Learning to Balance a Beam: The Effect of Instability

David Pfeiffer (pfeiffdd@ucmail.uc.edu)
Department of Psychology, University of Cincinnati
Cincinnati, OH 45221-0376, USA

Heidi Kloos (heidi.kloos@uc.edu)
Department of Psychology, University of Cincinnati
Cincinnati, OH 45221-0376, USA

Daniel P. Bullard (bullardp@mail.uc.edu)
School of Education, University of Cincinnati
Cincinnati OH 45221-0022, USA

Abstract

When presented with a problem-solving task, children sometimes fail to discern relevant pieces of information. Instead, they base their judgment on irrelevant information, sometimes ignoring corrective feedback. How could experience highlight relevant information? Using insights from complexity science, the current paper tests the usefulness of adding instability, or noise, to a child's experience. The idea is that an appropriate amount of instability flattens the attractor space of mistaken performance, allowing children to explore aspects of the environment perhaps considered irrelevant. To test this idea, we asked children between 4 and 9 years of age to place beams on a fulcrum where they would balance. Instability was conceptualized using beams for which the weight distribution was difficult to discern. While 4- to 5-year-olds and adults were unaffected by the noise manipulation, possibly for different reasons, 7- to 9-year-olds balanced the beams better when instability was at an intermediate level.

Keywords: cognitive development; science learning; torque;

Introduction

How does a child know what to pay attention to? For example, how does a child know that a new word pertains to the overall shape of an object rather than the color of it, or the shape of some part of it? Or how does a child know that two events are causally related? Answers to such questions fall in one of two categories: (1) they assume that specific top-down knowledge allows the child to sort messy variation into relevant and irrelevant information[1], or (2) they assume that general attentional processes allow the child to decipher rich statistical data[2]. Both types of approaches—the focus on specific knowledge, or the focus

on general attentional mechanism—have generated a large number of empirical findings, often pitting postulated knowledge against statistical contingencies (e.g., Kuhl, 2000). However, neither of them can fully account for the existing data, as we explain next.

Knowledge-*rich* approaches—those that attribute meaningful performance to mental structures of some sort—fall short when performance changes as a function of miniscule changes in the task context. For example, how could a knowledge-based account explain that the A-not-B error disappears when the infant's posture is briefly changed prior to a B trial (e.g., Smith & Thelen, 2003)? Or how could it explain that toddlers pay attention to the solidity of objects when the object is fully visible, but not when it is hidden (e.g., Keen, 2003)? Of course one could always postulate a new representation, schema, or belief for each unique performance. Children might have implicit knowledge of solidity, for example, but lack explicit knowledge of this concept. However, this solution is neither parsimonious nor predictive and testable.

On the other hand, knowledge-*lean* approaches—those that attribute meaningful performance to the mind's ability to extract statistical regularities—fall short in cases in which children blatantly ignore even the most straight-forward evidence. In particular, children appear blind to statistical contingencies when they have an expectation that conflicts with the statistical contingencies. For example, both children and adults sometimes mistakenly expect two variables to be correlated, leading to the phenomenon of illusory correlations (e.g., Chapman & Chapman, 1969). Or children fail to learn how two variables correlate if that correlation does not match with their larger frame of references (e.g., Kloos, 2007).

The current approach takes a step back and acknowledges that performance is based on both a child's specific expectations and a child's general ability to track statistical contingencies. In fact, there might be much more that matters in the immediate task context, including a child's idiosyncratic history and readiness, details of the verbal

---

[1] Example references include Gopnik et al., 2001; Gelman & Markman, 1986; Carpenter, Call & Tomasello, 2005; Murphy & Medin, 1985; Song, Onishi, Baillargeon & Fisher, 2008; Hamlin, Newman & Wynn, 2009.

[2] Example references include Sloutsky, 2003; Mix, Huttenlocher & Levine, 2002; Smith & Samuelson, 2006; Gómez, 2002; Saffran, Aslin & Newport, 1996.

instructions, and apparently irrelevant arrangements of the task materials. One could conceive of these factors as aligning themselves in a ratio of constraints that make up the attractor space of performance (cf., e.g., Spivey & Dale, 2004). In this terminology, consistent performance—such as doing the same thing even as the external circumstances change—reflects a deep attractor well. And inconsistent performance, even when the task context stays the same, reflects several very shallow attractor spaces that are easily traded for each other.

