Investigating the Explore/Exploit Trade-off in Adult Causal Inferences

Erik Herbst (e.herbst@mail.utoronto.edu)
Department of Psychology, University of Toronto, Canada

Chris Lucas (c.lucas@ed.ac.uk)
School of Informatics, University of Edinburgh, United Kingdom

Daphna Buchsbaum (buchsbaum@psych.utoronto.ca)
Department of Psychology, University of Toronto, Canada

Abstract

We explore how adults learn counterintuitive causal relationships, and whether they discover hypotheses by revising their beliefs incrementally. We examined how adults learned a novel and unusual causal rule when presented with data that initially appeared to conform to a simpler, more salient rule. Adults watched a video of several blocks placed sequentially on a blicket detector, and were then asked to determine the underlying causal structure. In the near condition the true rule was complex, but could be found by making incremental improvements to the simple and salient initial hypothesis. The distant condition was governed by a simpler rule, but to adopt that rule participants had to set aside their initial beliefs, rather than revising them incrementally. Adults performed better in the near condition, despite this rule being more complex, providing some of the first evidence for an explore-exploit trade-off in inference, analogous to the trade-off in active learning.

Keywords: causality, Bayesian inference, hypothesis search, process model

Background

Any time we make plans, predict the future, or attempt to understand why events occurred in the past, we are relying on causal knowledge. In acquiring this knowledge, we must draw conclusions from sparse, noisy, and ambiguous evidence. We gain the ability to make sense of this limited information at an early age, with causal thinking showing signs of emergence even in infancy (Sobel & Kirkham, 2006; 2007; Walker & Gopnik, 2014). By adulthood, our frameworks for interpreting causal phenomena become much more complex and able to accommodate diverse areas of knowledge (Kemp, Goodman, & Tenenbaum, 2007).

Despite its usefulness, sometimes our ability to generalize from past causal inferences can lead us astray, as in the case where we encounter a new causal relationship that is rare or strange by the standards of our past experience. For instance, we might expect that either of two switches will turn on a lamp, when in fact the lamp turns on when the switches are in matched positions. While our causal learning process is generally accurate and adaptive (e.g., Griffiths & Tenenbaum, 2005), in the current paper we claim – in the spirit of previous "rational process" models (e.g. Sanborn, Griffiths, & Navarro, 2010) – that human causal beliefs are updated in a limited or local fashion that is efficient but subject to systematic failures under certain conditions. This is especially true when the initial hypothesis is at a local optimum – the best hypothesis within reach, but not the best overall – and when the true causal structure is distant from our initial hypothesis in some hypothesis space. Suppose you break out in a rash every time you buy your favourite candy bar from a vending machine. After searching for the proper cause, you would probably conclude that you are allergic to the candy as soon as it comes to mind. You may be unlikely to consider that you are actually reacting to the coins used to purchase the candy bar, even if this is indeed the case. In this case, discovering the real cause requires abandoning your working hypothesis, rather than just incrementally refining it.

Bayesian Models of Causal Inference

Several researchers have attempted to explain learning of novel causal relationships using hierarchical Bayesian models of inference (e.g. Griffiths, Sobel, Tenenbaum, & Gopnik, 2011; Griffiths, Kemp, & Tenenbaum, 2008). Recent evidence demonstrates that adults and children can successfully modify their causal beliefs in light of new and surprising evidence in a manner that suggests Bayesian inference strategies (e.g., Griffiths, Sobel, Tenenbaum, & Gopnik, 2011; Lucas, Bridgers, Griffiths, & Gopnik, 2014). Through this process, learners also create and update higher-level models of how causal relationships operate in general. Regardless of whether human cognition functions exactly this way, hierarchical Bayesian models have accurately predicted human causal learning (Kemp, Goodman, & Tenenbaum, 2007; Lu, Yuille, Lijeholm, Cheng, & Holyoak, 2006; Lucas & Griffiths, 2010; Ullman, Goodman, & Tenenbaum, 2012).

