UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Causal Reasoning in Infants and Adults: Revisiting backwards-blocking

Permalink

Journal

Authors

Benton, Deon T.
Rakison, David H.

Publication Date

2016

Peer reviewed

Causal Reasoning in Infants and Adults: Revisiting backwards-blocking

Deon T. Benton (dtbenton@andrew.cmu.edu)
Department of Psychology, 1098 Morewood Avenue
Pittsburgh, PA 15232 USA

David H. Rakison (rakison@andrew.cmu.edu)
Department of Psychology, 1098 Morewood Avenue
Pittsburgh, PA 15232 USA

Abstract

Causal learning is a fundamental ability that enables human reasoners to learn about the complex interactions in the world around them. The available evidence with children and adults, however, suggests that the mechanism or set of mechanisms that underpins causal perception and causal reasoning are not well understood; that is, it is unclear whether causal perception and causal reasoning are underpinned by a Bayesian mechanism, associative mechanism, or both. It has been suggested that a Bayesian mechanism, rather than an associative mechanism, underpins causal reasoning because such a mechanism can better explain the putative backward-blocking finding in children and adults (e.g., Sobel, Tenenbaum, & Gopnik, 2004). In this paper, we report two experiments to examine to what extent infants and adults exhibit backward blocking and whether humans' ability to reason about causal events is underpinned by an associative mechanism, a Bayesian mechanism, or both.

Keywords: causality; infants; adults; causal reasoning; causal learning; causal perception; infant and child development

Introduction

The emergence of causal perception

Of the skills that enable humans to understand the events they experience in the world, perhaps few are as important as the ability to learn about causality in the world. This is a key cognitive ability that enables infants, children, and adults to perceive and encode cause-and-effect relationships and reason about the effects of interventions on those relationships. However, before humans can reason explicitly about causal events through interventions, they must be able first as infants to perceive simple cause-and-effect relations between objects in the world. The ability to perceive cause-and-effect relations has been studied most extensively using Michottian launching events. The reason Michottian launching events have been used in particular is because they are among the simplest events in which to observe the causal relationship between two objects.

In studies that use these sequences, infants typically are habituated to and tested on either a direct-launching, a delayed-launching, launching-without-collision, or a delayed-plus-no-collision event (Figure 1). In the direct-launching event, one object travels across a stage, ostensibly makes contact with a second, stationary object located mid-screen, at which point the first object stops moving and the second object begins to move. The delayed-launching event is similar to the direct-launching event except the second objects moves after a brief delay upon contact from the first object. The launching-without-collision event is also similar to the direct-launching event except that a small spatial gap is inserted between when the first object stops moving and when the second object begins to move. Finally, in the delayed-plus-no-collision event, the second object begins to move after a brief delay and in the presence of a spatial gap.
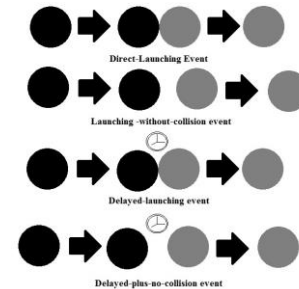

Figure 1: Launching events

Research that has employed launching-event sequences has established that causal perception emerges in development between 4½ and 10 months of age (e.g., Cohen et al., 1998; Rakison & Krogh, 2012). In a classic study on infants' ability to perceive causality in launching-event sequences, for example, Leslie and Keeble (1987) found that 6½-month-old infants who were habituated to a direct-launching sequence—in which a red cube caused a green cube to move through contact—dishabituated to the reversal of the event—in which the green cube now caused the red cube to move. Subsequent work by Cohen and Amsel (1998) showed that causal perception undergoes a developmental transition, whereby 4- and 5½-month-olds responded to the continuity of motion and spatiotemporal relations between the objects in the events, whereas 6¼-month-olds responded on the basis of causality. In a separate study, Schlottman and Surian (1999) found that 9-month-olds will perceive causality in launching events that incorporate a gap only when the objects involved in those events are imbued with animacy cues. Finally, Oakes and Cohen (1990) found that 10-month-olds, but not 6½-month-olds, responded on the basis of causality when realistic stimuli were used instead of simple geometric figures, and Rakison and Krogh (2012) found that 4½-month-olds showed evidence of causal perception when provided with real-world causal-action experience using Velcro-covered

mittens. Taken together, this research suggests that causal perception emerges between 4½ and 10 months of age.