There is one main advantage of the attractor approach: questions about the relative contribution of an individual factor—say, the contribution of a specific representation, belief, or strategy—become superfluous. If all of the relevant factors align themselves in an interdependent network of constraints, the unique contribution of a single factor cannot be determined empirically (cf., Guastello & Guastello, 1998). Postulating an attractor approach allows us instead to investigate the degree to which changes to the attractor space lead to a successful change in performance. We explore one specific perturbation to the attractor space: namely, to flatten the attractor space by adding noise into task context.

Noise is defined as any form of inconsistency or variation in seemingly irrelevant features of the task. Intuitively, such noise is considered to have negative effects on learning: it might hinder a child's ability to pick up on relevant variations. However, there are some findings that demonstrate the benefit of noise (e.g., Gentner, Loewenstein & Hung, 2007; Green Hall & Magill, 1995). Most notable is a recent finding with a gear task (Stephen, Dixon, & Isenhower, in press). Adults and children had to predict the turning direction of the last gear in a chain, upon knowing the turning direction of the first gear in that chain. Children and adults alike arrived at a more sophisticated solution of the task quicker when the task context featured more noise. In particular, when the chain was moved randomly across the computer screen (vs. standing still), participants more readily switched from tracing each gear individually to paying attention to the parity of the gears. The argument is that the noise, and the resulting instability in the participant-task system, might make available otherwise unnoticed contingencies.

The goal of the current study is to extend this rather non-intuitive finding to a new task domain and a new noise manipulation. Children and adults were asked to balance beams on a fulcrum such that the beam would not tip. This task has rich statistical contingencies (e.g., moment-to-moment haptic feedback) and, at the same time, it is likely to elicit children's prior beliefs (e.g., that beams balance at their geometric middle, and not at its center of mass; Karmiloff-Smith & Inhelder, 1975; Bonawitz, 2007). As such, this task is likely to be affected by both top-down knowledge and by bottom-up attention to statistical structure. Could noise help children attend to relevant information?

Previous research involving balance beams has explored understanding of the weight-distance relationship by manipulating both variables (e.g., Amsel et al., 1996; Ferretti et al., 1985; Siegler, 1976). However, as a means of noise manipulation, we used beams of uniform length yet with difficult-to-perceive weight distribution. Initial calibration testing revealed that preschoolers (N = 64 4- to 5-year-olds) performed at chance when judging the weight distribution of beams that balanced 3cm off the geometric middle. On the other hand, their judgment was above chance for beams that balanced at the geometric middle, as well as for beams that balanced 6cm off the geometric middle. Attempting to balance the 3cm-beams is therefore likely to increase the noise in the system. Note that this noise manipulation differs in important ways from the noise manipulation in Stephens et al.'s gear task: rather than being extraneous to the task (as was the case when the gear chains moved across the screen), our noise manipulation pertains to the relevant task contingencies per se (the relation between perceived weight distribution and correct balancing point). To what extent does such noise manipulation help learning?

There were three conditions that differed in the beams presented to participants during the initial balancing trials. In the *all-noise condition*, all uneven beams balanced 3cm off their geometric middle. In contrast, in the *no-noise condition*, all uneven beams balanced 6cm off their geometric middle. Finally in the *some-noise condition*, half of the uneven beams balanced 3cm off their geometric middle, while the other uneven beams balanced 6cm off their geometric middle. The crucial test was during the last session of trials, which included 3cm-beams and 6cm-beams for all participants. Note that participants in the some-noise condition were given less experience with 6cm-beams than participants in the no-noise condition. This allowed us to pit noise against relevant experience.

To determine whether our manipulation works differently across development, two groups of children were included: preschoolers between 4 and 5 years of age, and elementary-school children between 7 and 9 years of age. Adults served as control group to gauge the endpoint of development. While preschoolers were previously found to have a strong geometric-center belief, 7- to 9-year-olds are more likely to ignore the geometric center and focus instead on the center of mass (cf. Karmiloff-Smith & Inhelder, 1975). Does this difference in balancing success make a difference in how children learn best?

