Don’t throw the (associative-learning) baby out with the bathwater just yet: Backwards-blocking reasoning with *multiple* potential causes in human children

Deon T. Benton1, David M. Sobel2 David Kamper2, Rebecca M. Beaton1,

1Vanderbilt University  
2Brown University

Address correspondence to Deon T. Benton, Department of Psychology and Human Development, Vanderbilt University, Peabody College, 230 Appleton Place #552, Nashville, TN 37235

Data availability statement: The code and network-modeling scripts are available upon reasonable request.

Conflicts of interests: none

Abstract

Causal reasoning is a fundamental cognitive ability that enables humans to learn about the complex interactions in the world around them. However, the available evidence suggests that the mechanism or set of mechanisms that underpin causal reasoning are not well understood. It is unclear, for example, whether causal reasoning is underpinned by a Bayesian mechanism, an associative mechanism, or both. Some theorists have argued that a Bayesian mechanism underpins causal reasoning because it can better account for backward-blocking (BB) and indirect screening-off (IS) findings in children and adults (e.g., Sobel, Tenenbaum, & Gopnik, 2004). However, the evidence is mixed about the extent to which learners engage in both kinds of reasoning. Here, we report three experiments that examine to what extent adults engage in BB and IS reasoning using the blicket-detector design (e.g., Gopnik et al., 2001), what mechanism best explains their behavior in this task, and under what conditions are adults’ causal ratings consistent with the predictions of the three competing computational and analytical models. The results of Experiment 1 revealed that adults’ causal ratings in the backwards-blocking condition (as well as in the indirect screening-off condition) were consistent with the predictions of the traditional and modified Rescorla-Wagner models when asked to reason about two objects. The results of the present study suggest that adults use associative processes to reason about two objects but a Bayesian-inference-like process to reason about three or more objects.

Keywords: causal reasoning; causal mechanisms; computational models; analytical models; associative learning; Bayesian inference

There is perhaps no ability that is more central for learning about how the world works than causal reasoning or the capacity to make inferences about cause-and-effect relations. This is a key cognitive ability because it enables human learners to encode causal relations to inform prediction and inference (e.g., Leslie & Keeble, 1987; Oakes & Cohen, 1990), to intervene on those relations to generate new effects (e.g., Gopnik et al., 2001), and counterfactually to reason about causal events to determine what would have happened if alternative actions were chosen (e.g., Harris, German, & Mills, 1996; Sobel, 2004).

To date, most studies on causal reasoning in human children have used the blicket-detector design. In these studies, children are introduced to a machine called the "blicket detector" and told that it lights up and plays music when certain objects—namely, "blickets"—are placed on it but not when non-blickets are placed on it. Children are then asked to determine which objects are blickets and to “make the machine go” by placing the blicket on the machine.

There have been several important findings to come from research with the blicket detector (e.g., Benton, Rakison, & Sobel, 2021; Gopnik & Sobel, 2000; Gopnik et al., 2001; Kimura & Gopnik, 2019; Meltzoff, Waismeyer, & Gopnik, 2012; Sobel & Kirkham, 2006; Sobel & Munro, 2006; Walker & Gopnik, 2014), but the finding that has generated that the most controversy was that by Sobel, Tenenbaum, and Gopnik (2004). They showed that by 4 years of age children can engage in two forms of causal reasoning called “backwards-blocking” (henceforth, BB) reasoning and “indirect screening-off” (henceforth, ISO) reasoning. BB reasoning is the process by which learners discount or “block” causal cues that are revealed to be redundant in producing some effect. ISO reasoning is the process by which learners discount or “screen off” a causal cue whose causal status is known unambiguously.

In their study, children were first shown that two novel objects, objects A and B, together caused the detector to activate when both objects were placed on the machine. Children were then shown that object A alone either failed to activate the detector (i.e., AB+ A-; ISO condition) or activated the detector when placed on it (i.e., AB+, A+; BB condition). Children in both conditions were then asked which of the two objects were blickets and to make the machine go by placing the blicket on the detector. Sobel et al. (2004) found that the 4-year-olds responded by placing object B on the machine in the ISO condition; these same children responded by placing object A on the machine in the BB condition. Subsequent research by Sobel and Munro (2009) found that 3-year-olds could also engage in BB and ISO reasoning if the machine was given the properties and desires of animate entities.