The emergence of causal reasoning

In contrast to causal perception, the ability to reason about causal events emerges later in development. Although developmental researchers have used a variety of tasks to study causal reasoning (for a review see Bullock, Gelman, & Baillargeon, 1982), we restrict our focus to research that has used the blicket-detector design. This is because (1) the blicket-detector task has been used most frequently to study causal reasoning in young children, and (2) an ultimate aim of ours is to examine to what extent the abilities to perceive causality in launching-event sequences and distinguish blickets from non-blickets in a blicket-detector task are underpinned by the same or different mechanisms (e.g., Gopnik et al., 2004; McClelland & Thompson, 2007).

In a typical blicket-detector study, children are introduced to a machine called the "blicket detector" and told that the machine lights up and plays music only when objects labeled blickets are placed on its surface. Children are then asked to determine which objects are blickets and which are not and to intervene to make the machine go. Research that has used this design has shown that 3- to 5-year-olds can make causal inferences with blicket-like objects that span the biological and psychological domains (Schulz & Gopnik, 2004) and that 18- to 30-month-olds can use higher-order relations between objects to make causal inferences (Walker & Gopnik, 2014). Of these findings, perhaps the most relevant from the perspective of the present experiments is the finding that children 2 years and older can use screening-off and backward-blocking reasoning (hereafter BB) to make inferences and generate interventions about the causal status of blicket objects (e.g., Gopnik et al., 2001; Sobel et al., 2004). For example, Gopnik et al. (2001) showed that when 2-, 3-, and 4-year-olds were shown an indirect screening-off (hereafter IS) event—in which together two objects, objects A and B, caused the detector to activate (i.e., AB+) and then an event in which object A alone fails to activate the detector (i.e., A-)—they categorized only object B as the cause. This ability to use IS reasoning—which children ostensibly share with adults—is a hallmark of causal reasoning that enables human learners to distinguish objects associated with an effect from those that produce an effect.

In addition, previous research showed that children use BB reasoning and base rates to reason about blickets. For example, Sobel et al. (2004) found that 4-year-old children who are shown a BB event—in which together objects A and B produce the effect (i.e., AB+) and then an event in which object A produces the effect (i.e., A+)—are less likely to categorize object B as a blicket compared to same-age children who are shown the IS test event if blickets are rare than common. Together, the BB and IS findings are important abilities because it has been suggested that contemporary associative models such as the Rescorla-Wagner model (henceforth, the RW model) fail naturally to

account for base rates and why object B is not treated equivalently across the BB and IS conditions. Indeed, in terms of the RW model (Rescorla & Wagner, 1972), object B should be treated equivalently across both the IS and the BB trials because the associative strength between B and the activation of the detector is the same in both cases. Given the failure of contemporary associative models to account for the BB finding, some researchers have proposed that causal learning is underpinned by a Bayesian-inference mechanism (discussed below). Putting this debate aside, the research on causal reasoning suggests that it emerges between 18 months (cf., Sobel & Kirkham, 2006) and 4 years of age.

Development gap and theoretical debates

Despite extensive work on causal perception and causal reasoning, little is known about the relation between these two abilities. For instance, it is unclear whether causal perception and causal reasoning are underpinned by the same or different mechanisms. That is, is causal perception underpinned by one mechanism and causal reasoning, another? Or, is causal perception and causal reasoning underpinned by the same mechanism? According to proponents of the domain-general view of causal learning—where the same all-purpose mechanisms govern learning—early causal perception and later causal reasoning are (1) the emergent consequences of continually enriching perceptual and cognitive systems and (2) are abilities that are underpinned by an associative learning mechanism. This position garners support from behavioral and computational research that showed that infants, young children, and computational (PDP) models use the correlational structure and the predictive statistics of causal events to process and encode their causal relations and that this ability develops over time (e.g., Cohen et al., 1998; McClelland & Thompson, 2007).