Method

Participants

Participants were 64 4- to 5-year-old preschoolers (range = 4.5 to 5.5, M = 5.0 years, 28 girls, 37 boys) and 61 7- to 9-year-old elementary-school children (range = 7.1 to 9.9, M = 8.6 years, 34 girls, 27 boys[3]). They were recruited from

---

[3] Pilot data from this age group revealed no developmental change in balancing performance, $r(12)$ = -0.09, p < 0.47.

urban and suburban daycare centers and elementary schools serving racially diverse working- to middle-class families. One 4-year-old was tested but excluded from the final sample due to lost interest. In addition, adult data were collected from 68 undergraduate students (M = 19.9, 49 women, 19 men), in exchange for course credit. Children and adults participated in one of the three experience conditions (all-noise, some-noise, no-noise), yielding approximately equal number of participants and equal gender distribution in each condition.

Materials

Twelve balance beams were constructed, each consisting of two wooden containers attached to the opposite ends of a rectangular piece of plywood. The containers were approximately 10 cm long, 4.5 cm high and 6 cm wide, and they weighed between 100 and 450 g[4]. The plywood piece supporting two containers weighed about 100 g, and was 30 cm long and 10 cm wide. The fulcrum consisted of a metal rod (2.5 cm wide, 2.5 cm tall and 56 cm long) attached to wooden platform (61 cm long, 7.5 cm wide). Figure 1 shows three such beams, balanced on the fulcrum, each with a different balancing point. The *0cm-beams* balanced at their geometrical center, and thus had an even weight distribution; the *3cm-beams* balanced 3 cm away from the beam's geometrical center, and the *6cm-beams* balanced 6 cm away from the geometrical center. There were three different versions for each beam type: a light version (173g, 407g, and 436g, respectively), a medium version (387g, 431g, and 530g, respectively), and a heavy version (651g, 757g, and 825g, respectively). All beams were painted uniformly white, with no visual indication about difference in weight distributions. The desired mass of a container was achieved by gluing lead weights and Styrofoam inside the containers.
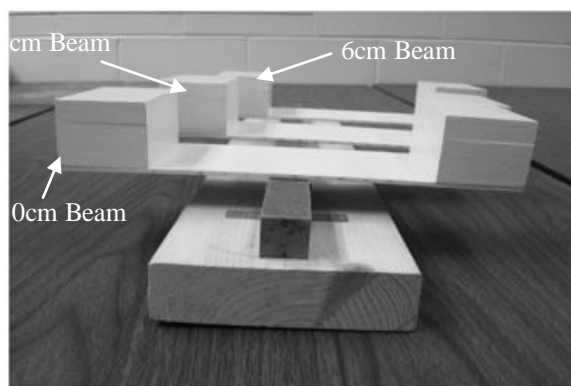

Figure 1: Three types of beams balanced on a fulcrum; 0cm, 3cm or 6cm away from the geometric center.

[4] The desired mass of a container was achieved by gluing lead weights and Styrofoam inside the containers.

Design and Procedure

Children and adults were tested individually, either in a quiet area of their school or in the lab. The experimenter sat across from the participant, with the fulcrum placed lengthwise between them. In an effort to maintain participants' interest in the task, a cover story was used, conveyed through a series of slides displayed on a Dell laptop computer. Participants were told:

> We're going to play a game with Penny the Poodle. You think you can help Penny the Poodle balance stuff? She brought some silly-looking things with her, and they are hard to balance. But here's the thing, Penny the Poodle does not like it when the beams fall down. Do you think you can help Penny the Poodle balance the beams so that they don't fall? I want you to pick it up and feel it and think first, and then put it up here where you think it won't fall.