These findings were interpreted to mean that human children can engage in BB reasoning and that this form of reasoning is underpinned by a Bayesian-inference mechanism. The crux of the Bayesian-inference account is that human learners use a simple form of Bayes’ rule to reason about causal events and to choose the causal hypothesis—within a space of hypotheses that is potentially super-exponentially large—that is most consistent with the observed data (e.g., Sobel et al., 2004; Gopnik & Wellman, 2012). Crucially, proponents of this perspective have argued that associative learning cannot underlie human causal reasoning. The associative-learning model that has received perhaps the most criticism is the traditional Rescorla-Wagner (henceforth, RW) model (e.g., Rescorla & Wagner, 1972; Griffiths et al., 2011; Sobel et al., 2004). One reason that this model has been called into question is that it predicts that object B should be treated equivalently across the BB and ISO conditions—this prediction is at variance with participants’ actual treatment of object B across these conditions. Although the traditional RW model suffers from this important limitation, as we will return to below this is not the associative-learning mechanism that could explain how children process BB events.

Some caution should be exercised either before accepting the conclusion that Bayesian inference underpins human causal reasoning or the claim that human children can engage in BB reasoning. One major reason concerns the fact that there are problems with how BB reasoning has been measured in previous research. For example, Sobel et al. (2004; see also Beckers et al., 2009; McCormack et al. 2009, Exp. 1; Sobel & Kirkham, 2006) operationally defined BB reasoning as greater B choices in the ISO condition than in the BB condition (although for alternative operationalizations, see De Houwer, Beckers, & Glautier, 2002; Larkin, Aitken, & Dickinson, 1998; Griffiths et al., 2011; Kruschke & Blair, 2000; Lovibond et al., 2003; Shanks, 1985; Van Hamme and Wasserman, 1994). This way of operationally defining BB reasoning was likely motivated by two key factors. First, if the causal status of object A—which can be determined unequivocally when object A is placed alone on the machine—causes participants retrospectively to reevaluate the causal status of object B, then participants *should* consider B to be less of a blicket in the BB condition than in the ISO condition.

There are two key limitations with this operationalization, however. First, as Beckers et al. (2005) and McCormack et al. (2009) pointed out, it cannot be determined why participants treated object B differently between the BB and ISO condition. Differential treatment of object B could have been due to a BB effect, an ISO effect, or both. Such differential treatment could have also resulted from the fact that participants observed a positive effect during the elemental (i.e., A+) phase in the BB condition but a negative effect during the elemental (i.e., A-) phase in the ISO condition. Crucially, this would not be a true retrospective reevaluation of object B by participants based on A’s *relation to and effect* *on* object B across both conditions (which is the intended inference).

The operationalization that we adopt here—which was first introduced by McCormack et al. (2009, Exp. 2)—eschews this limitation. On this operationalization, BB reasoning is assessed by comparing how participants treat object B following an AB+ A+ sequence of events (i.e., the BB experimental condition) to how participants treat object B following an AB+ C+ sequences of events (i.e., the BB control condition). These two conditions differ in terms of the object that is shown during the elemental phase (i.e., A or C) and that object’s *relation* to B (and thereby the potential impact that this object has on how B is treated). For example, in the BB experimental condition, a dependency is presumably established between objects A and B because both objects appear together during the compound phase of the condition. This means that the observed causal efficacy of object A during the subsequent elemental phase *should* affect participants’ (retrospective) treatment of object B. In contrast, in the BB control condition, object C never appeared with object B, which necessarily means that C’s causal status should not affect how participants evaluate object B. Crucially, the blicket effect itself is held constant such that, across both conditions and the compound and elemental phases the machine activates.

**An open question**

There is yet another reason to exercise caution before accepting the claim that human beings use Bayesian inference to engage in BB reasoning. This has to do with the fact that it is not known whether human children engage in BB reasoning for three (or more) objects. The is because most, if not all, of the studies on BB reasoning in human children have tended to use two objects (i.e., objects A and B). This research is important because it has revealed that BB reasoning may emerge by 3 years of age, but it leaves unaddressed whether children can engage in BB reasoning when asked to reason about three or more objects. This issue is worth addressing because if a Bayesian-inference mechanism is assumed to underpin human causal reasoning in the real world, then it is crucial to show that children continue to engage in BB reasoning even when reasoning about three (or more) objects. In other words, if a key goal is to elucidate the cognitive mechanisms by which children reason causally *in the real world*, then it is crucial that we understand how causal reasoning unfolds in situations that more closely approximate children’s natural environments.

One may question whether asking children to reason about three or four objects can really tell us more about the cognitive mechanisms that underlay causal reasoning than asking children to reason about two objects. This is because the two situations differ trivially by at most two potential causes. However, if Bayesian inference is the cognitive mechanism that underpins BB reasoning, then the difference between these two settings is far from trivial. This is because in the two-potential-cause setting, participants need only to determine which of *four* candidate causal hypotheses generated the observed data. In contrast, in the three- or even four-potential-cause setting, participants need to determine which of *eight* (in the case of 3 candidate causes) or *sixteen* (in the case of 4 candidate causes) hypotheses is the one that generated the observed data. This means that participants must consider four times as many causal hypotheses in the four-potential-cause setting as participants in the two-candidate-cause setting.