In contrast, according to proponents of the domain-specific position—where specific mechanisms process specific kinds of inputs—humans possess specialized modules or mechanisms that are designed specifically to process causal events (Leslie, 1995; Gopnik et al., 2001). For example, within the domain-specific position, some have posited that humans use a simple form of Bayes' rule to reason about the conditional probabilities in causal events and that this ability may be present from birth or shortly thereafter (e.g., Gopnik & Wellman, 2012). This argument is ostensibly supported by research by Sobel and Kirkham (2006, 2007) that showed (1) that infants 5 months of age and older use IS and BB reasoning in a modified habituation version of the blicket detector study, and (2) by research by Griffiths et al. (2011) that showed that adults engage in BB reasoning to reason about super pencils. However, these findings should be considered cautiously because (1) the evidence was mixed about whether infants in Sobel and Kirkham (2006, 2007) processed the events associatively or based on BB and IS reasoning, (2) the habituation task itself failed to preserve the conditional probabilities of previous

blicket studies, and (3) BB reasoning was observed only in one condition of four in Griffiths et al. (2011).

Based on the above limitations, a primary goal of the current set of experiments was to examine causal reasoning in infants and adults and examine to what extent 6-month-olds (Experiment 1) and adults (Experiment 2) use BB and IS reasoning to process causal events. We chose 6-month-olds because this is the age at which they begin to process launching-events causally. Experiment 1 used a novel task design—where the conditional probabilities between this and previous blicket-detector studies were identical—in which infants were habituated to an AB+ A+ event and then tested on A+, B+, A-, B- events (Figure 2). This design choice was important in two ways: (1) it bore closer resemblance to, and allowed direct comparisons with, the conditions presented in the blicket detector studies with children and (2) enabled us to test whether infants were processing the events associatively—in line with predictions made by the RW model—or according to BB reasoning (Figure 2). These predictions derive from previous research that examines infants' use of Bayesian inference (Sobel & Kirkham, 2006, 2007). Note that although the RW model has been ruled out as an informative model of causal reasoning in adults, it is still possible that it can predict the performance of young infants in a habituation task in which there are multiple trials. The specific aims of Experiment 2 were (1) to replicate with adults the conditions presented to children in previous blicket studies and (2) to examine to what extent adults engage in BB and IS reasoning. An additional aim of Experiment 2 was to explore whether adults process the causal events along a causal gradient (see the Conclusion section) in which some adults might process the events associatively, whereas others might process the events according to BB reasoning. An important strength of Experiment 2 was that the design enabled us to assess to what extent adults engage in BB reasoning by comparing pre- and post-ratings of B. With the exception of one study (Griffiths et al., 2011), most compare children's ratings of B in the BB and IS conditions; and the one study that conducted pre- and post-comparisons, a drop in the rating of B was observed in only 1 of 4 conditions.

Experiment 1

Methods

Subjects. Nineteen 6-month-old ($M = 6$ months; range: 5 months 15 days to 6 months 24 days) infants participated in the experiment.

Stimuli and Design. The habituation and test stimuli were computer-animated events that were presented on a computer-generated stage (Figure 2). In each of the two habituation movies, the red and blue circles entered the stage from the right and left (counterbalanced) and moved horizontally across the stage until they abutted a square that was located mid-screen at which point a sun appeared from the square. The second habituation movie was identical to the first except that only one of the two objects (object A)

moved horizontally across the stage. Following the habituation phase, infants were shown 4 test events (Figure 3).

Procedure. Each infant sat on their caretaker's lap facing the television monitor. The parent was instructed to abstain from any form of communication with the infant and to remain neutral to avoid biasing the infants' natural response to the habituation and test events. The caregiver was also naïve to the hypotheses and predictions of the experiments to eliminate the chance that the caregiver could reliably influence the infant's looking behavior during the experiment.

