracting over the objects and merely using the motion in the input layer to predict the motion at the output motion layer. To test for this possibility, we tested the network on novel scenes in which all motions were of the same causality type. If the network is generalizing the animacy of novel scenes and using this information to predict the motion, it should show a higher error when presented with a noncausal scene having a causal motion (*noncausal-to-causal switch*). The network should also show a higher error when presented with a causal scene having noncausal motions (*causal-to-noncausal switch*). To perform this comparison we compared the motion predicted by the network with the motion actually observed (e.g., always the direct motion for the *noncausal-to-causal* switch trials).

Finally, we examined the activations of the hidden units to see whether the network's representations group into causal and noncausal clusters. If this were found to be the case, it would provide an added demonstration that the network is learning to predict the causality of the scene based on the observed animacy relationship between the objects. We recorded the hidden-unit activations in the larger of the network's two hidden layers, and then performed K-means cluster analysis to determine whether this simple algorithm can correctly identify the representations of the causal and noncausal scenes. Because the network is recurrent, the representations evolve over time. We chose to examine the representations from the last timestep of each trial. At this point, all the motions are identical (i.e., they already occurred), meaning that any differences in representations can only result from the animacy relation of the objects in the scene, and a memory trace of what the motion was several timesteps prior.

Results

Table 2 shows the distribution of responses of trained networks tested on novel scenes. The results are averages of 10 networks trained and tested on separate 60-scene sets randomly constructed from the corpus containing the full 513 concepts. On average, the networks selected the direct motion 89.9% of the time when presented with a novel "causal" scene, and 18.2% of the time when presented with a novel "noncausal" scene. Collapsing across the types of non-causal motions, the network selected a noncausal response 81.8% when presented with a noncausal scene, but only 10.1% for the causal scene. The response patterns were significantly different in the two groups: Kruskal-Wallis test for group median, z = 13.9, p<.0005.

Table 2: Motion predictions (in percent) of novel causal and non-causal scenes.

	Direct Launch	Gap	Delay	Gap+Delay
Causal	89.90	0.35	0.00	9.76
Non-causal	18.21	29.39	35.14	17.25

To ensure that this response pattern did not merely indicate a sensitivity to the motion type, we compared the networks' motion predictions when presented with causal and *causal-to-noncausal* switch trials. As predicted, when networks were tested on noncausal scenes with causal motions, the errors were significantly higher than when tested on causal scenes with their expected causal motions: repeated measures mixed ANOVA with network as a random factor and scene-type as a fixed factor: $F(1,9) = 47.29$, p<.0005. To test that this effect did not result from a bias of the network to predict the direct motion, we also performed the complementary analysis of the *causal-to-noncausal* motion switch. As predicted, this showed the reverse pattern of a higher error to causal scenes with a noncausal motion than to causal scenes with a noncausal motion, $F(1,9) = 7.31$, p<.03. There was also a significant interaction between scene type and switch-type, $F(1,18) = 28.32$, p <.0005. The amount of error change from the "familiar" to the switch trials was not different in the two motion switch conditions, $t(18) = 0$, p>.5.

Lastly, we performed a K-means cluster analysis on the hidden activations produced by the network when presented with novel causal and noncausal scenes. The algorithm classified 80.7% of the causal scenes correctly, and 79.3% of the noncausal scenes correctly, χ^2=21.57, p <.0005. Note that this classification is not merely reflective of animacy of individual objects (both causal and noncausal scenes have animate and inanimate objects), but reveals a sensitivity to the relation between the objects as well as learning which kinds of motions go with which types of scenes. We can confirm this claim by determining whether the statistics of the input alone are sufficient to establish the causality nature of the scene.

Is learning about motion necessary?

It may be argued that given the initial clustering of the conceptual representations into animates and inanimates, the network may be able to cluster scenes into causal and noncausal even without learning to associate which motions go with what scenes. Because the input statistics are sufficient to cluster concepts into animates and inanimates, it may be possible to predict causality even without having to learn the correlation between animacy relations and the motion. Performing an identical K-means cluster analysis on an untrained network revealed that the input statistics were indeed sufficient to group together the causal scenes, but *not* the noncausal scenes. The untrained network correctly classified 80.7% of the causal scenes, but only 55.2% of the noncausal scenes. In other words, the animacy clustering in the corpus combined with the statistics shown in Table 1 were sufficient to cluster together the causal scenes, but insufficient to cluster together the noncausal scenes and separate their representations from those of the causal scenes.

Discussion

The current simulation demonstrates that a connectionist model can use animacy relations to predict the causality structure of both familiar and *novel* scenes. This is possible because (1) animate objects are more similar to each other than to inanimate objects, (2) causal events are more likely to have animate agents and inanimate recipients than noncausal events, (3) noncausal events are more likely to have animate (i.e., self-propelled) "recipients", and (4) causal and non-causal events correlate with different types of motions. The model produces these results without relying on any explicit representations of either animacy or causality.

The concept representations we used in the model were based on word co-occurrences in a large corpus. Clearly, infants do not have access to such a knowledge-base. However, considering the salience of animacy in these representations (e.g., it is evident from the first two principal components of an already highly reduced vector set), it is reasonable to suggest that infants, particularly those toward the end of the second year of life, are similarly sensitive to the animacy distinction (Rakison, 2005, in press). Here, we demonstrated how this sensitivity to animacy in turn enables an associative-learning device to learn to classify causal and noncausal events.