Crucially, this difference may have important implications for whether an associative-learning mechanism or a Bayesian-inference mechanism underlies causal reasoning in children. For instance, it is possible that when children’s information-processing abilities are taxed—such as when they are asked to reason about three (or more) objects (see the General Discussion for a fuller discussion)—they may resort to simpler modes of causal inference that are better explained by associative processes. Thus, if participants’ BB performance adheres to the predictions of an associative-learning mechanism (see below) in a setting that includes many potential causes, this would suggest that the conclusion that associative processes cannot explain human causal reasoning may be premature.

**Possible cognitive mechanisms underlying BB reasoning for multiple potential causes**

Given that a key goal of the current study was to elucidate how or by what cognitive mechanism children reasoned about the present causal events, a critical first step was to derive the predictions of a simple Bayesian model. We also wanted to introduce, discuss, and derive the predictions of an associative-learning mechanism—which, crucially, does not suffer from the same limitations as the traditional RW model—that could explain how children process BB events such as those used here. We specifically focus on an associative-learning process that is based on a simple counting strategy (see below). To our knowledge, this is the first time that such an associative mechanism (see below) has been offered to account for children’s BB performance. This is because the debate has tended to focus on whether causal reasoning in human children is best explained by Bayesian inference or associative learning based on the traditional RW model. We restrict our discussion below to each model’s predictions but interested readers should consult the Appendix for the formal details of the Bayesian model.

**Bayesian inference.** Proponents of the Bayesian-inference account maintain that human learners use a simple form of Bayes’ rule to reason about causal events. Specifically, this perspective maintains that a learners’ responsibility is to determine which hypothesis—within a space that contains potentially an infinite number of psychological hypotheses—is responsible for observed data. The proposed cognitive mechanism by which this is achieved is by combining learners’ prior beliefs about each hypothesis (in the absence of data; this is sometimes called the “prior”) with the likelihood that the observed data was produced by a particular hypothesis (this is sometimes called the “likelihood”). Crucially, learners will retain a hypothesis to the extent that it can produce the observed data; they will discard a hypothesis when it no longer can produce the data.

Given that learners were asked to reason about three potential causes (i.e., objects A-C) during the experimental trials in the both the BB and ISO conditions and four potential causes during the control trials in both the BB and ISO conditions (i.e., objects A-D), the corresponding psychological hypothesis spaces consist, respectively, of 8 and 16 hypotheses. To illustrate how such a psychological hypothesis space might look, Figure 1 below shows the hypothetical hypothesis space for three objects.

**Timeline

Description automatically generated with medium confidence**

Figure 1. The eight different causal hypotheses indicating the possible causal relations for a causal event that involves three objects and one blicket detector. *A*, *B*, and *C* correspond to the three objects that were used on the machine and *E* indicates the activation of the machine.

By application of Bayes’ rule, the prediction that this model makes for how participants should treat the objects after the BB main trial is shown below in Table 1.

|  |  |
| --- | --- |
| BB Experimental Condition – 3 objects | |
| Object A | 1 |
| Object B | p |
| Object C | *p* |
| BB Control Condition – 4 objects | |
| Object A | *p* |
| Object B | *p* |
| Object C | *p* |
| Object D | 1 |

Table 1. This table displays the predictions of the Bayesian model for the BB experimental and control trials.

As can be seen, this model predicts that for any probability *p,* following the AB+ A+ BB experimental or main event participants should be maximally confident that object A is a blicket but should treat objects B and C equivalently; that is, they should produce an equivalent number of “yes” responses when asked whether objects B and C are blickets. In contrast, the model predicts that for any probability *p*, following the ABC+ D+ BB control trials participants should be maximally confident that object D is a blicket after the BB control trials but should treat objects A-C equivalently. The predictions that this model makes after the ISO experimental and control trials are shown below in Table 2.

|  |  |
| --- | --- |
| ISO Experimental Condition – 3 objects | |
| Object A | 0 |
| Object B | p |
| Object C | *p* |
| ISO Control Condition – 4 objects | |
| Object A | *p* |
| Object B | *p* |
| Object C | *p* |
| Object D | 0 |

Table 2. This table displays the predictions of the Bayesian model for the ISO experimental and control trials.