Although Bayesian models accurately capture many aspects of human causal reasoning, they may not fully account for adults' relative difficulties in learning more unusual types of causal relationships. Specifically, Lucas and colleagues (2014) found that young children were more likely than adults to discover an unusual conjunctive causal relationship. Children and adults were tasked with inferring a causal principle after viewing a machine that activated when certain blocks or block combinations were placed on top of it. Even after viewing evidence that blocks only activated the machine in specific pairs (and not individually), adults had more difficulty than children with generalizing this principle to new blocks. One possibility for this finding is that adults are more biased by prior experiences—as they have observed that conjunctive relationships are relatively rare—which leads them to demand strong evidence before they infer a conjunctive relationship is present. Indeed, if cognition operates via Bayesian principles, there are conceivably instances in which rigid commitment to a prior may preclude learners from uncovering the true nature of a causal relationship. However, this may not apply in novel causal situations with which adults have limited experience. Moreover, adults are cognitively different than children beyond simply having more experience, so differences in causal reasoning may in fact be the by-product of some developmental change.

As an alternative to simply having different priors, adults' relative difficulty with conjunctive causal relationships may be explained in terms of the *process* by which they explore and weigh new hypotheses in light of their current beliefs. It is typically impossible to evaluate all potential hypotheses (of which there may be an infinite number). Bayesian inference is often intractable in practice for complex problems, so human inferences must sometimes depart from the Bayesian ideal. Nonetheless, there is evidence that people may be resource rational observers, making approximately Bayesian inferences in ways that make efficient use of limited time and memory (Bonawitz, Denison, Gopnik, & Griffiths, 2014; Sanborn, Griffiths, & Navarro, 2010). As for possible processes underlying these approximations, some empirical phenomena, such as order effects, offer clues. If learners make inferences from a complete set of data, as traditional Bayesian models assume, then they should not be influenced by the order in which stimuli are presented. Nevertheless, humans are sensitive to presentation order (Danks & Schwartz, 2006; Sanborn, Griffiths, & Navarro, 2010). One explanation for these order effects is that people arrive at solutions by considering a small number of hypotheses at any single moment in time, and updating or replacing them sequentially with more data – sometimes losing information and leading to small but systematic errors. More recently, Bayesian process models have been proposed to explain these patterns of errors by drawing analogies to Monte Carlo sampling methods that permit tractable and efficient inference in applied statistics and machine learning (Abbott, Hamrick, & Griffiths, 2013; Shi, Griffiths, Feldman, & Sanborn, 2010).

Inference techniques are often modelled using Monte Carlo methods that update sequentially and incrementally. These methods allow hypotheses to be revised by sampling from the posterior, without computing the posterior distribution in its entirety. Markov chain Monte Carlo sampling is a popular and efficient subclass of Monte Carlo methods, and it is marked by a degree of stickiness or inertia, in which people hew more closely to their initial hypotheses than a truly optimal Bayesian learner would. This family of models predicts that individuals will tend toward inferences that are similar to their prior beliefs. For example, one study showed that when people made inferences about a causal system, they tended toward solutions that required the fewest single edits to their initial hypothesis, where a single edit is an addition, subtraction, or reversal of a causal link (Bramley, Dayan, Griffiths, & Lagnado, 2017). Therefore, causal process models can account for multiple limitations on causal learning, and have recently been shown to explain phenomena such as classical anchoring (Lieder, Griffiths, Huys, & Goodman, 2017). Learners can be constrained not only by priors, but also the similarity of candidate hypotheses to their current beliefs, perhaps precluding them from finding too-distant hypotheses.

The Explore-Exploit Trade-off in Inference

These findings could reflect a cognitive tradeoff in development that affects how learners search through hypotheses. When presented with a wide range of possibilities, individuals must often decide whether to employ a general, shallow search or a narrow, deep one. This decision is analogous to the explore-exploit tradeoff, whereby decision-makers must allocate cognitive resources to either exploit previous knowledge or explore alternatives (Sutton & Barto, 1998). Adults may be more inclined to exploit, by searching nearby solutions extensively—and less likely to explore hypotheses that require unusual, low-probability edits to the current hypothesis. With limitations on the number of hypotheses a learner can consider, exploitation-biased adult learners could plausibly benefit from focusing cognitive resources on hypotheses that are refinements of an initial proposal that is plausible and informed by long experience. This will increase efficiency of finding adequate solutions but potentially limit access to distant alternatives. Conversely, exploration-focused learners (young children, perhaps) may spread out their search over a more diverse range of possibilities. Although this approach sacrifices the ability to efficiently refine already-reasonable hypotheses, it may grant access to unusual solutions that would be unreachable with a more conservative search.