	Associative (RW model)	Bayes (BB reasoning)
A-	Longer looking	Longer looking
B+	Equal looking	Longer looking
B-	Longer looking	Equal looking

Figure 2: Looking-time predictions to the A-, B+, and B- test events relative to the A+ habituation event.

During the habituation phase of the experiment, infants were presented with the two habituation events, as outlined in the previous section. In other words, infants were presented with the AB+ event, in which objects A and B caused the sun to appear from the box, and the A+ event, in which object A singly caused the sun to appear from the box. Whether the object (A or B) entered the stage from the right or from the left was counterbalanced across infants.
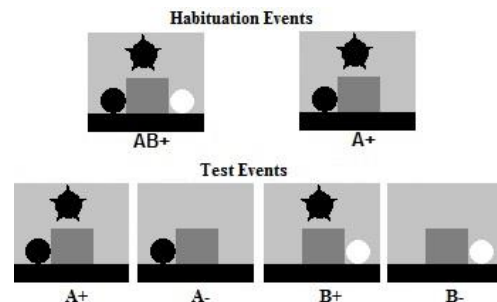

Figure 3: The habituation and test movies

When the habituation criterion was reached or when 16 trials had been presented, the four-trial test phase began. Infants were excluded from the final analyses if they did not reach the criterion within 16 trials. Two infants were excluded for this reason. The test events (i.e., A+, A-, B+, B-) were presented using a Latin square to ensure that event presentation was counterbalanced.

Results

The first analyses compared $\log_{10}$ looking times to the last three habituation trials to $\log_{10}$ looking times to the familiar A+ test event. The rationale for this analysis was to examine whether infants had, in fact, habituated to the familiar (A+) habituation event. Infants' fixation times were entered into a one-way between-subjects ANOVA, which revealed that infants' looking times to the familiar (A+) test event ($M =$

0.81 , *SD* = 0.39) did not reliably differ from their looking times to the last three habituation events (*M* = 0.87, *SD* = 0.06), $F(1, 36) = 0.35$, $p = 0.56$. This analysis revealed that infants had indeed habituated to the familiar event (A+).

The second analysis examined whether there was an effect of sex on looking time to the four test events. In particular, We examined whether male or female infants differed in the amount of time they looked at the four test events. Thus, looking times were entered into a 2 (Sex: male vs. female) x 4 (Test type: A+, A-, B+, B-) mixed-design ANOVA, with sex as the between-subjects variable and test type as the within-subjects variable. Neither the main effects for sex, $F(1, 17) = 0.14$, $p = 0.71$, or test type, $F(3, 51) = 0.71$, $p = 0.55$, nor the interaction (sex × test type) was significant, $F(3, 51) = 1.55$, $p = 0.21$. The data for sex was subsequently collapsed.
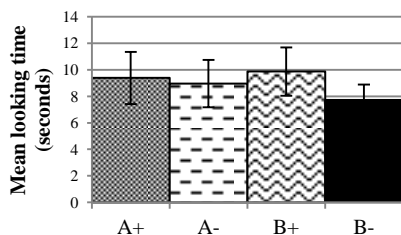

Figure 4: Results from Experiment 1: Infants' mean looking times as a function of test trial.

The primary analysis examined whether infants differed in their looking times to the four test events. The rationale for this analysis was that if infants showed differential looking to the four test events, then it would be possible to examine whether infants were processing the events associatively or according to Bayesian inference (as outlined in the Introduction). A repeated measures ANOVA with test type (A+, A-, B+, B-) as the within-subjects factor revealed that infants did not differ reliably in their looking times to the four test events, $F(3, 54) = 0.28$, $p = 0.84$. This analysis suggests that infants were processing the events neither associatively nor in a way that is consistent with previous BB findings.

