Cognition 183 (2019) 57–66

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

Mature counterfactual reasoning in 4- and 5-year-olds

Angela Nyhout, Patricia A. Ganea

Applied Psychology & Human Development, University of Toronto, 252 Bloor Street West, Toronto, Ontario M5S 1V6, Canada

ARTICLE INFO

Keywords:
Counterfactual reasoning
Causal reasoning
Physical causation

ABSTRACT

Counterfactual reasoning is a hallmark of the human imagination. Recently, researchers have argued that children do not display genuine counterfactual reasoning until they can reason about events that are overdetermined and consider the removal of one of multiple causes that lead to the same outcome. This ability has been shown to emerge between 6 and 12 years of age. In 3 experiments, we used an overdetermined physical causation task to investigate preschoolers' ability to reason counterfactually. In Experiment 1a, preschoolers (N = 96) were presented with a "blicket-detector" machine. Children saw both overdetermined (2 causal blocks on a box) and single-cause trials (1 causal and 1 non-causal block) and were asked what would have happened if one of the two blocks had not been placed on the box. Four-year-olds' performance was above chance on both trial types, and 5-year-olds' performance was at ceiling, whereas 3-year-olds did not perform above chance on any trial types. These findings were replicated in Experiment 1b with 4- and 5-year-olds (N = 40) using more complex question wording. In Experiment 2 (N = 40, 4- and 5-year-olds), we introduced a temporal delay between the placement of the first and second block to test the robustness of children's counterfactual reasoning. Even on this more difficult version of the task, performance was significantly above chance. Given a clear and novel causal structure, preschoolers display adult-like counterfactual reasoning.

1. Introduction

The ability to mentally manipulate representations of past events is a cornerstone of the human imagination. Known as counterfactual reasoning, this ability is important to understanding the causes of events, and adapting one's behaviour in the future (Byrne, 2016; Epstude & Roese, 2008). For instance, an individual who concludes that, had she not eaten mangos, she would not have a rash, may decide in future to forgo opportunities to eat the offending fruit. Counterfactual reasoning, however foundational and adaptive, has recently been suggested to be beyond the ability of young children (Rafetseder, Cristi-Vargas, & Perner, 2010; Rafetseder & Perner, 2014; Rafetseder, Schwitalla, & Perner 2013). In the present study, we investigated whether, given a sufficiently simple and clear physical causation task, 3- to 5-year-old children would be able to demonstrate mature counterfactual reasoning.

There is considerable debate over when the ability to think counterfactually reaches maturity, with previous research finding the ability develops in the preschool years (e.g., Buchsbaum, Bridgers, Weisberg, & Gopnik, 2012; Harris, German, & Mills, 1996) to as late as adolescence (Rafetseder et al., 2010, 2013). Part of the disagreement has to do with how exactly counterfactual thinking is conceptualized. Some researchers view counterfactual thinking broadly. Generally, these theorists propose that counterfactual thinking is early to develop, and shares a common underlying basis with other abilities to imagine deviations from reality, such as pretend play and future thinking (Buchsbaum, Bridgers, Skolnick Weisberg, & Gopnik, 2012; Weisberg & Gopnik, 2013). According to these theories, counterfactuals may take the form of future hypotheticals, timeless conditionals, or past counterfactuals.

The earliest research on children's counterfactual reasoning suggested that the ability was available to children by age 4. In a seminal study, Harris et al. (1996) presented children with short vignettes involving simple cause-and-effect relations and asked them counterfactual questions. In one, a character walks across a clean floor with muddy shoes and makes the floor all dirty. After listening to the story, children were asked what would have happened if she had taken her shoes off. Three-and-a-half-year-olds were able to correctly respond that the floor would be clean. Several other studies have found that children begin engaging in counterfactual reasoning by age 4 (e.g., Beck, Robinson, Carroll, & Apperly, 2006; German & Nichols, 2003; Riggs, Peterson, Robinson, & Mitchell, 1998; Robinson & Beck, 2000, Study 1).

Others take a narrower view of counterfactual thinking, holding that counterfactual thinking specifically concerns alternatives to *past* events and is qualitatively different from other abilities to imagine

E-mail addresses: angela.nyhout@utoronto.ca (A. Nyhout), patricia.ganea@utoronto.ca (P.A. Ganea).

https://doi.org/10.1016/j.cognition.2018.10.027
Received 17 April 2018; Received in revised form 21 October 2018; Accepted 30 October 2018
0010-0277/

alternatives to reality (e.g., Beck, 2016). For example, Beck proposes that the reasoner must simultaneously hold in mind both the way things are and the way things could have been, which carries considerable executive demands (Beck, 2016; Beck & Riggs, 2014). Rafetseder and colleagues distinguish between genuine counterfactual and basic conditional reasoning, and have argued that children in Harris et al. (1996) and similar tasks could have relied on the latter (Leahy, Rafetseder, Perner, 2014; Rafetseder et al., 2010; 2013; Rafetseder & Perner, 2014). Basic conditional reasoning involves using one's general knowledge of causal regularities (e.g., clean shoes mean clean floors). In contrast, *mature* counterfactual reasoning involves respecting the *nearest possible world constraint*, which stipulates that one should change only those features of an event that are causally dependent on a counterfactual antecedent, and hold all else constant (Edgington, 2011).

To test this proposal, Rafetseder et al., (2013) presented children with scenarios similar to those used by Harris et al. (1996, Study 1), but introduced a second cause of the outcome, yielding a *causally overdetermined* outcome. Instead of just one character, Susie entering the house with muddy shoes, a second character, Max, also entered with his muddy shoes. Now, if asked what would have happened if Susie had taken her shoes off, children who are engaging in counterfactual reasoning should respond that the floor would still be dirty (because Max still had his dirty shoes on). Those who are relying on basic conditional reasoning should answer that the floor would be clean, because they are reasoning based on the counterfactual premise without regard for the sequence of events. Rafetseder et al. (2010, 2013) found support for their hypothesis. It was not until children were 12 years old that they demonstrated mature counterfactual reasoning, answering that the floor would still be dirty.