Thus, the inferential explore-exploit trade-off may have interesting implications for the process of selecting between competing hypotheses. This selection process has been modelled using Bayesian algorithms for both children and adults (Bonawitz, Denison, Gopnik, & Griffiths, 2014; Denison, Bonawitz, Gopnik, & Griffiths, 2013; Lieder, Griffiths, & Goodman, 2012; Sanborn, Griffiths, & Navarro, 2010), but relatively little previous work has examined adults' potential tendencies toward exploitation. As one possible example of how hypothesis search may reflect an exploitation bias, Gopnik and colleagues have likened human belief updating to simulated annealing; just as the heating and gradual cooling of metal can increase its malleability, so can a gradual "cooling" of an inference method corresponding to an increasingly conservative search policy lead to better inferences (Gopnik, Griffiths, & Lucas, 2015; Lucas, Bridgers, Griffiths, & Gopnik, 2014). For instance, while young children may use high-temperature searches, considering a wide range of hypotheses with relatively equal probability, adults' searches are "cooler" and more narrow in scope. Although commitment to priors may still matter, simulated annealing allows us to examine which types of hypotheses are considered. High-temperature searches are more likely to discard adequate hypotheses, but may allow individuals to escape local optima and discover unlikely solutions that are potentially better. In contrast, low-temperature searches can quickly converge to good solutions if fewer low-probability edits are required to get there, but may otherwise get trapped in local optima. With this in mind, adults may have more difficulty discovering unusual causal relationships because their search is too focused and too close to their initial guesses to accommodate distant ideas.

The purpose of our current studies is to test the hypothesis that belief updating in adults is *exploitation-biased*. To accomplish this, we designed a task encouraging participants to generate a particular initial hypothesis about a novel causal relationship. Evidence that contradicted this hypothesis was then presented, causing participants to modify their beliefs. The true causal structure took one of two forms corresponding to two experimental conditions. In the *near condition*, the correct causal structure was closer to the initial hypothesis but designed to be relatively complex. In the *distant condition*, the correct causal structure was simpler but possibly harder to reach when making incremental changes from the initial hypothesis, which is a local optimum. Thus, we hoped to determine the breadth of hypotheses that participants were willing to entertain. If adults' search process is more exploitation-biased, we should expect the near-hypothesis solution would be more easily found than the distant one, even if both rules are a priori equally unlikely. However, if adults' failure to infer unlikely causal relationships is simply due to the low prior probability that they place on these relationships, then they should be equally unlikely to consider either solution.

Experiment 1: Investigating the Explore-Exploit Tradeoff in Inference

Participants Participants were 90 adult US residents, recruited through Amazon Mechanical Turk and paid a base rate of \$1 for their time. An additional \$1 bonus was given to the top 10% performers as an additional incentive. Participants were divided randomly among near (n = 45) and distant (n = 45) conditions. Six participants from the near condition and seven from the distant condition were excluded due to failure to correctly answer attention manipulation tasks.

Materials and Procedure The methods used in this study are similar to those used in previous blicket tasks (e.g. Gopnik & Sobel, 2000), except that animated video stimuli were presented online using Qualtrics survey software (similar to Buchsbaum et al., 2012). Participants were asked to examine several blocks and determine which blocks are blickets. They were informed that blickets are blocks that activate the blicket detector, and were shown a video of an animated blicket detector activating and not activating. Participants then watched a five-minute animation depicting 20 blocks being consecutively placed onto the blicket detector. If the block was a blicket, the detector lit up and a sound played. The blocks were sorted into blicket/non-blicket categories and left on screen for participants to study.

Whether a block was a blicket depended on specific aspects of the block pattern. Each block had a coloured background (red or blue) and several small red or blue triangles in a fixed pattern (see Figures 1 and 2). The block pattern was such

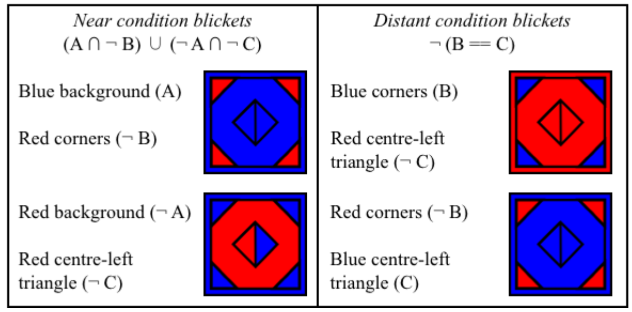

Figure 1: Examples of blickets in the near condition (left) and the distant condition (right).