As is shown in Table 2, this model predicts that for any probability *p,* following the AB+ A- ISO experimental or main event participants should be maximally confident that object A is not a blicket but should treat objects B and C equivalently; that is, they should produce an equivalent number of “yes” responses when asked whether objects B and C are blickets. In contrast, the model predicts that for any probability *p*, following the ABC+ D- ISO control trials participants should be maximally confident that object D is not a blicket after the ISO control trials but should treat objects A-C equivalently. In sum, a simple Bayesian model predicts that learners should be maximally confident about the status of a candidate cause when it is shown in isolation but should treat objects that are shown in combination (and never alone) equivalently.

**Associative learning: a simple “counting” cognitive mechanism.** In contrast to a simple Bayesian model, an alternative cognitive mechanism that children may rely on is one that is based on a simple associative “counting” strategy. Such a mechanism would be sensitive to frequency with which (i.e., the number of times that) a given object—either individually or in combination with other objects—appeared with the blicket effect. To understand how this process might work mechanistically, consider the ABC+ D+ BB control trial. If this process is the one that best explains how learners arrive at their causal judgements in the current study, then learners should treat objects A-D equivalently following the BB control event. This is because all four objects would have been paired with the machine’s activation an equal number of times. In other words, object A would have been seen with the machine’s activation exactly once; object B would have been seen with the machine’s activation exactly once; object C would have been seen with the machine’s activation exactly once; and object D would have been seen with the machine’s activation exactly once. Below in Table 4 are the predictions that this account makes for all four conditions.

|  |  |  |  |
| --- | --- | --- | --- |
| The predictions of a simple associative-based counting mechanism | | | |
| BB main – ABC+ A+ | | | |
| Is A a blicket? | Is B a blicket? | Is C a blicket? | Is D a blicket? |
| +2 | +1 | +1 | N/A |
| BB control – ABC+ D+ | | | |
| Is A a blicket? | Is B a blicket? | Is C a blicket? | Is D a blicket? |
| +1 | +1 | +1 | +1 |
| ISO main – ABC+ A- | | | |
| Is A a blicket? | Is B a blicket? | Is C a blicket? | Is D a blicket? |
| 0 | +1 | +1 | N/A |
| ISO control – ABC+ D- | | | |
| Is A a blicket? | Is B a blicket? | Is C a blicket? | Is D a blicket? |
| +1 | +1 | +1 | 0 |

Table 4. This table shows the predictions a simple counting mechanism makes for the BB experimental and control trials and the ISO experimental and control trials. A 0 indicates that the number of times that a particular object was paired with the machine’s activation was exactly cancelled out by the number of times that it was paired with the machine’s inactivation. A +1 indicates that a particular object was paired with the machine’s activation once. A +2 indicates that a particular object was paired with the machine’s activation twice. Learners are said preferentially tos choose objects with larger values.

As shown in the table above, this account predicts that for the BB main trials children should say that object A is a blicket significantly more often than either B or C but that their treatment of B and C should not differ. This is because objects B and C would have been paired with the machine exactly once. In contrast, object A would have been paired with the machine twice. Similarly, this account predicts that during the BB control trials participants’ treatment of all four objects should not differ. This prediction results from the fact that all four objects would have been paired with the machine’s activation exactly once. In contrast, this account predicts that for the ISO experimental trials participants should not consider object A to be a blicket but should be split in their treatment of objects B and C because B and C would have been paired with the machine’s activation an equal number of times. Likewise, during the ISO control trials, this account predicts that participants should not consider object D to be a blicket but should be split in their treatment of objects A, B, and C.

**The present investigation**

The present investigation had two goals. First, it was designed to determine whether 5- and 6-year-olds could engage in BB reasoning when asked to reason about three and four objects and when a more appropriate measure of BB reasoning was used. Second, it was designed to gain greater insight into how—that is, by what underlying cognitive mechanism—children reasoned about the present causal events. We aimed specifically to determine which of the two cognitive mechanisms best accounted for children’s causal inferences in the present context. Given that some previous research operationally defined BB reasoning as greater treatment of object B in the BB condition compared to the ISO condition, participants in the current experiment also experienced the ISO condition (in a between-subjects manipulation).

**Current study**

Five- and 6-year-old children were introduced to a computer-animated machine called the “blicket detector” and were told that their task was to determine which objects made the machine. They were told that objects that made the machine “go” were called blickets; objects that did not make the machine go were not called blickets. Participants then received either two BB main trials and two BB control trials or two ISO main trials and two ISO control trials. Participants in both conditions were then asked to indicate whether the objects in each trial were blickets. Participants were randomly assigned to the BB or ISO conditions.