The experimenter then demonstrated how to lift a beam and orient it properly, without balancing the beam on the fulcrum. Figure 2 shows a schematic of how a beam was supposed to be held by the participant. Testing trials started immediately. Superlab® software (Version 2.0) was used to determine the order of beams and record participants' responses. For each trial, the experimenter first handed the beam to the participant. In the case of uneven beams, care was taken to counterbalance the orientation of the beam (such that the heavier side was sometimes in the right and sometimes in the left hand of the participant). The participant was first asked to wield the beam with two hands to determine if one side of the beam was heavier, and, if so, which side was heavier. Without providing any feedback, the participant was then asked to balance the beam on the fulcrum in such a way that it would not fall. The experimenter recorded the initial placement of the beam using small markings drawn surreptitiously along the height of the plywood plank. Once the beam was placed on the fulcrum, the participant was allowed to slide the beam from left to right until the beam balanced. After maximally two minutes, a new trial started (i.e., if the participant could not balance the beam within two minutes, the trial was terminated and the child was encouraged to "try another beam").
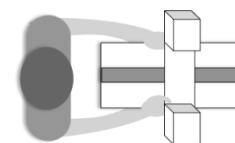

Figure 2: Schematic view-from-above of how children were asked to balance a beam.

There were six consecutive sessions, with the last session being identical across conditions: This last session was designed to test the effect of experience gained during the first five sessions (it included three 3cm-beams and three

6cm-beams). In the *all-noise condition*, each of the first five sessions included two 3cm-beams and two 0cm-beams, resulting in ten beams of each type across the five sessions. Conversely, in the *no-noise condition*, each session included two 6cm-beams and two 0cm-beams (again resulting in ten beams for each type across the five sessions). Finally, in the *some-noise condition*, each session included one 3cm-beam, one 6cm-beam, and two 0cm-beams per session, resulting in ten even beams and five of each of the asymmetric beams across the five sessions. Note that the total number of beams during the five sessions stayed the same in all three conditions, with half of the beams being even, and with half of the beams being uneven. The only difference was whether the noisy 3cm-beam was included or not. Order of beams within a session was randomized, and each session was followed by a short break during which participants were told that they reached the next level in the activity.

Results

Each trial yielded two scores, one to reflect whether the mass distribution of a beam was judged correctly (i.e., *weight judgment*), and one to reflect the initial placement of the beam on the fulcrum (i.e., *beam placement*). For weight judgments, responses were coded according to whether the two sides of the beam were judged as equal in weight (correct for even beams) or not equal (correct for uneven beams). Judgment data was missing for 12 out of the 5,018 trials across participants.

For beam placement, performance was coded in terms of the absolute distance between actual placement of the beam and correct balancing point. Placing even beams at the geometric center yields a placement error of 0cm, and placing uneven beams at the geometric center yields a placement error of 3cm and 6cm, respectively. Placement data was missing on 14 of the 5,018 trials. Note that placing uneven beams on the lighter side of the beam would yield a very large error. This latter beam placement never occurred for the 6cm-beams, and it occurred seldom for the 3cm-beam. Thus, using the absolute error is likely to inflate only the error for even beams (when error is expected to be low to start with), but not the error for uneven beams.

We first present a series of preliminary analyses (1) to confirm that the weight distribution of 3cm-beams was difficult to detect and (2) to uncover possible learning effects during the five initial sessions. We then turn to the main analysis of describing performance in the sixth and last session, to assess the effect of our noise manipulation.

Preliminary Analyses

The first preliminary analysis pertains to whether the weight judgment for 3cm-beams was lower than the weight judgments of the other two beam types (0cm- and 6cm-beam). Figure 3 shows the mean proportion correct weight judgments for each beam type, separated by age group. As expected, judging the weight distribution of 3cm-beams was more difficult than of the other two beam types. Repeated-measure ANOVAs, one for each age group, revealed large quadratic effects of beam type, $Fs > 40.9$, $ps < .001$. Note that mean judgments of 3cm-beams did not exceed chance for preschoolers, $t < 1.1$ (all other mean judgments were above chance, $ts > 4.3$). And while the difference in adults' judgment success between 0cm- and 3cm-beams was marginally significant, $t = 1.95$, $p < .08$, all other pair-wise difference to the 3cm-beams were significant.
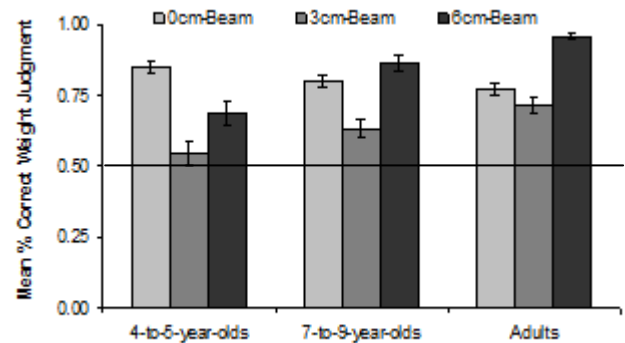