that the background colour was the most obvious and visually striking feature. For the first 15 blocks (the *initial rule-consistent* blocks), the background colour appeared to determine whether the blocks activated the machine—i.e. blocks with one background colour consistently activated the machine, while the others did not. Inspired by an experimental manipulation in Williams and Lombrozo (2010; 2013), this was designed to lead participants to an initial causal hypothesis based on the objects' most salient feature. The final five blocks (the *initial rule-violating* blocks), however, violated this initial hypothesis; the blocks that did and did not activate the machine had the opposite background colour as before. Thus participants needed to modify their initial hypothesis to capture the optimal solution.

The true rule separating blickets from non-blickets varied based on condition. This true rule determined whether a block was a blicket 100% of the time. In the near condition, the background colour was related to whether a block was a blicket, whereas in the distant condition the background colour was unrelated. Each block had five binary features (Figure 1), which could vary by colour on each block (background, corners, centre-left triangle, centre-right triangle, and border), giving a total of 32 different colour combinations. In the near condition, blocks were blickets based on a combination of the background colour and the colour of two secondary features. In the distant condition, only the colour of these two secondary features determined whether a block was a blicket, while the background colour was irrelevant.

Thus, the five features could be labeled as follows: one primary feature (A), two relevant secondary features (B and C), and two irrelevant secondary features (D and E). In the distant condition, the optimal rule for determining whether a block is a blicket—that is, the simplest rule that perfectly explains the data—can be written as $R = (B == C)$, whereas the optimal rule in the near condition can be written as $R = (A \cap \neg B) \cup (\neg A \cap \neg C)$. These rules were designed to seem arbitrary to naïve participants and minimize the role of the participants' prior knowledge. In the near condition, there is a consistently-improving path of single edits to transition from the initial hypothesis, $R = A$, to the correct rule, where a single edit consists of adding or subtracting a variable or changing an operator (e.g. changing $R = A$ to $R = A \cap \neg B$; Goodman & Tenenbaum, 2008 use a similar approach for searching a hypothesis space). In the distant condition, the single-edit path to the correct rule requires edits that initially worsen the hypothesis (e.g. removing A as a relevant variable). Participants must therefore ignore the ineffectiveness of these local edits and keep exploring to find the correct solution. Thus, if adults use a Bayesian single-edit search process with an exploit bias, participants should be less likely to abandon $R = A$, and thus should perform more poorly in the distant condition, where $R = A$ is the local optimum.

The lists of blocks seen by participants in the near and distant conditions were generated randomly with the following constraints: a) there were ample block feature combinations that participants did not see, so that they could be tested on these blocks later, and b) the rules and edit paths conformed to the specifications in the previous paragraph. Thus, the final sets of blocks were as follows: near condition participants saw 11 blickets (3 initial rule-violating) and 9 non-blickets (2 initial rule-violating), whereas distant condition participants saw 10 blickets (2 initial rule-violating) and 10 non-blickets (3 initial rule-violating). The differences in block numbers were necessary due to the constraints of the conditions.

Following the presentation of all of the blickets, participants saw a blicket rating task, in which they were asked to judge whether a randomized series of eight blocks were blickets. For each block, participants rated how certain they were that it was, or was not, a blicket, on a seven-point Likert scale ranging from "definitely a blicket" to "definitely not a blicket". Blocks were balanced by background colour, blicket/non-blicket status, and whether they had already been presented in the observation stage. Participants received a score between -3 and 3 for each block based on accuracy and certainty, and the sum of these scores determined their final score for this task. Next, participants completed a forced-choice task, where they chose which of two blocks was more likely to activate the blicket detector, for a series of four pairs. Blocks were selected randomly such that there were an equal number of initial rule-consistent and initial rule-violating blocks, and blocks in each pair differed from each other in background colour and whether they were a blicket. Participants received a point for each correct block judgment.

Afterwards, the participants were asked to describe the causal rule they had inferred. They were then told to imagine that a new rule was suggested by a friend, and asked if they preferred this rule over their own. This rule always represented the correct causal structure. The purpose of this question was to ensure that any differences between the two conditions were not due to participants finding the near rule inherently more plausible or likely than the distant one. The participants' rule preference was measured using a seven-point scale. Finally, each participant received questions to test their task comprehension and an instructional manipulation task to control for inattention, similar to the one used by Oppenheimer, Meyvis, and Davidenko (2009).