**Method**

**Participants.** Participants were 31 5-year-olds (16 boys and 15 girls) and 31 6-year-olds (17 boys and 14 girls). Sample size was determined based on previous studies on BB reasoning in human children (e.g., Gopnik & Sobel, 2000; Griffiths et al., 2011; Sobel et al., 2004). Two children were excluded from analysis for failing to participate (*N* = 1) or missing video (which made coding their responses impossible) (*N* = 1). Although most children were from white, middle-class backgrounds, a range of ethnicities that resembled the diversity in the population were represented. All children were tested in a quiet room at a children’s museum.

**Materials.** The “device” used in the experiments presented here was a computer-animated version of the blicket detector. The device was a white rectangle with a black border that measured 5.99 cm × 23.47 cm. If the device was “on”, the white region of the rectangle turned blue. If the device was “off”, the white region remained white. In addition, a maximum of 4 differently colored circles were used, and each circle measured 2.67 cm × 2.67 cm (see Figure 2 below). The machine was designed such that it activated immediately when a circle that was predetermined to be a blicket contacted it. At the start of any given trial, three (for the BB experimental trials) or four equally spaced (for the BB control trials) circles appeared above the blicket machine. Finally, the videos contained a built-in script, which experimenters were instructed to read. All video events were created in Microsoft PowerPoint.

**Procedure.** Participants were either tested in a quiet room on campus or in a quiet room in local children’s science museum. At the beginning of the experiment, all participants were shown a pretraining video. The video consisted of a rectangular base (i.e., the previously mentioned blicket detector) and two shapes (i.e., a gray triangle and a gray pentagon). Crucially, these shapes were unrelated to the circles that were used during the main portion of the experiment. The pretraining phase began with the triangle (object A) and pentagon (object B), which were located side-by-side and above the machine. Object A then descended until it contacted and immediately activated the machine (i.e., the white region changed from white to blue). Object A then ascended until it returned to its starting position above the machine. Object B then descended until it contacted and failed to activate the machine. Object B then returned to its starting position. Finally, both objects descended until they contacted the machine, which immediately activated. Participants were then asked whether each object was a blicket. This event was identical to the “one-cause” event in Gopnik et al. (2001) and was included to ensure that participants could reason about blicket objects.

Diagram

Description automatically generated

Figure 2. Schematic of one of the two BB experimental events. The upper-right portion of the figure shows the BB event as it unfolded across time. The lower-left portion of the figure shows the three objects and the text, “Is this one a blicket?” above each object across time.

Following the pretraining phase, participants were given four test trials—either the two BB experimental trials and 2 BB control trials or two ISO experimental trials and 2 ISO control trials—in counterbalanced order using a Latin square. Differently colored objects were used across all trials to prevent carryover effects.

The two BB main trials began with three differently colored objects, which were located above the machine. The text, “Look, I have these three toys. Let’s find the blickets. Watch what happens” appeared above the objects. All three objects (i.e., objects A, B, and C) then descended until they contacted and activated the machine. At this point, the text, “Look, these also make the machine go!” appeared above the objects. The objects then ascended to their starting positions. The left- or right-most (counterbalanced) object (i.e., object A) then descended until it contacted and immediately activated the machine. The text, “Look, this one makes the machine go!” then appeared above the objects. This object then returned to its starting position. Children were then asked verbally to respond with a “yes” or “no” if they thought that each object was a blicket. Specifically, the text, “Is this one a blicket?” with a downward-facing arrow then appeared above each object, and participants were asked to indicate whether each object was a blicket. The first and second BB experimental trials were identical except for the object colors.

The two BB control trials began with four differently colored objects (i.e., objects A, B, C, and D), which were located above the machine. Objects A, B, and C then descended until they contacted and activated the machine. Object D then descended by itself until it contacted and activated the machine. The left-right position of object D was counterbalanced. Children were then asked whether each object was a blicket. Critically, whereas object A descended with the remaining two objects in the BB experimental trials, object D did not descend with the remaining three objects in the control trials. The experimental and control trials used the same text.

Finally, the ISO experimental and control conditions were identical to the BB experimental and control conditions except that objects A (during the ISO main trials) and D (during the ISO control trials) failed to activate the machine (see Table 1 for a schematic).

|  |  |  |  |
| --- | --- | --- | --- |
| Schematic of Experiment 1 | | | |
|  | Compound | Elemental | Test |
| BB experimental trial | ABC+ | A+ | Is A/B/C a blicket? |
| BB control trial | ABC+ | D+ | Is A/B/C/D a blicket? |
| ISO experimental trial | ABC+ | A- | Is A/B/C a blicket? |
| ISO control trial | ABC+ | D- | Is A/B/C/D a blicket? |

Table 5. The +/- signs corresponds to whether the machine activates (+) or not (-)

**ResultsChart, bar chart

Description automatically generated**Figure 3. Participants’ mean responses to whether each object was a blicket across the conditions and trial types (i.e., eventType). Bars show standard error.