Figure 3: Mean proportion of correct weight judgment, separated by beam type and age group. Error bars reflect standard errors.

Breaking it down by session, weight judgment of 3cm-beams did not improve from the first to the last session. For example, when comparing the combined success rate of weight judgment in the first two sessions with that of the combined fourth and fifth session, performance for each age group and condition remained unchanged, $ts < 1$. Thus, the weight distribution of 3cm-beams was indeed more difficult to judge than the weight distribution of the other beams—giving credibility of our noise manipulation.

How did placement performance change across the first five initial sessions? Recall that all children had the same number of even beams during the first five sessions of the experiment. Considering only the placement error for even beams, a session-by-age ANOVA revealed no significant main effects of session or age, nor a significant session by age interaction. Across sessions, mean placement error (in cm) was very low, with $M = 0.46$cm for 4- to 5-year-olds, $M = 0.32$cm for 7- to 9-year-olds, and $M = 0.49$cm for adults. For uneven beams (3cm- and 6cm-beams), recall that conditions differed in the kinds of beams presented during the initial five sessions. There were no 6cm-beams in the first five sessions of the all-noise condition, and there were no 3cm-beams in the first five sessions of the no-noises condition Figure 4 shows the mean placement error of these beams across the first five sessions. A series of trend analyses revealed that learning (i.e., improved placement across sessions) took place only for the 6cm-beams placed by 7- to 9-year-old and adults, $F(1,38) > 24.0$, $ps < .001$.

Note that 4- to 5-year-olds had higher overall placement errors than 7- to 9-year-old children, who in turn had higher overall placement errors than adults. Finally, condition did not have a noticeable effect on placement performance

during the initial five sessions: For example, whether performing in the all-noise condition or the some-noise condition, 4- to 5-year-olds placed most of the 3cm-beams between two and three cm off the correct balancing point.
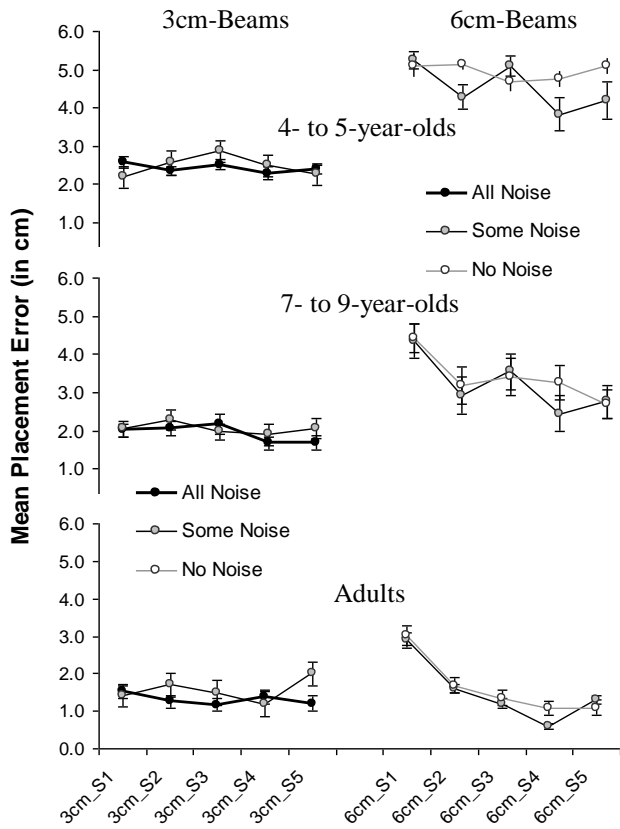

Figure 4: Placement performance for each age group of uneven beams across the five training sessions.