Results and Discussion If adults' hypothesis search strategy is exploitation-biased, participants in the near condition will perform better on both tasks than those in the distant condition. The results supported our predictions. For the forced-choice task, a 2x2 ANOVA was run with condition (distant/near) and rule consistency (initial rule-consistent/violating) as factors (see Table 1 for a score summary). Near condition participants outscored those in the distant condition, $F(1, 84) = 6.46$, $p = .01$, $MSE = 0.26$. Participants also scored higher for initial rule-consistent blocks, than for rule-violating blocks, $F(1, 84) = 226$, $p < .001$, $MSE = 0.34$. There was no significant interaction effect, $F(1, 84) = 0.154$, $p > .69$, $MSE = 0.34$.

For the blicket rating task, a 2x2 mixed ANOVA (condition x rule consistency) was run (see Table 2 for a score summary). The analysis found that participants were much more likely to confidently identify initial rule-consistent blocks than initial rule-violating blocks $F(1, 84) = 131$, $p < .001$, $MSE = 15.32$, suggesting that the salience manipulation was effective and participants were influenced by the background colour. Supporting our forced-choice results, there was a marginally significant effect of condition, $F(1, 84) = 3.77$, $p = .06$, $MSE = 11.87$, with a mean score of 7.51 for the near condition and 4.63 for the distant condition (scores ranged from -24 to 24).

Intriguingly, and unlike in the forced-choice task, there was also a significant interaction effect, $F(1, 84) = 3.34$, $p = .04$, $MSE = 15.32$. This is a result of participants in the near condition performing better than those in the distant condition on initial rule-consistent blocks, but equally poorly on initial rule-violating blocks. To assess whether this interaction was due to differences in confidence for some blocks, an additional 2x2 mixed ANOVA (condition x rule consistency) was run to investigate participants' certainty ratings when evaluating blocks. The analysis showed no main effect of condition, $F(1, 84) = 2.30$, $p > .13$, $MSE = 0.69$. Mean confidence ratings were relatively near ceiling in both conditions (greater than 2 out of 3), which may partially explain the lack of a main effect. However, participants were more certain of their answers when rating initial rule-consistent blocks than when rating rule-violating blocks, $F(1, 84) = 22.0$, $p < .001$, $MSE = 0.32$. There was also a highly significant interaction effect between condition and rule-consistency, $F(1, 84) = 13.1$, $p < .001$, $MSE = 0.32$, driven by participants in the near condition having more certainty for initial rule-consistent blocks than for rule-inconsistent blocks, suggesting that while participants in the near condition were better able to correctly categorize both initial rule-violating and initial rule-consistent blocks, they were most confident about the latter.

Additional one-sample t-tests examined whether participants scored better than would be expected by chance. For the forced-choice task, participants correctly classified blocks as blickets and non-blickets significantly better than chance in the near condition, $t(42) = 5.82$, $p < .001$, but *not* in the distant condition, $t(42) = 1.31$, $p = 0.20$. In the blicket rating task, however, participants classified blocks better than

Table 1: Mean scores and SE for forced-choice task. Total scores range from 0 to 4, and scores for initial rule-consistent and initial rule-violating blocks range from 0 to 2.

Condition	Near	Distant
Total score	2.53($\pm$0.10)	2.24($\pm$0.12)
Rule-consistent	1.90($\pm$0.08)	1.82($\pm$0.07)
Rule-violating	0.77($\pm$0.13)	0.42($\pm$0.07)

Table 2: Mean scores and SE for blicket rating task. Total scores range from -24 to 24, and scores in each sub-category range from -12 to 12.

Condition	Near	Distant
Total score	8.00($\pm$1.04)	4.87($\pm$1.26)
Rule-consistent	9.59($\pm$0.51)	6.39($\pm$0.72)
Rule-violating	-1.59($\pm$1.01)	-1.53($\pm$1.06)

chance in both the near condition, $t(42) = 7.69$, $p < .001$, and the distant condition $t(42) = 4.13$, $p < .001$. The at-chance performance of distant condition participants in the forced-choice task may simply reflect the low number of trials compared to the blicket rating task.