Figure 3 shows the results for this experiment. The dependent measure was the number of times that participants responded “Yes” to the “Is this a blicket” question. Thus, across two trials, the maximum number of times that a participant could respond “Yes” was 2; the minimum number of times that a participant could respond “Yes” was 1. Using this dependent measure, the data were entered into a four-way linear model with Age (5-year-olds vs. 6-year-olds) and Condition (BB vs. ISO) as the between-subjects factor and Trial Type (experimental vs. control) and Objects (A vs. B vs. C vs. D) as the within-subjects factors. This analysis revealed a main effect of Condition, *F*(1, 380) = 11.66, *p* < .001, a main effect of Objects, *F*(3, 380) = 8.09, *p* < .001, and a main effect of Trial Type, *F*(1, 380) = 11.96, *p* < .001.These significant main effects were qualified by a significant two-way interaction between Condition and Objects, *F*(3, 380) = 24.15, *p* < .001, a significant two-way interaction between Condition and Trial Type, *F*(1, 380) = 5.88, *p* < .001. These two-way interactions were further qualified by a significant three-way interaction between Condition, Objects, and Trial Type, *F*(2, 380) = 19.88, *p* < .001. This three-way interaction is shown in Figure 3.

We followed up this three-way interaction with separate one-way linear models for the main and control trials within the BB and ISO conditions. The Objects factor was treated as the sole within-subjects factor in these follow-up analyses. The first one-way linear model for the control trials within the BB condition did not reveal a significant effect of Objects, *F*(3, 118) = 0.26, *p* = .85. This means that participants treated the objects similarly during the control trials of the BB condition. In contrast, the second one-way linear model for the main trials within the BB condition revealed a significant main effect of Objects, *F*(2, 84) = 10.25, *p* = .001. This main effect reflected the fact that participants considered object A to be more of a blicket (*M* = 1.97, *SD* = 0.18) than object B (*M* = 1.25, *SD* = 0.84), *t*(27) = 4.42, *p<* .001, or object C (*M* = 1.32, *SD* = 0.82), *t*(27) = 4.12, *p* < .001.

The third and fourth one-way linear models for the main and control trials within the ISO condition both revealed a significant main effect of Objects, both *F*’s > 23.24, both *p*’s < .001. This reflected the fact that participants considered object A (*M* = 0.48, *SD* = 0.85) to be significantly less of a blicket than objects B (*M* = 1.68, *SD* = 0.69) and C (*M* = 1.72, *SD* = 0.61) during the main condition, both *t*’s > -5.82, both *p*’s < .0001, and object D (*M* = 0.67, *SD* = 0.66) to be less of a blicket than objects A (*M* = 1.74, *SD* = 059), B (*M* = 1.66, *SD* = 0.60), and C (*M* = 1.74, *SD* = 0.61) during the control trials, all *t*’s > -6.01, all *p*’s < .001.

**Assessing BB under the new operationalization of BB reasoning**

To examine whether there was evidence of BB reasoning according to the new operationalization of BB reasoning—in which participants’ treatment of some of the redundant causes is compared to their treatment of other redundant causes *within* the BB condition—data for the redundant causes within the BB experimental and control conditions were entered into a two-way linear model with Objects (A, B, and C) and Trial Type (main vs. control) as the within-subjects factors. This analysis revealed only a main effect of Trial Type, *F*(1, 144) = 4.90, *p* = .03, which reflected the fact that participants were more likely to respond that an object was a blicket during the control trials (*M* = 1.60, *SD* = 0.71) than during the main trials (*M* = 1.29, *SD* = 0.82). Crucially, neither the main effect of Objects, *F*(2, 144) = 0.07, *p* = .93, nor the interaction between Objects and Trial Type*, F*(1, 144) = 0.02, *p* = .89, was significant. This reflected the fact that participants treated the redundant causes equally within the BB main and control conditions.

**Assessing BB under the old operationalization of BB reasoning**

To examine whether there was evidence of BB reasoning according to the old operationalization of BB reasoning—in which participants’ treatment of some of the redundant causes is compared to their treatment of other redundant causes *between* the BB and ISO conditions—data for the redundant causes between the BB and ISO conditions were entered into a three-way linear model with Condition (BB vs. ISO) as the sole between-subjects factor, and Objects (A, B, and C) and Trial Type (main vs. control) as the within-subjects factors. This analysis yielded only a main effect of Condition, *F*(1, 277) = 7.60, *p* < .01. This reflected the fact that participants were more likely to call an object a blicket in the ISO condition (*M* = 1.71, *SD* = 0.59) compared to the BB condition (*M* = 1.48, *SD* = 0.77). These results indicated that participants treated the redundant causes equivalently between the BB and ISO main and control trials. More broadly, these results provide no support for BB reasoning according to the new or old operationalizations of BB reasoning.