Some recent findings cast doubt on the interpretation that children cannot engage in mature counterfactual reasoning until adolescence. Nyhout, Henke, and Ganea (2017) suggested that Rafetseder et al. (2013) tasks may have underestimated children's performance by mischaracterizing the causal structure of events children were representing. If children did not have the requisite causal knowledge, or represented the causal relations in a way that did not conform to the expectations of the researchers, then they could not be expected to arrive at the “correct” counterfactual response. For example, children in the previous study may have inferred that Max would have done the same as Susie (i.e., take his shoes off). Nyhout et al. (2017) investigated children's ability to think counterfactually about overdetermined outcomes, and manipulated the causal relation between antecedents, such that for half of the children the two antecedent events were causally connected to one another, and for the other half the antecedents were causally disconnected. By the age of 8, children could reason counterfactually about various types of scenarios and took into account the causal relationship between events. Six-year-olds performed well when the antecedents were causally connected and thus the causal structure of the events more specified, making less room for unwarranted inferences.

In another recent study, McCormack, Ho, Gribben, O'Connor, and Hoerl (2018) presented children aged 4 through 8 with an overdetermined physical causation task. Two ramps led towards a gap with a toy pig in the middle. If a disc was rolled down either ramp, it would knock the pig over. A disc could be blocked by inserting a peg along its ramp. McCormack and colleagues asked children both subtractive (“If I had not rolled the red disc…”) and additive (“If I had put a peg in here…”) counterfactual questions. In overdetermined cases, the pig still would have fallen over because the other disc still would have rolled. They found that children's performance on these types of trials exceeded chance between the ages of 6 and 7 and reached ceiling between the ages of 8 and 9. Although 4- and 5-year-olds' performance did not exceed chance on overdetermined trials, their performance was better than younger children's in Rafetseder et al.'s (2013) studies. This developmental trend appears to be consistent with Nyhout et al.'s (2017) findings. The evidence so far suggests that mature counterfactual reasoning is in place by 6 years of age. However, previous research may have mischaracterized children's causal inferences by providing opaque causal structures, and therefore failed to find evidence for robust counterfactual thinking early in development.

In the present study, we presented children with clear causal structures about overdetermined events in the physical domain. We argue that previous studies (McCormack et al., 2018; Nyhout et al., 2017; Rafetseder et al., 2013) have underestimated children's counterfactual reasoning ability. In the case of previous narrative tasks (Nyhout et al., 2017; Rafetseder et al., 2013), the causal structure was not fully transparent and children may have made unwarranted inferences about the events in question. In the case of the one previous physical causation task (McCormack et al., 2018), children were required to learn a number of rules about how the device functioned and therefore their representation of the causal structure may not have been robust. Our argument is not that apparent failures of counterfactual reasoning were due to *failures* of causal reasoning. Rather, we suggest that previous studies may not have done enough to ensure that children's comprehension of the causal structure was both correct and sufficient to answer counterfactual questions, if children possessed the ability to do so.

We identified a context in which children show early and sophisticated causal reasoning. The “blicket-detector” task is a commonly used paradigm in studies of causal reasoning, and involves presenting children with a novel toy or machine that lights up when some types of objects are placed on it (e.g., Gopnik & Sobel, 2000). Numerous studies of children's causal reasoning suggest that children understand the causal structure of blicket-detector paradigms from a young age (e.g., Gopnik, Sobel, Schulz, & Glymour, 2001; Sobel, Tenenbaum, & Gopnik, 2004). The version of the blicket-detector task used in the current study was structurally similar to the vignettes presented to children in the previous studies of counterfactual reasoning (Nyhout et al., 2017; Rafetseder et al., 2013). Two independent causes (blocks) led to the same outcome (the toy lighting up), yielding an overdetermined outcome.

The blicket-detector paradigm has been used by researchers in some previous studies of children's counterfactual inferences (McCormack, Butterfill, Hoerl, & Burns, 2009; McCormack, Simms, McGourty, & Beckers, 2013). McCormack et al. (2009) presented trials in which blocks were placed singly or in pairs on the blicket-detector. Some blocks were never placed singly on the box, and therefore the question was whether children would use their knowledge of another block's causal status to infer the causal role of the second block. On both causal (e.g., “Is this one a blicket?”) and counterfactual questions (e.g., “Do you think it would have gone off if I hadn't put this one on?”), 5 and 6-year-olds used their knowledge of one block to infer the causal role of the other, whereas 4-year-olds did not. This task differs from our own in a few important respects. First, the causal learning phase in McCormack et al.'s study was much more complex. In the present studies, children first saw each block placed singly on the box and therefore there was no question of its causal efficacy. McCormack et al.'s task required causal *inferencing*, whereas our own required causal learning via observation. Second, the counterfactual questions in McCormack et al.'s study did not pit counterfactual reasoning against basic conditional reasoning. It remains an open question how children will perform on a variant of the blicket-detector task in which the causal structure is clear, and counterfactual questions concern unambiguously overdetermined outcomes. Counterfactual questions about overdetermined outcomes provide a valuable litmus test of children's counterfactual reasoning, as they pit the answers provided by mature counterfactual and basic conditional reasoning against one another.

When asked a counterfactual question about the removal of one event that contributes to an overdetermined outcome, those employing mature counterfactual reasoning respond that the outcome would still have occurred, whereas those using basic conditional reasoning answer that the outcome would *not* have occurred, because they reason on the