Despite the fact that the main analysis failed to yield a significant finding, it is possible that infants processed the events based on associative learning or Bayesian inference. In particular, by comparing looking times to the A+ and the A- test trials for infants who received these two trials first and looking times to the B+ and B- test trials for infants who received these trials first, it is possible to determine whether infants processed the events based on Bayesian inference or associative learning. Thus, separate paired-sample t-tests were used to compare infants' looking times to the A+ and A- test trials and looking times to the B+ and B- test trials. The analysis that compared looking times to the A+ and A- test trials revealed that infants looked equally long at the A+ test trial (M = 15.64 s, SD = 11.84 s) and A- test trial (M = 12.72 s, SD = 5.81 s), $t(4) = 0.44$, $p = 0.68$. The analysis that compared looking time to the B+ and B-

test trials revealed that infants looked equally long at the B+ test trial (M = 9.58 s, SD = 5.45 s) and B- test trial (M = 11.17 s, SD = 7.22 s), $t(5) = 0.39$, $p = 0.72$. Considered together, this set of analyses reveals that infants were not processing the events based on associative learning or Bayesian inference.

Discussion

The null results from Experiment 1 indicated that it is unclear whether infants processed the events associatively or according to previous BB findings; that is, the results neither showed that infants looked longer at the A- and B- test events relative to the A+ and B+ test events as would be predicted from an associative (PDP) model nor did they show that infants looked longer at the A- relative to the remaining three test events as would be predicted from a Bayesian perspective. One possible explanation for this finding is that the events may have been too complex for infants to process compared to those in previous causal-perception studies. Indeed, standard Michottian launching events are typically simpler and involve far less dynamic cues—which can be difficult for 6-month-olds to process, much less according to BB reasoning—than the events used in the present experiment.

Nonetheless, the results from Experiment 1 suggest tentatively that 6-month-olds in the current design cannot solve (or reason) about BB events in associative or Bayesian way despite previous research by Sobel and Kirkham (2006, 2007) that suggests that Bayesian inference is present by at least 5 months (c.f., Shultz, 2007). Despite the fact that the results from Experiment 1 provided inconclusive evidence about whether 6-month-olds used BB reasoning to process the causal events, it is still possible that in the context of a standard blicket-detector study, adults will use BB and IS reasoning to process causal events. This was the goal of Experiment 2.

Experiment 2

Methods

Subjects. Sixty college students were recruited from Carnegie Mellon University to participate in Experiment 2.

Stimuli and Design. A device similar to the blicket detector in previous studies was used in this study. The device was 5" x 7" x 3" and was made of wood (painted black) with a white lucite top. The machine operated via a remote control that was attached to the end of an electric wire that was attached to the side of the machine. When the button was depressed and the object predetermined to be the "blicket" was placed on the surface of the detector, the music and the lights began to play and flash. The button was not pressed, and hence the music and lights did not play or flash, when the object predetermined not to be the blicket was placed on the detector's surface.

Eight cube and cylinder objects, each of different color and approximately 1" in diameter, were used. No objects of

the same shape were used to demonstrate the effect of blicketness and the object that was designated as the blicket was counterbalanced across subjects. Two unrelated objects were used in the pretest phase of the experiment.

Procedure. Participants were tested in a quiet testing room. Participants were introduced to the machine and told that it was called a blicket detector that activated only when blickets were placed on it at the beginning of the experiment. They were instructed also that their job was to determine which objects were blickets and which were not. Participants were then given two pretest trials to ensure that they understood the purpose of the experiment. In one of the pretest trials, one of the two unrelated objects activated the machine and was labeled the blicket, and the other of the two objects (both randomly determined) did not activate the machine.

Following this pretest phase, participants were then given four test trials (counterbalanced). These test trials paralleled those in previous blicket-detector studies and included the one-cause (1C), two-cause (2C), indirect screening-off (IS), and backward-blocking (BB) trials. The 1C and 2C trials served as the controls to ensure that participants understood the test events. Participants were instructed to rate on a scale of 0 (definitely not) – 100 (definitely is) that each object in the pair was a blicket both before and after a trial. In the 1C trial, object A activated the machine when placed alone on the detector but object B did not when placed on the detector alone. Both objects were then placed on the machine twice, which activated. In the 2C trial, object A activated the detector 3 of the 3 times it was placed alone on the detector, whereas object B activated the machine 2 of the 3 times it was placed alone on the detector. In the BB trial, both objects A and B were placed on the detector twice, which activated both times. Object A was then placed on the detector by itself and the detector activated. The IS trial was identical to the backward-blocking trial except that object A did not activate the detector.