Learning about motions, in particular, is important for correctly classifying noncausal events. The added difficulty of learning about noncausal events compared to causal events has been empirically confirmed. Rakison (2005) has found that 14-month-olds successfully learned animacy relations (i.e., dynamic agent, passive recipient), but also learned relations inconsistent with the real world—static agent, dynamic recipient). Sixteen-month olds, on the other hand, only learned the relation consistent with the real world. An extended version of the current model has been used to replicate this developmental trend (Rakison & Lupyan, 2006).

Our results obviate the need for theory-driven top-down processing or specialized modules for explaining the emergence of causality concepts in infancy. The results also strongly argue that *Original Sim*—the idea that the abundance of potential correlations that can be encoded means that it is impossible to know which ones are important for category membership (Keil, 1981)—is misguided. It is true that one does not "know" which features are important for category membership—in this case for categorizing causal and noncausal events—but one does need to know which features are important. While the problem space in the current simulation is rudimentary compared to the real world, it is nothing if not complicated. The "features" used in the representations do not correspond to any ordinary features—there are no units for shape, color, size, and of course, no explicit representations of animacy. Nevertheless, there is no problem with "finding" the features that are relevant to animacy or causality—the statistical structure of the input coupled with learning about motion properties is sufficient to predict causality. While we agree that theory-driven processing (e.g., Gopnik & Nazzi, 2003) may be required to *reason explicitly* about causal relations, such explicit reasoning is not required to categorize causal and noncausal events.

References

Chaput, H.H. & Cohen, L.B. (2001). A Model of Infant Causal Perception and Its Development, Proceedings of the 23rd Annual Conference of the Cognitive Science Society, 182-187. Hillsdale, NJ: Erlbaum.

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology, 29*, 421-433.

Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science, 14,* 179-211.

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science, 14,* 79-106.

Gilmore, R. O., & Thomas, H. (2002). Examining individual differences in infants' habituation patterns using objectives quantitative techniques. *Infant Behavior & Development, 25,* 399-412.

Gopnik, A. & Nazzi, T. (2003). Words, kinds and causal powers: A theory theory perspective on early naming and categorization. In D. Rakison, & L. Oakes (Eds.) *Early categorization*. Oxford: Oxford University Press.

Keil, F.C. (1981). Constraints on knowledge and cognitive development. *Psychological Review, 88,* 197-227.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25,* 265-288.

Leslie, A. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-148). New York: Cambridge University Press.

Leslie, A. (1995). A theory of agency. In D. Sperber, D. Premack, & A.J. Premack (Eds.), *Causal cognition* (pp. 121-141). Oxford: Clarendon.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, and Computers, 28*, 203-208.

Mandler, J.M. (1992). How to build a baby II. Conceptual primitives. *Psychological Review, 99,* 587-604.

Mandler, J.M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development, 1*, 3-36.

Mandler, J.M. (2003). Conceptual categorization. In D.H. Rakison & L.M. Oakes (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 103-131). New York: Oxford University Press.

Michotte, A. (1946/1963). The perception of causality. New York: Basic Books.

Sirois, S. & Mareschal, D. (2002). Models of habituation in infancy. *Trends in Cognitive Sciences*, *6*, 293-298.

McRae, K., Cree, G.S., Seidenberg, M.S., & McNorgan, C. (in press). Semantic feature production norms for a large set of living and nonliving things. *Special issue of Behavioral Research Methods, Instrumentation, and Computers*.

Oakes, L.M., & Cohen, L.B. (1990). Infant perception of a causal event. *Cognitive Development, 5*, 193-207.

Premack, D. (1990). The infants' theory of self-propelled objects. *Cognition, 36*, 1-16.

Quinn, P. C., & Eimas, P. D. (2000). The emergence of category representations during infancy: Are separate perceptual and conceptual processes required? *Journal of Cognition and Development, 1,* 55-61.

Quinn, P.C., Johnson, M., Mareschal, D., Rakison, D., & Younger, B. (2000). Response to Mandler and Smith: A dual process framework for understanding early categorization? *Infancy*, *1*, 111-122.

Rakison, D.H. (2003). Parts, categorization, and the animate-inanimate distinction in infancy. In D. H. Rakison, & Oakes, L. M. (Eds.), *Early concept and category development: Making sense of the blooming buzzing confusion*. New York: Oxford University Press.

Rakison, D.H.. & Hahn, E. (2004). The mechanisms of early categorization and induction: Smart or Dumb Infants? In R. Kail (Ed.), *Advances in Child Development and Behavior*. *Vol 32*. New York: Academic Press.

Rakison, D. H. (2005). A secret agent? How infants learn about the identity of objects in a causal scene. *Journal of Experimental Child Psychology, 91*, 271-296

Rakison, D. H. (in press). Make the first move: How infants learn about the identity of self-propelled objects. *Developmental Psychology*. .

Rakison, D.H. & Lupyan, G. (2006) Developing object concepts in infancy: An associative learning perspective. *Manuscript in preparation*

Rohde, D. L. T. (1999) LENS: The light, efficient network simulator. CMU-CS-99-164. Carnegie Mellon University School of Computer Science.

Rohde, D. L. T., Gonnerman, L., & Plaut, D. C. (2006). An improved model of semantic similarity based on lexical co-occurrence. *Manuscript under review*.