Finally, we looked at participants' preference for the correct rule over their own. Participants in the distant condition significantly preferred the correct friend's rule over their own rule, $t(42) = 4.78$, $p < .001$, while participants in the near condition did not, $t(42) = 1.55$, $p = .13$. Participants in the distant condition also preferred the friend's rule significantly more than those in the near condition, $t(75) = 2.09$, $p = .04$. This supports our hypothesis that participants in the distant condition had not previously considered the distant rule, rather than that they considered it, but dismissed it as unlikely.

Experiment 2: A priori rule preference

Although the main study compared the extent to which participants preferred the correct rule over their own, it did not examine the rules in both conditions side-by-side. This study investigated adults' a priori preference for either the near or the distant rule without differentiating data. This was to confirm that differences in causal learning and rule preference between conditions in Experiment 1 were not due to an intuitive preference for the near rule before seeing any data.

Participants Participants were 51 adult US residents, recruited through Amazon Mechanical Turk (MTurk) and paid a base rate of $0.50 for their time.

Materials and Procedure As in the previous study, participants were told that blickets were blocks that activated the blicket detector, and saw an animated blicket detector activating and not activating. Unlike the previous study, however, participants only saw one block placed on the machine, causing it to activate. They were then told the two possible rules, and that both rules accurately described this block, but that only one rule was the correct rule for identifying blocks that activate the machine. Participants were asked to choose which rule they thought was more likely to be correct. These rules were identical to the near rule and the distant rule from the previous study, and the blicket that participants saw was chosen from a set of blocks that conformed to both rules. Finally, after selecting a rule, participants explained why they chose that rule and rated their confidence in their decision, ranging from 1 (just guessing) to 7 (completely certain). This confidence rating was turned into a score ranging from -7 (completely certain the near rule is correct) to 7 (completely certain the distant rule is correct) for statistical analysis.

Results and Discussion Of the 51 participants, 22 preferred the near rule and 29 preferred the distant rule, $p = .41$, exact binomial test. A one-sample t-test demonstrated that the rule preference scores, $M = 0.25$, $SE = 0.50$, did not significantly differ from chance, $t(49) = 0.71$, $p = 0.48$. Thus, participants did not prefer one rule over the other, suggesting that it was not an a priori preference for the near rule driving the results of Experiment 1.

General Discussion

The findings obtained by these studies lend support to the exploitation-biased search hypothesis. We expect that exploitation-biased searches of the hypothesis space will be more likely to discover rules close to the initial hypothesis, and less likely to discover more distant rules, even if they are less complex. As predicted, participants were more accurate at classifying blocks in the near condition than the distant condition. This is especially notable given that participants in Experiment 2 found both rules equally a priori plausible, which supports that the near rule is at least as complex as the distant rule. This in turn makes it less likely that the differences between conditions can be explained by differently-weighted prior probabilities. Participants performed better in the near condition, where the true rule was arguably more complex, but was comparatively easier to discover from the salient starting point due to the consistently-improving edit path, than in the distant condition, where the true rule was simpler, but where the salient rule was a local optimum. This suggests that adults are searching through their hypothesis space in an exploitation-biased manner.

Nevertheless, participants were better able to identify initial rule-consistent blocks than initial rule-violating ones in both tasks. This suggests that the strength of one's priors may still play a role in conjunction with the exploitation bias. However, this difference in performance suggests intriguing future research avenues—in particular, the finding in the blicket rating task that participants in the near condition scored higher than those in the distant condition on initial rule-consistent but not initial rule-violating blocks. This seems to be driven largely by participants' relative certainty toward initial rule-consistent blocks in the near condition, rather than their accuracy at categorizing the blocks (as measured by the forced choice task). Future studies might assess how nearness to an initial hypothesis affects the certainty of judgments of causal relationships.

It is still unclear, however, if these difficulties in discovering certain causal relationships are the result of a developmental process. Consequently, we plan to expand this study to directly compare adults with children, to examine whether children possess these same search-related difficulties. If these findings are the result of a developmental shift toward exploitation-based search strategies, then exploration-oriented children could perform just as well—if not better—than adults in tasks such as those in this study. Children should also perform equally well in both experimental conditions, or perhaps even better in the distant condition than in the near one. Particularly, this may be the case if children see the near rule as a priori less likely. When comparing children's and adults' performance, it may also be useful to note differences in time spent on each task, as it might generate additional insights about their hypothesis search process. Although participants in the current studies had unlimited time to complete each task, timing data were not recorded.