Results

To analyze whether adults' ratings of objects A and B differed for the BB, IS, 1C and 2C test trials, a repeated-measures ANOVA with ratings of objects A and B for each test trial as the within-subjects factor revealed that participants' ratings of both objects differed between each test trial, $F(7, 413) = 69.58$, $p < .001$. To examine whether the pre-ratings of A and B differed from the post-ratings of A and B for each test trial, paired samples t tests with Bonferroni corrections were conducted

The first paired-samples t-test for the 1C test trial revealed that the post-rating of A ($M = 94.92$, $SD = 15.58$) increased significantly from the pre-rating of A ($M = 49.62$, $SD = 16.96$), $t(59) = -15.34$, $p < .006$. In contrast, the post-rating of B ($M = 10.33$, $SD = 21.68$) decreased significantly from the pre-rating of B ($M = 50.08$, $SD = 15.66$), $t(59) = 11.03$, $p < .006$. These results replicate the 1C condition in previous blicket-detector studies.

The second paired-samples t-test for the 2C test trial revealed that the post-rating of A ($M = 94.75$, $SD = 11.33$) increased significantly from the pre-rating of A ($M = 50.60$, $SD = 18.25$), $t(59) = -15.78$, $p < .006$. Likewise, post-rating of B ($M = 79.10$, $SD = 19.42$ significantly increased from the pre-rating of B ($M = 48.75$, $SD = 16.46$), $t(59) = -10.79$, $p < .006$. These results also replicate previous 2C results.

The third paired-samples t-test for the IS test trial revealed that participants' post-rating of A ($M = 12.03$, $SD = 27.06$) decreased significantly from their pre-rating of A ($M = 55.77$, $SD = 18.39$), $t(59) = 10.20$, $p < .006$. In contrast, the post-rating of B ($M = 89.48$, $SD = 16.60$) increased significantly from the pre-rating of B ($M = 49.83$, $SD = 17.32$), $t(59) = -12.12$, $p < .006$. This result suggests that adults used IS reasoning.
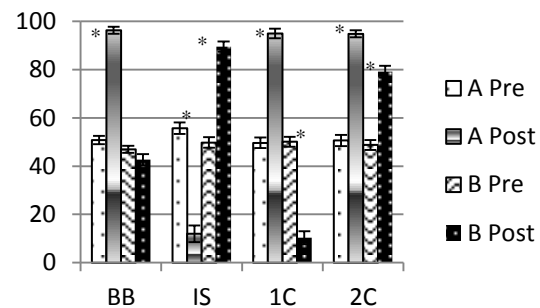

Figure 5: Results of Experiment 2: Participants pre- and post-ratings of objects A and B as a function of condition. Asterisks indicate significance at $p < .05$ between A pre and A post and B pre and B post pairs.

The final paired-samples t-test for the backward-blocking test trial revealed that participants' post-rating of A ($M = 96.25$, $SD = 11.07$) increased significantly from their pre-ratings of A ($M = 50.83$, $SD = 13.38$), $t(59) = -19.73$, $p < .006$. The pre-rating of B ($M = 46.92$, $SD = 11.28$), however, did not differ from the post-rating of B ($M = 42.67$, $SD = 18.14$), $t(59) = 1.84$, $p = .07$, demonstrating the absence of BB.

Given the absence of the BB effect, we conducted an additional analysis to examine whether the effect was moderated by a tendency for some participants to use BB reasoning and others to use associative reasoning. A repeated-measures ANOVA revealed that pre-ratings of B that varied between 0 and 50 ($M = 42$, $SD = 1.87$) did not differ reliably from post-ratings of B ($M = 46.83$, $SD = 3.25$), whereas pre-ratings of B that varied between 50 and 100 ($M = 51.83$, $SD = 1.87$) were significantly higher than post-ratings of B ($M = 38.5$, $SD = 3.25$), $F(1, 58) = 20.52$, $p < .001$. These same participants did not differ in their use of IS reasoning. This provides preliminary (but speculative) evidence that participants processed the causal events along a causal gradient; that is, not all participants showed BB reasoning. This is an important finding because, in contrast to our finding, previous research suggests that BB reasoning is a fundamental human ability.