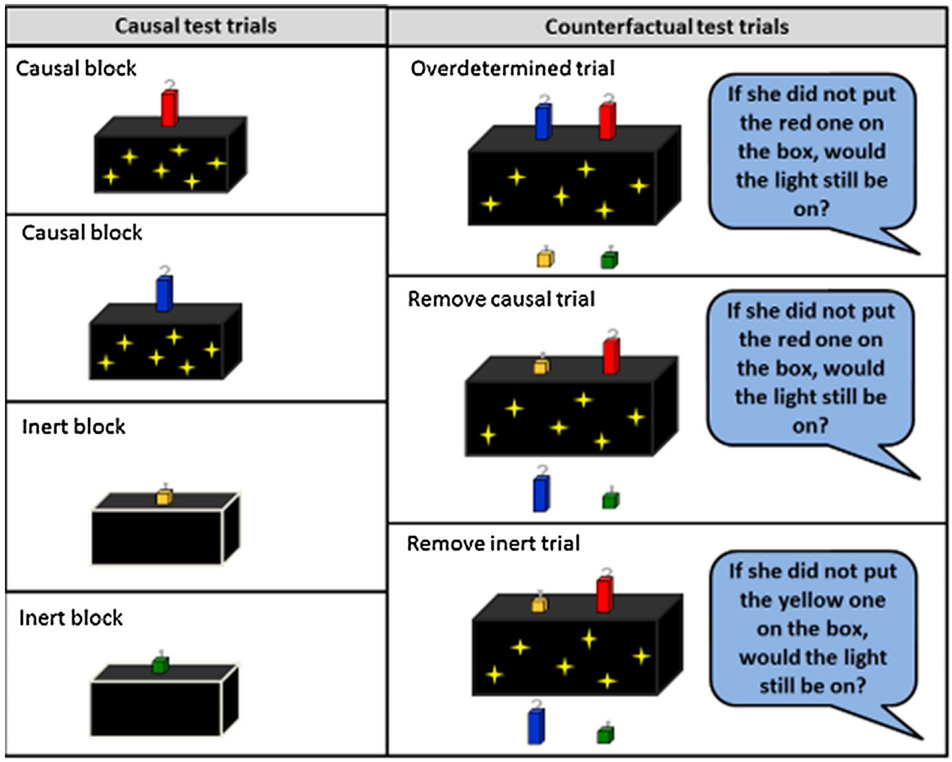

Fig. 1. Schematic of causal and counterfactual test trials used in Experiment 1a and 1b. Exp 1b used different question wording.

basis of the information in the premise without regard for the actual sequence of events. Recall that children in previous studies of counterfactual reasoning about overdetermined outcomes have not passed until at least age 6 (McCormack et al., 2018; Nyhout et al., 2017) to as late as age 12 (Rafetseder et al., 2013). If children's difficulty in previous studies stemmed in part from errors they made in representing the causal structure of events, then presenting children with a clear causal structure should bolster their performance, and we may see evidence for mature counterfactual reasoning at a younger age.

2. Experiment 1a

2.1. Method

2.1.1. Participants

The final sample included 96 children (M = 4.58, SD = 0.76 years, 51 girls): 32 older 3-year-olds (M = 3.73, SD = 0.16, range = 3.50–3.99), 32 4-year-olds (M = 4.52, SD = 0.31, range = 4.02–4.99 years), and 32 5-year-olds (M = 5.49, SD = 0.30, range = 5.04–5.97 years). Sixteen additional children were tested but excluded due to insufficient English exposure (n = 13), failing to answer causal questions correctly (n = 1), inattention (n = 1), and experimenter error (n = 1). Participants were recruited and tested at a science museum (n = 40) or in our university laboratory (n = 56). The sample was representative of the community from which it was drawn, with families reporting their ethnicity as White (40%), mixed ethnicity (27%), Chinese (6%), West Asian (2%), Latin American (2%), South Asian (2%), Aboriginal (2%), Black (1%), and Korean (1%). In 67% of families, at least one parent had a university degree or higher, 10% had a community college diploma, 4% had a high school diploma, and 1% had some high school. Demographic information was not specified by 18% of families.

2.1.2. Design and procedure

Children were shown a video in which an adult actor sat at a table with a black wooden shoebox in front of her. Drilled into the shoebox were several holes with small white lights poking through. The actor placed four blocks in front of the box and demonstrated their function by placing them on the box one-by-one. Two of the blocks switched the lights on (causal blocks), and two did nothing (inert blocks). Each block was a different colour (red, blue, green, and yellow) to allow for unambiguous reference, and the causal versus inert blocks differed in size (tall or short) and the type of appendage they had on the top (hooks or screws). The colour, size, and appendage of which blocks were causal were counterbalanced across children. For example, in one counterbalancing order, the actor demonstrated that tall red and blue blocks with hooks on them each activated the box, whereas the short yellow and green blocks with screws on them did not. The actor referred only to the colour of the blocks (e.g., "Let's try the red one!") and did not label the size or the appendage.

On *causal test trials*, after the actor had demonstrated each block individually, the experimenter asked the child about the function of each block to ensure the child understood their causal status (e.g., "Did the red one make the lights switch on?"). She then asked the child, "Can you remind me, which blocks made the lights switch on?" and "Which blocks did not make the lights switch on?". This served to ensure the child learned and remembered the causal status of each block before entering the counterfactual phase. All included participants answered these questions correctly. One child answered causal questions incorrectly and was excluded from analyses.

On *counterfactual test trials*, the actor placed two blocks on the box at a time. Each child saw 6 trials in which she placed either the two causal blocks, or one causal and one non-causal block on the box at the same time. After the actor placed the blocks on the box, the experimenter paused the video and asked the child a counterfactual question: "The light switched on! If she did not put the (colour) block on the box, would the light still be on?". There were three different trial types: (1) *overdetermined*, in which she placed both causal blocks on the box at the same time and the child was asked about the removal of one, (2) *single cause: remove causal trials*, in which she placed one causal and one inert block on the box, and the child was asked about the removal of the causal block, and (3) *single cause: remove inert trials*, in which she placed one causal and one inert block, and the child was asked about the removal of the inert block. Children received two trials of each type. Trial order was counterbalanced in 4 pseudo-random orders, with the requirement that trials of the same type not appear back-to-back. A schematic of trial types in presented in Fig. 1.

Children's answers were recorded on paper by the experimenter during the session and later scored, and sessions were video recorded for later reliability coding. A second coder, blind to the purpose of the study, coded 29 participants' data (30%). Coding agreement was excellent (97.5% agreement), $\kappa = 0.935$, $p < .001$.

2.2. Results

Nine participants (7 3-year-olds, 2 4-year-olds) answered "yes" to all questions and 9 participants (6 3-year-olds, 3 4-year-olds) answered "no" to all questions. We ran analyses both using children's actual scores, and corrected scores for which we changed children's scores to zero for all trials if they answered yes or no to all questions, because these potentially biased strategies yielded correct answers on 4/6 and 2/6 trials, respectively. In all results sections, we report the results using the actual scores, and report only when the pattern using the corrected scores deviated. In nearly all cases across the three experiments, the two methods of coding results yielded the same pattern of results.