Main Analysis

For the main analysis, we looked at participants' placement of the beams during the last testing session: when the task was to balance three 3cm-beams and three 6cm- beams. The main question pertains to whether our noise manipulation affected the placement error of 6cm-beams. Figure 5 shows the mean absolute error for this beam type during the last session.

One-factor ANOVAs, one for each age group, revealed a quadratic effect of condition for 7- to 9-year-olds, $F(2, 58) = 3.2$, p < .05 (see red circle in Fig. 5). Children in the some-noise condition had smaller placement errors ($M = 2.3cm$) than children in the no-noise condition ($M = 3.1cm$) or children in the all-noise condition ($M = 3.6cm$). In fact, whether children in this age group were exposed to no noise or only noise, placement errors were equally high. The same effect of condition was not observed for 4- to 5-year-olds (their mean error was uniformly high, independent of condition) not was it found for adults (their mean error was uniformly low). These findings show that noise, in the

proper proportion and context, can indeed contribute to improved performance.
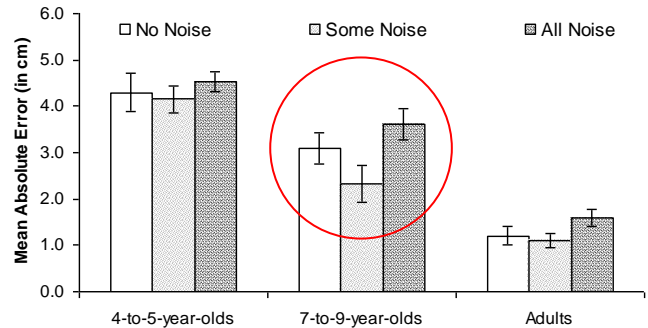

Figure 5: Mean absolute placement error of 6cm-beams in the last session, separated by condition and age group. Error bars show standard errors.

Discussion

Our goal was to determine whether experience infused with some instability (or noise) leads to learning. The task was to balance beams on a fulcrum, and the crucial manipulation was whether participants were given experience with 'noisy' beams, those which had a difficult-to-perceive weight distribution (3cm-beams). Our findings show the predicted effect for 7- to 9-year-olds. In this age group, children who were exposed to some of the noisy beams made fewer errors balancing the 6cm-beams in the final testing session than children who were either presented with only 6cm-beams (no-noise condition) or only 3cm-beams (all-noise condition). In fact, final balancing performance of 6cm-beam in these latter conditions did not differ from each other. This suggests that placement error of 6cm-beams was high, whether children had a lot of experience with 6cm-beams or none at all. Placement error only decreased when children were exposed to some 3cm-beams.

Note that participants started out with a similar degree of placement errors, independently of condition. Second, there was no effect of condition on children's placement of 3cm-beams during the final session of trials. Therefore, the effect of condition on the placement of 6cm-beams in the final session of trials cannot be attributed to differences in initial competence or differences in overall improvement. Instead, results suggest that the moderate amount of noise, introduced by the 3cm-beams, highlighted relevant information in the task. Put another way, the moderate amount of noise might have flattened the attractor of the belief that visually symmetrical beams balance at their geometric center, allowing children to explore the less salient haptic information of weight distribution.

Why did our manipulation fail to affect preschoolers and adults? For adults, the issue might be a ceiling/floor effect: balancing performance of 6cm-beams might have been too

successful to pick up on a possible effect of condition. The steep learning curve during the initial five sessions supports this conclusion. Furthermore, adults were quite successful at judging the weight distribution of 3cm-beams—possibly reducing the degree of perceived noise. In contrast, preschoolers had considerable difficulty judging the weight distribution of 3cm-beams: their performance was a chance. This might have introduced too much noise in their experience, reducing the potential benefits of noise in the learning context. These speculations support the idea that the benefits of noise might follow a U-shaped trajectory: too much noise might hurt learning, as does too little noise—at least when learning includes overcoming misconceptions.