In the future, it may be useful to develop a more explicit process model to measure hypothesis distance. Although the near-hypothesis rule is closer to the salient hypothesis, in that adding and subtracting particular predicates improves the hypothesis toward the correct rule, this may not accurately represent how individuals process locality. In other words, we lack a precise model for how people move between rules, and thus exactly how far $R = (B == C)$ is from $R = A$, and how much harder it is to find $R = (A \cap \neg B) \cup (\neg A \cap \neg C)$. In future experiments, this process model will need to be clarified.

Overall, our results demonstrating that adults are able to discover a true causal structure nearer to an initial hypothesis more readily than a distant causal structure of equal or *greater* complexity provides compelling initial evidence for an explore-exploit trade-off in causal inferences. This may help inform future research on how individuals generate new hypotheses about everyday causal interactions.

References

Abbott, J., Hamrick, J., & Griffiths, T. (2013). Approximating Bayesian inference with a sparse distributed memory system. *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 16861691). Berlin.

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-Stay, Lose-Sample: A simple sequential algorithm for approximating Bayesian inference. *Cognitive psychology, 74*, 35–65.

Bonawitz, E., Denison, S., Griffiths, T.L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: sampling in cognitive development.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neuraths ship: Approximate algorithms for online causal learning. Psychological review, 124(3), 301.

Buchsbaum, D., Bridgers, S., Whalen, A., Seiver, E., Griffiths, T. L., & Gopnik, A. (2012). Do I know that you know what you know? Modeling testimony in causal inference. In *Proceedings of the 34th annual conference of the cognitive science society.*

Danks, D., & Schwartz, S. (2006). Effects of causal strength on learning from biased sequences. In *Proceedings of the 28th annual meeting of the cognitive science society.*

Denison, S., Bonawitz, E., Gopnik, A., & Griffiths, T. L. (2013). Rational variability in children's causal inferences: The sampling hypothesis. *Cognition, 126*(2), 285–300.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2014). *Bayesian data analysis* (Vol. 2). London: Chapman & Hall/CRC.

Glymour, C. N. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology.* MIT press.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A Rational Analysis of RuleBased Concept Learning. *Cognitive Science, 32*(1), 108-154.

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When younger learners can be better (or at least more open-minded) than older ones. *Current Directions in Psychological Science, 24*(2), 87-92.

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

Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. In R.Sun (Ed.), *Cambridge handbook of computational psychology* (pp. 59100). New York: Cambridge University Press.

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

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive psychology, 51*(4), 334-384.

Kemp, C., Goodman, N. D., & Tenenbaum, J. B. (2007). Learning causal schemata. *Cognitive Science Society.*

Lieder, F., Griffiths, T. L., & Goodman, N. D. (2012). Burn-in, bias, and the rationality of anchoring. *Advances in Neural Information Processing Systems, 25.*

Lieder, F., Griffiths, T., Huys, Q. J., & Goodman, N. D. (2017). The anchoring bias reflects rational use of cognitive resources. *Psychonomic Bulletin & Review.*

Lu, H., Yuille, A., Lijeholm, M., Cheng, P. W., & Holyoak, K. J. (2006). Modeling causal learning using Bayesian generic priors on generative and preventive powers.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition, 131*(2), 284-299.

Lucas, C. G., & Griffiths, T. L. (2010). Learning the form of causal relationships using hierarchical Bayesian models. *Cognitive Science, 34*(1), 113-147.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867-872.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological review, 117*(4), 1144.

Shi, L., Griffiths, T. L., Feldman, N. H., & Sanborn, A. N. (2010). Exemplar models as a mechanism for performing Bayesian inference. *Psychonomic bulletin & review, 17*(4), 443-464.

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., & Kirkham, N. Z. (2007). Bayes nets and babies: Infants' developing statistical reasoning abilities and their representation of causal knowledge. *Developmental science, 10*(3), 298-306.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1). Cambridge: MIT press.

Ullman, T. D., Goodman, N. D., & Tenenbaum, J. B. (2012). Theory learning as stochastic search in the language of thought. *Cognitive Development, 27*(4), 455-480.

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

Williams, J. J., & Lombrozo, T. (2010). The role of explanation in discovery and generalization: evidence from category learning. *Cognitive Science, 34*(5), 776-806.

Williams, J. J., & Lombrozo, T. (2013). Explanation and prior knowledge interact to guide learning. *Cognitive psychology, 66*(1), 55-84.