We ran a generalized estimating equation (GEE), a semi-parametric regression technique that accounts for covariation between measures in modelling repeated measures or correlated data. Included in the model were trial type (overdetermined, remove causal, remove inert trials) as a within-subjects categorical variable, age as a continuous covariate, and score (out of 2) as the dependent variable. The model had a multinomial probability distribution and a cumulative logit function. The effect of trial type on score was not significant, $p = .922$. The effect of age on score was significant ($B = 0.90$, $SE = 0.29$, Wald $\chi^2(1) = 9.80$, $p = .002$, 95% CI = [0.34, 1.46]). The trial type by age interaction was also non-significant, $p = .938$.

Table 1 displays the percentage of children in each age group scoring 2 out of 2 on each trial type. For analyses against chance, we compared the proportion of children scoring 2 out of 2 to a chance distribution of 0.25 using binomial tests. We applied Bonferroni correction and adopted an alpha-value of 0.005 (0.05/9). On all three trial types, the proportion of 3-, 4-, and 5-year-olds scoring 2/2 was significantly higher than chance, p-values < 0.001, with the exception of 3-year-olds' performance on remove inert trials, $p = .012$. Using corrected scores, however, the proportion of 3-year-olds scoring 2/2 did not exceed chance on *overdetermined trials*, $p = .012$, *remove causal trials*, $p = .077$, and *remove inert trials*, $p = .419$.

Table 1
Percentage of children scoring 2 out of 2 on each trial type in Experiments 1a, 1b, and 2 Percentage using corrected scores in parentheses. In cases where no children showed a bias to answer yes or no to all questions, original and corrected scores are identical.

	Overdetermined	Remove causal	Remove inert
Experiment 1a			
3-year-olds (n = 32)	65.6** (43.5)	56.3** (37.5)	43.8 (21.9)
4-year-olds (n = 32)	71.9** (65.6)**	65.6** (56.3)**	62.5** (56.3)**
5-year-olds (n = 32)	90.6** (90.6)**	84.4** (84.4)**	78.1** (78.1)**
Total	76** (66.6)**	68.8** (59.4)**	61.5** (52.1)*
Experiment 1b			
4-year-olds (n = 20)	90** (80)**	55* (55)*	60* (50)
5-year-olds (n = 20)	90** (85)**	70** (70)**	75** (70)*
Total	90** (82.5)**	62.5** (62.5)**	67.5** (60)**
Experiment 2			
4-year-olds (n = 20)	75** (70)**	80** (70)**	60* (55)*
5-year-olds (n = 20)	85** (80)**	80** (75)**	75** (70)**
Total	80** (75)**	80** (72.5)**	67.5** (62.5)**

Binomial test (chance = 25%).
** $p < .001$.
* $p < .005$.

Considering performance across trials types, if children were using a strategy consistent with counterfactual reasoning, they should answer "Yes" to the counterfactual question on the *overdetermined* and *non-causal* trials, but not on the *remove causal* trials. On the other hand, if children were employing basic conditional reasoning, they should answer "No" to all questions, answering on the basis of the information contained in the counterfactual premise only (i.e., if a given block wasn't placed on the box, the lights wouldn't have switched on). This simpler reasoning strategy would have led to above chance performance on *remove causal* trials, but not on *overdetermined* and *remove inert* trials (see Fig. 2 for a graphic display of these strategies across trial types). Only 9 participants (6 3-year-olds, 3 4-year-olds) showed a response pattern consistent with basic conditional reasoning. Table 2 displays the percentage of children who displayed the following reasoning strategies: counterfactual (2/2 on all trial types), basic conditional ("no" to all trial types), overdetermined correct (with 2/2 on overdetermined trials, and incorrect or mixed responses on single causal trials), "yes" to all, or mixed (e.g., 1/2 on each trial type). This classification is conservative, as it treats only children with perfect scores as counterfactual reasoners. The modal reasoning strategy for 4- and 5-year-olds was counterfactual, whereas 3-year-olds did not show a tendency toward any of the reasoning strategies. We compared the proportion of children displaying each reasoning strategy to the proportion that would be expected by chance, using binomial tests and a Bonferroni-corrected alpha value of 0.003 (0.05/15). Chance levels were calculated by considering the number of ways participants could receive each reasoning strategy classification out of the total 27 response patterns (3 trial types, with scores of 0, 1, and 2 possible on each), and are displayed in Table 2. Significantly more 3- ($p = .001$), 4- ($p < .001$), and 5-year-olds ($p < .001$) were classified as counterfactual reasoners than predicted by chance (chance = 3.7%). The only other groups significantly above chance were 3-year-olds classified as basic conditional reasoners ($p < .001$) and 3-year-olds who responded yes to all questions ($p < .001$). In all three age groups, significantly fewer children displayed a mixed response pattern than predicted by chance, $ps < 0.001$.

3. Experiment 1b

To increase children's chance of success on the task, we did not use typical counterfactual language in Experiment 1a, given that it is syntactically complex. It is possible that the question wording we used led children not to reason counterfactually, but instead perhaps to make predictions or reason about the functions of the blocks. In a second experiment, we changed the question wording to the past subjunctive to ensure questions conformed to typical counterfactual syntax (Iatridou, 2000).

3.1. Method

3.1.1. Participants

The final sample included 40 children (M = 5.01, SD = 0.66, range = 4.02–5.88 years): 20 4-year-olds (M = 4.42, SD = 0.33, range = 4.02–4.92 years), and 20 5-year-olds (M = 5.60, SD = 0.24, range = 5.09–5.88 years). Eight additional children were tested, but their data were excluded from analyses due to insufficient English exposure (n = 6), incorrectly answering causal questions (n = 1), or inattention (n = 1). Participants were recruited and tested at a science museum (n = 17) or in our university laboratory (n = 23). Families reported their ethnicity as White (25%), mixed ethnicity (20%), South Asian (10%), Chinese (5%), and Black (3%). In 56% of families, at least one parent had a university degree or higher and 8% had a community college diploma. Demographic information was not specified by 38% of