Discussion

The results from Experiment 2 indicated that participants' post-ratings of objects A and B differed reliably from their pre-ratings of objects A and B in all test trials except for the BB test trial where no significant difference was observed. This BB result is particularly interesting because it contravenes the prediction about BB made by the Bayesian perspective; that is, if participants engaged in BB reasoning, then a significant decrease in the rating of B between the pre- and post-rating phases should have been observed according to this view. However, associative models (e.g., the RW model) make no such prediction about a drop in the rating of B as a blicket. The results of Experiment 2 seemed to support both perspectives in which some adults engaged in BB reasoning whereas others engaged in associative reasoning. In general, the results from Experiment 2 replicated those with children in previous blicket-detector studies.

Conclusions

The null results from Experiment 1 indicated that the 6-month-olds processed causal events neither associatively or in terms of BB. This result suggests that BB reasoning and, to a lesser extent the ability to use a simple form of Bayes' rule as has been assumed, may not be present from birth or shortly thereafter. Instead, this ability may develop as infants learn about causal events in the world. For example, infants 6 months of age and younger may process the features of the objects or the paths that the objects take independently, whereas older infants may process the relations between the objects and begin to parse the events in an associative- or Bayesian-like way. Ongoing research is testing this hypothesis with older infants.

The results from Experiment 2 replicated three of the four test-trials given to children and showed that adults use IS reasoning to reason about blicket events. For the BB trial, however, adults did not rate B differently between the pre- and post-rating phases. Note that we tested the same number of adults in this study as in Griffiths et al. (2011). Nonetheless, there was evidence that participants whose pre-ratings of B varied between 0 and 50 appeared to process the events associatively, whereas participants whose pre-ratings of B varied between 50 and 100 appeared to use BB reasoning. This effect also appeared to be restricted to the BB trial, which suggests that the different modes of reasoning apply only to BB trials. This result suggests that adults may process causal events along a gradient; that is, some adults may process the events associatively, others may process the events according to Bayesian inference, and still others may use a combination of both to process causal events.

In summary, the results of the present experiments reveal that (1) 6-month-olds processed BB events neither associatively nor in terms of Bayesian inference and (2) that adults may process causal events along a causal gradient rather than in a strict Bayesian or associative way.

References

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. *The developmental psychology of time*, 209-254.

Cohen, L. B., & Amsel, G. (1998). Precursors to infants' perception of the causality of a simple event. *Infant behavior and development*, 21(4), 713-731.

Cohen, L. B., Amsel, G., Redford, M. A., & Casasola, M. (1998). The development of infant causal perception. *Perceptual development: Visual, auditory, and speech perception in infancy*, 167-209.

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.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological review*, 111(1), 3.

Griffiths, T. L., Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2011). Bayes and blickets: Effects of knowledge on causal induction in children and adults. *Cognitive Science*, 35(8), 1407-1455.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality?. *Cognition*, 25(3), 265-288.

McClelland, J. L., & Thompson, R. M. (2007). Using domain-general principles to explain children's causal reasoning abilities. *Developmental Science*, 10(3), 333-356.

Oakes, L. M., & Cohen, L. B. (1990). Infant perception of a causal event. *Cognitive Development*, 5(2), 193-207.

Rakison, D. H., & Krogh, L. (2012). Does causal action facilitate causal perception in infants younger than 6 months of age?. *Developmental science*, 15(1), 43-53.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning: Current research and theory*.

Schlottmann, A., & Surian, L. (1999). Do 9-month-olds perceive causation-at-a-distance?. *Perception-London-*, 28, 1105-1114.

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Developmental psychology*, 40(2), 162.

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: the development of causal reasoning in toddlers and infants. *Developmental psychology*, 42(6), 1103.

Sobel, D. M., Tenenbaum, J. B., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive science*, 28(3), 303-333.