General Discussion

This study had two aims. The first aim was to examine whether 5- and 6-year-olds would engage in BB reasoning when asked to reason about 3 objects. This departs from the typical convention of using two objects to study causal reasoning in human children. We evaluated BB reasoning under both an old and new operationalization. The second aim was to clarify the debate on *how* children reason about the present causal events. Specifically, we were interested inn whether participants used a Bayesian-inference mechanism or an associative-based counting mechanism to process the current causal events.

With respect to the first aim, there was no evidence that children engaged in BB reasoning. This was true regardless of how we measured BB reasoning. This finding extends previous research to show that when children are asked to reason about three objects, they do not engage in BB reasoning. With respect to the second aim, the data were most consistent with an associative-based counting mechanism. This is because children’s willingness to say that an object was a blicket depended on the frequency with which that object was paired with the machine’s activation; the more frequently that the object was paired with the machine’s activation, the more likely children were to say that the object was a blicket. This finding is significant because it has been suggested that causal reasoning in human children is best explained by Bayesian inference and rational processes rather than by associative processes. The present data suggest that whether children engage in Bayesian inference depends on how causal reasoning is measured and the number of objects about which they are asked to reason.

One potential criticism of this study is that the results are inconsistent with the findings from previous studies on BB reasoning in human children. Such previous research showed that children do engage in BB reasoning when asked to reason about two objects; the current study showed that children do not engage in BB reasoning when asked to reason about three objects. However, we believe that the present results extend (rather than are at odds with) such previous research. Specifically, the present study demonstrates that when children’s information-processing capacities are stretched such as when they are asked to reason about multiple potential causes, they may deploy and rely on simpler associative processes. Although at the level of individual objects the difference between three and four objects is miniscule, by contrast the corresponding increase in the size of the underlying psychological hypothesis space is substantial. Such an increase in the size of the underlying psychological hypothesis space may have important ramifications on the cognitive mechanism that gets deployed by children, especially if children are sensitive to and affected by this increase. For example, children who are asked to reason about two candidate causes need only to represent and choose among *four* candidate causal hypotheses (i.e., 2n, where *n* is the number of potential causes). Four candidate causal hypotheses may be within the information-processing capacities of 5- and 6-year-olds. In contrast, children who are asked to reason about three candidate causes must now consider *eight* candidate causal hypotheses, which may well be outside the limits of their restricted information-processing capacities for the developing child.

It turns out that there is a wealth of data that is consistent with this general proposal (Doebel & Zelazo, 2015; Frye, Zelazo, & Palfai, 1995; Zelazo, Frye, & Rapus, 1996; Zelazo et al., 2003). One such recent study by Kenderla and Kibbe (2023) showed that when 8- and 10-year-old children’s information-processing abilities were stretched in a virtual memory game—such as when children were asked to find three cards that shared one feature and differed on another feature—they relied less on working memory and more on manual exploration. Given that children are not required actively to maintain information in memory when manually exploring, manual exploration is an ostensibly simpler, less cognitively effortful strategy than one that requires working memory. In a similar vein, Richland, Morrison, and Holyoak (2006) found that 3- and 4-year-old children made more featural and relational errors when asked to reason about multiple relations or when a salient distractor was made to compete with the critical relation than when asked to reason about a single relation without a distractor. Finally, there is evidence that preschool-age children's performance on theory-of-mind and social-problem-solving tasks is adversely affected when they are first made to complete tasks that taxed their information-processing abilities compared to when such capacities were not taxed (Caporaso & Marcovitch, 2021; Powell & Carey, 2017; Steinbeis, 2018).

Together, this research demonstrates that although children can process information at higher levels, if the task that they are given requires information-processing abilities that extend beyond what they possess, then there will be a tendency for them to process information at lower levels and to rely on less sophisticated strategies and cognitive mechanisms. This may provide a developmental explanation for why children in the present study did not engage in BB reasoning or show evidence that they relied on Bayesian inference. A testable prediction of this account is that there should be a point at which children go from using a simple associative-based counting mechanisms in contexts like the present one to more rationale processes like Bayesian inference. This issue should be explored more fully in future research.

A second potential criticism is that we cannot be sure that a simple Bayesian-inference mechanism underpinned participants’ performance in the present study. For example, if participants assumed that blickets were common in the present context—which is plausible given how frequently the detector activated in the present study—then participants should be *less* likely to block redundant causes; in other words, participants should be *more* likely to treat all potential blickets (expect for the ones that are explicitly shown not to be causal) equally. This could explain participants’ performance in the BB control condition—in that condition, participants treated all objects equally. However, this explanation cannot explain *all* the present data. This is because this explanation predicts that participants should have also treated objects A-C equivalently in the BB experimental condition as well, but this was not the case: Participants treated object A differently than either objects B or C in the BB experimental condition.