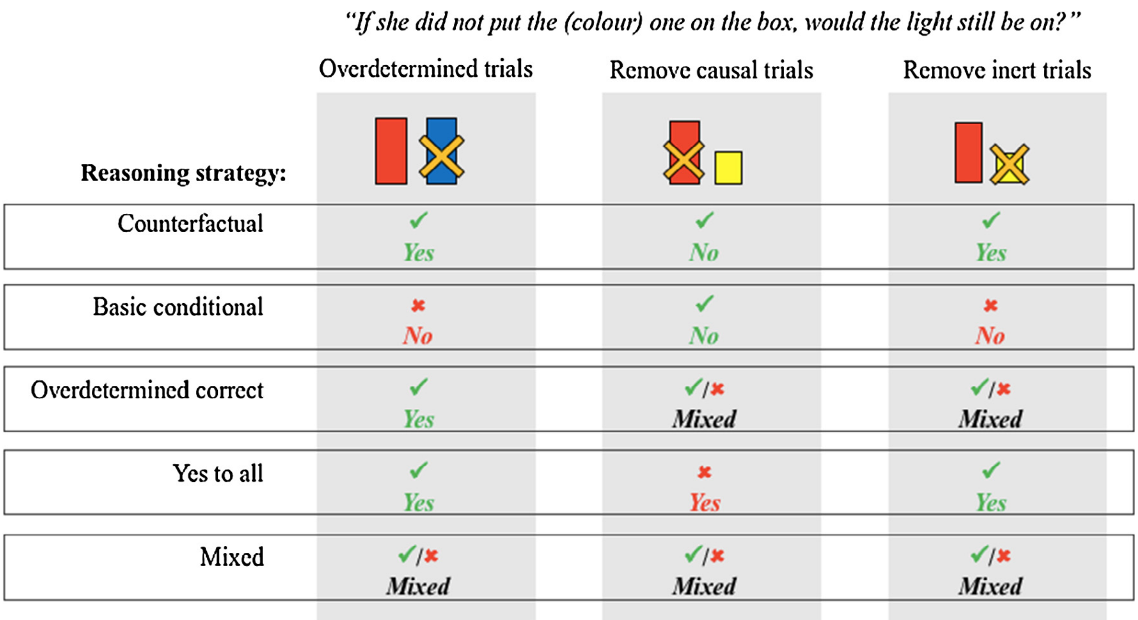

Fig. 2. Responses given by each of the reasoning strategies across trial types. Checkmarks/Xs indicate whether the reasoning strategy gives the correct or incorrect answer. *Yes/No/Mixed* refers to the responses children using each reasoning strategy give in response to the test question. In this example, causal blocks are red and blue, inert block is yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Percentage of children displaying each of the reasoning strategies. Differences from chance marked by asterisks. Chance levels vary by reasoning strategy and are displayed in the first row, according to the number of response patterns that followed each reasoning strategy. (Totals do not sum to 1, because the 27th response pattern was to respond incorrectly to all questions. No children showed this pattern.)

	Counterfactual	Basic conditional	Overdetermined correct	Yes to all	Mixed
Chance	*3.7%* (1/27)	*3.7%* (1/27)	*26%* (7/27)	*3.7%* (1/27)	*59%* (16/27)
Experiment 1a					
3-year-olds	15.6%*	18.8%**	21.9%	28.1%**	15.6%**
4-year-olds	43.8%**	9.4%	21.9%	6.3%	18.8%**
5-year-olds	65.6%**	0%	25%	0%	9.4%**
Experiment 1b					
4-year-olds	35%**	0%	45%	10%	10%**
5-year-olds	55%**	0%	30%	5%	10%**
Experiment 2					
4-year-olds	45%**	10%	25%	5%	15%**
5-year-olds	60%**	5%	20%	5%	10%**

** $p < .001$.
* $p < .003$.

families.

3.1.2. Design and procedure

The task was identical to that in Experiment 1a, with the exception of the wording of the counterfactual questions. All questions were asked as follows (changes from Experiment 1a underlined): "The light switched on! If she had not put the [colour] one on the box, would the light still have switched on?".

Coding took place as in Experiment 1a for 12/40 participants (30%). Coding agreement was excellent (95.5% agreement), κ = 0.72, $p < .001$.

3.2. Results and discussion

Analyses were carried out as in Experiment 1a. Three participants (2 four-year-olds and 1 five-year-old) answered "yes" to all questions. Again, we report results using children's actual scores and report deviations when analyzed using scores corrected for yes/no biases. In a GEE with trial type as a within-subjects variable, age as a continuous covariate, and score (out of 2) as the dependent variable, the effect of trial type on score ($p = .582$), age on score ($p = .380$), and the trial type by age interaction ($p = .752$) were all non-significant.

Using binomial tests with a Bonferroni-corrected alpha-value of 0.005, the proportion of children scoring 2/2 on all three trial types was significantly greater than expected by chance, all *p*-values < 0.001. On all three trial types, the proportion of 5-year-olds scoring 2/2 was significantly greater than chance, all *p*-values < 0.001. Among 4-year-olds, the proportion scoring 2/2 was significantly above chance on *overdetermined trials*, $p < .001$, *remove causal trials*, $p = .004$, and *remove inert trials*, $p = .001$. Using corrected scores, the proportion of 4-year-olds scoring 2/2 was not significantly above chance on *remove inert trials*, $p = .014$.

As in Exp 1a, the pattern of responses was not consistent with a basic conditional reasoning strategy. No participants displayed this reasoning strategy in the current study. Among both 4- and 5-year-olds, the most common reasoning strategies were counterfactual reasoning, or to answer overdetermined questions correctly (see Table 2). We compared the proportion of children displaying each reasoning strategy to the proportion that would be expected by chance, with a Bonferroni-corrected alpha value of 0.005 (0.05/10). Significantly more 4- and 5-year-olds were classified as counterfactual reasoners than predicted by chance, $ps < 0.001$. In both age groups, significantly fewer children displayed a mixed response pattern than predicted by chance, $ps < 0.001$.

The results of Experiments 1a and 1b suggest that children can reason counterfactually about overdetermined outcomes by the age of 4. The ability to consider overdetermined outcomes is a valuable litmus test for mature counterfactual thinking, as it pits genuine counterfactual reasoning against basic conditional reasoning. The results of the study suggest that, given a clear and simple task, preschoolers can engage in mature counterfactual reasoning.