Acknowledgments

References

Amsel, E., Goodman, G. Savoie, D., & Clark, M. (1996). The Development of Reasoning about Causal and Noncausal Influences on Levers. *Child Development*, *67*(4), 1624-1646.

Bonawitz, E.B. & Schulz, L. (2008). Why Learning Can Be Hard: Preschooler's Causal Inferences. *Symposium on Naturally-Inspired Artificial Intelligence, Proceedings of AAAI*.

Carpenter, M., Call, J. & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science, 8*(1), F13-F20.

Chapman, L.J. & Chapman, J.P. (1969). Illusory Correlation as an Obstacle to the Use of Valid Psychodiagnostic Signs. *Journal of Abnormal Psychology, 74*(3), 271-280.

Ferretti, R.P., Butterfield, E.C., Cahn, A. & Kerkman, D. (1985). The classification of children's knowledge: development on the balance-scale and inclined-plane tasks. *Journal of Experimental Child Psychology, 39*(1), 131-160.

Gelman, S. A. & Markman, E.M. (1986). Categories and induction in young children. *Cognition, 23*(3), 183-209.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*. 285-307.

Gómez, R. L. (2002). Variability and Detection of Invariant Structure. *Psychological Science, 13*(5), 431-436.

Gopnik, A., Sobel, D. M., Schulz, L.E. & Glymour, C. (2001). Causal Learning Mechanisms in Very Young Children: Two-, Three-, and Four-Year-Olds Infer Causal Relations From Patterns of Variation and Covariation. *Developmental Psychology, 37*(5), 620-629.

Green Hall, K. & Magill, R.A. (1995). Variability of Practice and Contextual Interference in Motor Skill Learning. *Journal of Motor Behavior, 27*(4), 299-309.

Guastello, S. J. & Guastello, D.D. (1998). Origins of Coordination and Team Effectiveness: A Perspective From Game Theory and Nonlinear Dynamics. *Journal of Applied Psychology, 83*(3), 423-437.

Hamlin, J. K., Newman, G. E. & Wynn, K. (2009). Eight-Month-Old Infants Infer Unfulfilled Goals, Despite Ambiguous Physical Evidence. *Infancy, 14*(5), 579-590.

Karmiloff-Smith, A., & Inhelder, B. (1975). If you want to get ahead, get a theory. *Cognition, 3*, 195-212.

Keen, R. (2003). Representation of Objects and Events: Why Do Infants Look So Smart and Toddlers Look So Dumb? *Current Directions in Psychological Science, 12*(3), 79-83.

Kloos, H. (2007). Interlinking physical beliefs: Children's bias towards logical congruence. *Cognition, 103*(2), 227-252.

Kuhl, P. K. (2000). A New View of Language Acquisition. *Proceedings of the National Academy of Sciences of the United States of America, 97*(22), 11850-11857.

Mix, K. S., Huttenlocher, J. & Levine, S.C. (2002). Multiple Cues for Quantification in Infancy: Is Number One of Them? *Psychological Bulletin, 128*(2), 278-294.

Murphy, G. L. & Medin, D. L. (1985). The Role of Theories in Conceptual Coherence. *Psychological Review, 92*(3), 289-316.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical Learning by 8-month-old Infants. Science, 274, 1926-1928.

Siegler, R. S., (1976). Three Aspects of Cognitive Development. *Cognitive Science, 8*(4), 481-520.

Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences, 7*(6), 246-251.

Smith, L. B. & Samuelson, L. (2006). An Attentional Learning Account of the Shape Bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology, 42*(6), 1339-1343.

Smith, L. B. & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7*(8), 343–348.

Song, H. J., Onishi, K. H., Baillargeon, R. & Fisher, C. (2008). Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition, 109*(3), 295-315.

Spivey, M. J. & Dale, R. (2004). On the Continuity of Mind: Toward a Dynamical Account of Cognition. *Psychology of Learning and Motivation, 45*, 87-142.

Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (in press). Dynamics of representational change: Entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance*.