Nonetheless, because we did not systematically manipulate base-rate information, this alternative explanation cannot yet be ruled out entirely. However, if we are correct that participants do not rely on Bayesian inference when asked to reason about multiple causes, we predict that their performance in this proposed future study would not differ from participants’ performance in the current study. However, if children’s causal judgements are shown to be affected by base-rate information, such that their BB reasoning performance changes with changes to base-rate information, then this would suggest that participants may use Bayesian inference to reason about multiple candidate cause after all, at least when a Bayesian-inference mechanism is primed by explicitly and systematically manipulating base-rate information.

**Conclusion**

These potential criticisms notwithstanding, these experiments constitute one of the first systematic attempts to examine BB and ISO reasoning in human children in the context of three objects. A longstanding view has been that the cognitive mechanism by which human beings reason about causal events is Bayesian inference (e.g., Gopnik et al., 2004) rather than associative processes. The experiments reported here support a different conclusion: an associative-learning counting mechanism supports 5- to 6-year-old children’s reasoning about multiple potential causes.

References

Beckers, T., Vandorpe, S., Debeys, I., & De Houwer, J. (2009). Three-year-olds’ retrospective revaluation in the blicket detector task: Backward blocking or recovery from overshadowing?. *Experimental Psychology*, *56*(1), 27-32.

Benton, D. T., Rakison, D. H., & Sobel, D. M. (2021). When correlation equals causation: A behavioral and computational account of second-order correlation learning in children. *Journal of Experimental Child Psychology*, *202*, 105008.

Caporaso, J. S., & Marcovitch, S. (2021). The effect of taxing situations on preschool children’s responses to peer conflict. *Cognitive Development*, *57*, 100989.

Doebel, S., & Zelazo, P. D. (2015). A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children. *Developmental Review*, *38*, 241-268.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive development*, *10*(4), 483-527.

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child development*, *71*(5), 1205-1222.

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., & Wellman, H. M. (2012). Reconstructing constructivism: causal models, Bayesian learning mechanisms, and the theory theory. *Psychological bulletin*, *138*(6), 1085.

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.

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*(3), 233-259.

Houwer, J. D., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *The Quarterly Journal of Experimental Psychology: Section A*, *55*(3), 965-985.

Kenderla, P., & Kibbe, M. M. (2023). Explore versus store: Children strategically trade off reliance on exploration versus working memory during a complex task. *Journal of Experimental Child Psychology*, *225*, 105535.

Kimura, K., & Gopnik, A. (2019). Rational higher‐order belief revision in young children. *Child Development*, *90*(1), 91-97.

Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin and Review*, *7*(4), 636-645.

Larkin, M. J., Aitken, M. R., & Dickinson, A. (1998). Retrospective revaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1331.

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

Lovibond, P. F. (2003). Causal beliefs and conditioned responses: retrospective revaluation induced by experience and by instruction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 97.

McCormack, T., Butterfill, S., Hoerl, C., & Burns, P. (2009). Cue competition effects and young children’s causal and counterfactual inferences. *Developmental psychology*, *45*(6), 1563.

Meltzoff, A. N., Waismeyer, A., & Gopnik, A. (2012). Learning about causes from people: observational causal learning in 24-month-old infants. *Developmental psychology*, *48*(5), 1215.

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

Powell, L. J., & Carey, S. (2017). Executive function depletion in children and its impact on theory of mind. *Cognition*, *164*, 150-162.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Classical conditioning II: Current research and theory, 2, 64-99.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children’s development of analogical reasoning: Insights from scene analogy problems. *Journal of experimental child psychology*, *94*(3), 249-273.

Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *The Quarterly Journal of Experimental Psychology Section B*, *37*(1b), 1-21.

Sobel, D. M. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology*, *22*(1), 37-58.

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., & Munro, S. (2006, August). When Mr. Blicket wants it, children are Bayesian. In *Proceedings of the Cognitive Science Society* (pp. 810-816).

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.

Steinbeis, N. (2018). Taxing behavioral control diminishes sharing and costly punishment in childhood. *Developmental science*, *21*(1), e12492.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and motivation*, *25*(2), 127-151.

Walker, C. M., & Gopnik, A. (2014). Toddlers infer higher-order relational principles in causal learning. *Psychological science*, *25*(1), 161-169.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive development*, *11*(1), 37-63.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.