Although children's performance on this task rules out the possibility that they rely on basic conditional reasoning when considering counterfactual premises at this age, two alternate explanations for children's successful performance remain. First, when asked a counterfactual question, rather than considering the counterfactual possibility that one of the blocks was not placed on the blocks, children could instead have ignored the counterfactually-removed block and considered only the function of the remaining block using the information visible on the screen. Second, as with many previous studies of counterfactual reasoning, we cannot be certain that children arrived at correct answers on counterfactual questions by mentally altering *past* events. Instead, they could have considered future hypothetical conditionals of the form, "next time, if she doesn't put the blue block on the box, the light will switch on." Indeed, children succeed earlier on tasks measuring future hypothetical reasoning than analogous tests of counterfactual reasoning (Beck et al., 2006). Physical causation tasks seem to be particularly prone to this issue given that these causes are likely to be considered deterministic and therefore one can be relatively certain that *next time* events will unfold in the same way. Our goal to reduce processing demands, by keeping the final outcome on the screen and removing the temporal delay between antecedents, may have also opened the task up to the possibility that children could make future hypothetical, rather than counterfactual inferences.

What is needed is a task in which children must answer with reference to a sequence of *past* events and not just what is currently visible on the screen. In a second experiment, we used a different version of the blicket-detector task to do just that.

4. Experiment 2

In Experiment 2, we tested a more challenging variant of the blicket-detector task with 4- and 5-year-olds. Blocks were added to the box one at a time with a slight delay between the placement of the first and second block. This was done to ensure children had to recall the sequence of events to arrive at the correct counterfactual representation, and could not rely on static information on the screen that could be subject to a simpler form of reasoning. For example, a child who observed the blue block and then the green block added to the box and was then asked about the removal of the green block would have to recall that the blue block was already on the box and the light was already on at the time the green block was added. This structure is similar to that used in prior research using narrative tasks with overdetermined scenarios (Rafetseder et al., 2013; Nyhout et al., 2017). In the muddy shoes example, children hear about Susie who enters the kitchen first and muddies the floor, and then Max follows and dirties the floor even more. When reasoning about the effect of Susie removing her shoes, children have to bring to mind the sequence of events and recollect the effect of each individual person. Similarly, in the current task, children had to recall the sequence of events and consider the alternative outcome of one block's removal based on the function of the additional block. The counterfactual test questions were asked with a blank screen in front of the child so that children could not easily answer by only reverting to the function of the remaining block.

Success on this variant of the task would provide further compelling evidence that young children can indeed engage in mature counterfactual reasoning, because they had to answer with reference to the *past* sequence of events, and could not answer by considering the problem as a future hypothetical. Because the task places demands on children's memory and temporal reasoning, we asked memory questions before asking each counterfactual question to ensure they remembered the sequence of events they were asked to reason about.

4.1. Method

4.1.1. Participants

The final sample included 40 children (M = 5.05, SD = 0.54 years, 22 girls): 20 4-year-olds (M = 4.58, SD = 0.23, range = 4.05–4.95 years) and 20 5-year-olds (M = 5.53, SD = 0.25, range = 5.08–5.90 years). Five additional children were tested but excluded due to insufficient English language exposure (n = 3) and inattention (n = 2). Participants were recruited and tested at a science museum (n = 30) or in our university laboratory (n = 10). Families reported their ethnicity as mixed ethnicity (38%), White (29%), South Asian (12%), Chinese (12%), Latin American (6%), and other ethnicity (3%). In 85% of families, at least one parent had a university degree or higher, 9% had a community college diploma, and 6% had a high school diploma. Demographic information was not specified by 24% of families.

4.1.2. Design and procedure

Children were shown a video using the same blicket-detector as in Experiment 1. In this case, the actor used only 3 blocks. Two causal blocks (blue and green tall blocks) were on the left side of the box from the child's view, and one inert block (red short block) was on the right.

On *causal test trials*, the actor demonstrated each block individually. As in Experiment 1, the experimenter asked the child about the function of each block one-by-one and then asked the child, "Can you remind me, which blocks made the lights switch on?" and "Which block did not make the lights switch on?". All participants answered these questions correctly.

On *counterfactual test trials*, unlike in Experiment 1, the actor placed one block on the box, and then placed a second block on the box. The lights always activated when the first (or only) causal block was placed on the box and remained on for the entirety of the trial. Each child saw a total of 6 trials. After the actor placed the blocks on the box and a black screen was displayed, the experimenter paused the video. She first asked the child 2 memory questions: "Which block did she put on the box first?" and "Which block did she put on the box next?". The experimenter corrected children who answered either of these questions incorrectly by replaying the trial. Incorrect responses were rare, with children answering 35 out of a total of 480 memory questions incorrectly (7.3%). The experimenter then reiterated the order of the blocks' placement (e.g., "So she put the blue block and then the green block on the box."). This section was included to ensure that children remembered which blocks were used and the order they were placed in to reduce the possibility that incorrect responses to counterfactual questions would be due to memory failures. The experimenter then asked the child a counterfactual question about the removal of the *second block*. In contrast to some previous studies (e.g., Rafetseder et al., 2013), we chose to ask questions only about the second event for a few reasons. Asking about the first event would have created a problem with the counterfactual question. Unlike in previous studies in which the overdetermined events had an additive effect (e.g., a floor getting muddy and muddier), the current study used a binary outcome (i.e., on or off) and therefore the second causal block on the overdetermined trials did not have an effect on the outcome. Because the second block did not switch the lights on, the answer to the question about the

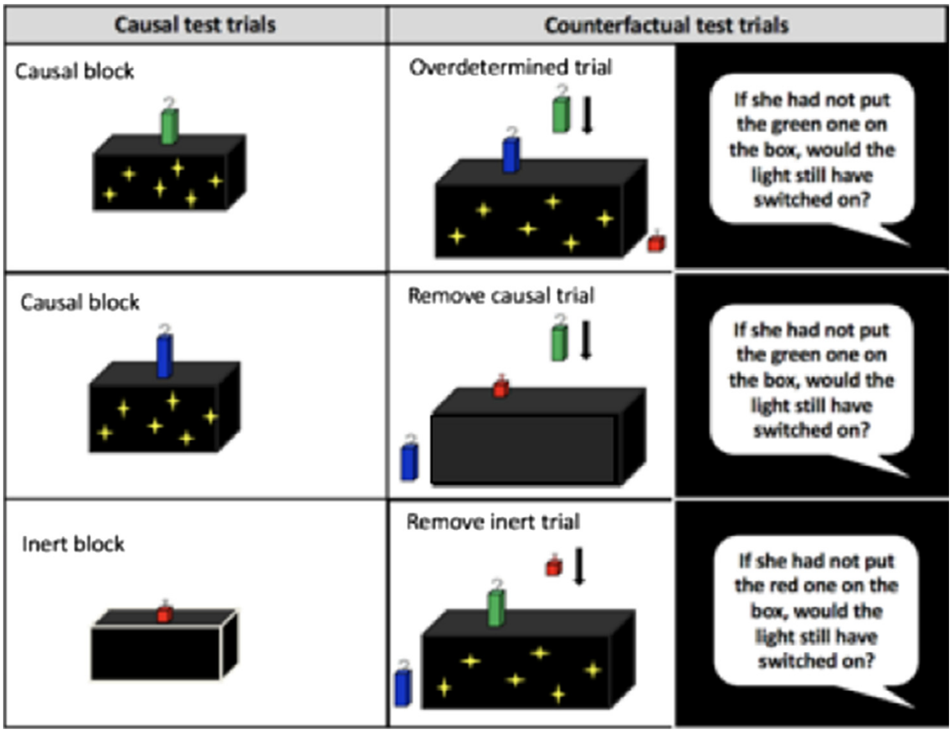

Fig. 3. Schematic of causal and counterfactual trials used in Experiment 2.

removal of the first block is unclear. Nyhout et al. (2017) directly compared children's responses to counterfactual questions about the removal of the first versus second event and did not find a significant difference overall in accuracy. Question wording was identical to Experiment 1b. There were again 3 trial types, and each child received 2 trials of each type, depending on the combination and order in which the blocks were placed on the box: (1) *overdetermined:* first a causal block, then another causal block (e.g., green, then blue), (2) *remove inert:* first a causal block, then an inert block (e.g., green, then red), or (3) *remove causal:* first an inert block, then a causal block (e.g., red, then green). Trial order was counterbalanced in 4 pseudo-random orders, with the requirement that trials of the same type not appear back-to-back. A schematic of trial types is presented in Fig. 3.

Reliability coding took place as in Experiments 1a and 1b for 13/40 participants (32.5%). Coding agreement was excellent (97.3% agreement), $\kappa = 0.87$, $p < .001$.

4.2. Results and discussion

Results were analyzed as in Experiments 1a and 1b. Two participants (1 4-year-old, 1 5-year-old) answered "yes" to all questions and 3 (2 4-year-olds, 1 5-year-old) answered "no" to all questions. On a GEE with trial type as a within-subjects variable, age as a continuous covariate, and score (out of 2) as the dependent variable, the effect of trial type on score ($p = .992$), age on score ($p = .445$), and the trial type by age interaction ($p = .931$) were all non-significant.

Table 1 displays the percentage of children scoring 2 out of 2 on each trial type. We again compared the proportion of children scoring 2 out of 2 to a chance distribution of 0.25 using binomial tests, with a Bonferroni-corrected alpha-value of 0.005. Overall, the proportion of children scoring 2/2 was significantly greater than expected by chance for all trial types, both when collapsing across age and looking at 4- and 5-year-olds separately on each trial type. In all cases, $p < .001$, with the exception of 4-year-olds' performance on *remove inert* trials, where $p = .001$.

As in Experiments 1a and 1b, children's response pattern was not consistent with basic conditional reasoning. Most children showed a counterfactual reasoning response pattern, and this was the overwhelming pattern among 5-year-olds. Reasoning strategy was slightly more mixed among 4-year-olds, but the majority showed either counterfactual reasoning or responded correctly to both overdetermined trials (see Table 2). We again compared the proportion of children displaying each reasoning strategy to chance, using binomial tests and a Bonferroni-corrected alpha value of 0.005. Results mirrored those in Experiment 1b, with significantly more 4- and 5-year-olds classified as counterfactual reasoners and fewer displaying a mixed response pattern than predicted by chance, *ps* < 0.001.

In Experiment 2, children were required to recall the preceding events, contemplate the removal of the second of these events and consider its effect on the outcome. These results indicate that children's good performance in Experiments 1a and 1b was not due to their reliance on basic conditional reasoning or a simpler heuristic, such as using the visual information to imagine a future hypothetical. Even when a temporal delay was introduced between the placement of the 2 blocks and a blank screen was displayed when children were asked counterfactual questions, 4- and 5-year-olds demonstrated mature counterfactual thinking.

5. General discussion

In the present studies, we asked whether preschoolers could reason counterfactually by respecting the nearest possible world constraint, which stipulates that the reasoner must change only those features that are causally dependent on the counterfactual antecedent and hold all else constant (Edgington, 2011; Rafetseder et al., 2010). We presented children with a task with a clear causal structure that has been used extensively in studies of children's causal learning (Gopnik & Sobel, 2000). Four- and 5-year-olds answered a range of different counterfactual questions with a high degree of accuracy, including questions about overdetermined events, and about single-cause events in which either the removal of the causal or inert block had to be contemplated. In previous studies, children have failed on structurally similar, but more complex physical causation tasks (McCormack et al., 2018) or narrative-based tasks (Nyhout et al., 2017; Rafetseder et al., 2010;

2013) until middle to late childhood.

Three-and-a-half year-olds' performance was chance-like across all trial types when the more conservative corrected scores are taken, suggesting that before the age of four, children may not reason counterfactually in an adult-like way. However, the grammar of counterfactual questions may have been too complex for the youngest children (Kuczaj & Daly, 1979), even with the relatively simpler grammar of the counterfactual question in Experiment 1a. Four-year-olds answered counterfactual questions on overdetermined trials with a high degree of accuracy, but showed more variability in performance on single-cause trials, a finding that on first glance may seem inconsistent with previous results (Rafetseder et al., 2013; Rafetseder & Perner, 2018). This may have been because single-cause trials required the child to recall two different functions (i.e., causal and inert), which may be more demanding on their memory than recalling that two blocks had the same function. In contrast, in previous narrative-based tasks (Rafetseder et al., 2013; Rafetseder & Perner, 2018) the reasoner had to recall only which events happened, but not the causal efficacy of those events. In Experiment 2, when we provided additional memory supports, 4-year-olds' performance was above chance on all trial types, suggesting that their difficulty in Experiment 1b may have stemmed from memory errors. (Although there was some variation in 4- and 5-year-olds' performance across the three experiments, it should be noted that these differences were not significant.) Three-year-olds' performance may also have been impeded by these memory demands. Although their performance did not exceed chance on any trials, they performed slightly better on overdetermined than single-cause trials (66% scoring 2/2 on overdetermined trials compared to 44% on remove inert trials). Overall, three-year-olds did not appear to employ a counterfactual reasoning strategy compared to the alternative forms of reasoning we considered, including basic conditional reasoning or answering "yes" to all questions.

The findings of the present study raise several questions about the reasons for children's difficulty in previous studies, but they suggest that it was not due to immaturity in their counterfactual reasoning abilities. Although we have highlighted the importance of clarity of causal structure throughout this paper, we consider other possible explanations for children's earlier success in the current studies compared to previous studies, many of which may drive the child's representation of causal structure.

In many previous studies, children were required to make an *inference* about the state of the world given a counterfactual antecedent (McCormack et al., 2018; Nyhout et al., 2017; Rafetseder et al., 2013). For example, children needed to infer the state of the floor if Susie had taken her shoes off but Max had left his on, given that they never saw the floor dirtied by Max alone. In contrast, children in the present studies could access a relevant representation from memory to answer counterfactual questions (e.g., remember the state of the box after the blue block was added, but before the red block was added in Exp 2). However, this difference in reliance on inference vs. memory cannot fully account for the earlier success children showed in our study, as children failed to reason counterfactually even in some previous studies in which a relevant memory representation was available (Rafetseder et al., 2010; Rafetseder & Perner, 2010). For example, Rafetseder and Perner (2010) created scenarios in which a doctor was in an unusual (e.g., the park) or canonical location (e.g., the hospital) when he was called to an emergency at a swimming pool. When asked where he would be if he had not been called to the pool, children engaging in counterfactual reasoning responded by saying he would be in the park (presumably by recalling the previous situation) and those engaging in basic conditional reasoning responded that he would be at the hospital. These results suggest that the default among children whose reasoning is non-counterfactual is not to undo an event in memory to access a relevant prior representation, but to rely on general knowledge. Thus, inferencing vs. remembering does not appear to drive the difference in results between our study and previous ones. Given these and other findings, we also do not take reasoning that makes use of representations in memory to be non-counterfactual. We expect that in many everyday cases, counterfactual reasoning involves accessing a prior representation rather than simulating one anew.

The majority of previous studies of children's counterfactual reasoning have used stories involving *agents* as stimuli. We consider here several potentially relevant differences between these story-based tasks and the current task. First, whereas previous studies have focused on causes involving agents, the current study featured physical causes. Children's understanding of physical causes emerges earlier (Baillargeon, 2002) than their understanding of causes in other domains including psychology (Flavell, Green, & Flavell, 1995) and biology (Gelman & Wellman, 1991). This earlier causal understanding of the physical domain may mean that children are able to answer counterfactual questions earlier (Sobel, 2011), perhaps because they have a more accurate representation of the underlying causal structure. Certain types of causal reasoning involving agents are present in early childhood (e.g., desire reasoning), and children may show an earlier ability to reason counterfactually in these cases (Sobel, 2011).

Moreover, one can be more confident when considering counterfactuals about physical causes because causes in physical systems are typically deterministic, whereas those involving agents are more likely to be probabilistic. As Glymour (2007) has argued, "[t]he more human action is involved, the more indeterministic things seem." (p. 231). Consistent with this argument, Strickland, Silver, and Keil (2017) found that adults judged physical and psychological causes as having different causal structures across a range of tasks. In particular, physical events were more likely than psychological events to be seen as deterministic, linear causal chains. This bias to consider certain causal systems as more or less deterministic may also influence the ease with which individuals compute alternative representations when asked to reason counterfactually. Preschoolers, too, consider physical causes to be deterministic and infer hidden causes when a known cause acts indeterministically (Schulz & Sommerville, 2006). If one represents the relation between a set of variables to be deterministic, then answering a counterfactual question should be straightforward compared to cases where one expects that, given a slightly different set of circumstances, the relationship between a set of variables could have changed. A related distinction is Woodward's (2006) notion of sensitive versus insensitive causes. Sensitive causal relations are those that, in the face of various counterfactual changes, would not continue to hold, whereas insensitive causal relations would be maintained even with various types of departures from reality. Causal relations involving agents (as in Nyhout et al., 2017; Rafetseder et al., 2010; 2013) may be seen as more sensitive than physical causal relations (as in the present study).

Importantly, none of the explanations we have outlined undermine the claim that children in the present study demonstrated mature counterfactual reasoning. Future research may examine influences on children's counterfactual reasoning, including different domains (e.g., physical causes vs. agents), and causes that are deterministic or not.

A final consideration is what the findings of the present studies tell us about the relation between causal and counterfactual reasoning. Significant debate exists among both philosophers and psychologists, though a full explication of this debate is beyond the scope of the current paper. One group of theories suggests that counterfactuals are implicated when drawing causal inferences, either by acting as input to causal inferences (e.g., Harris et al., 1996; Lewis, 1973; Mackie, 1974) or as a corollary of causal inferencing (e.g., Gopnik & Schulz, 2007; Pearl, 2000; Schulz, Gopnik, & Glymour, 2007; Woodward, 2003), such that when one draws a causal inference (e.g., X causes Y), one commits to the counterfactual (e.g., a change to X would lead to a change to Y). In contrast, others propose that causation is primary and that counterfactual reasoning depends on causal inferences (e.g., Edgington, 2011) or on domain-specific causal knowledge (Sobel, 2011). The evidence to date on the relation between causal and counterfactual reasoning is mixed, with some research indicating a link between the two (e.g., Harris et al., 1996; Sobel, 2004), and other research finding limited involvement for counterfactual reasoning on only certain types of causal reasoning tasks (e.g., Frosch, McCormack, Lagnado, & Burns, 2012; McCormack, Frosch, & Burns, 2011; McCormack, Frosch, Patrick, & Lagnado, 2015).

The fact that 3-year-olds in Exp 1a did not answer counterfactual questions correctly while clearly understanding the causal structure of the task may, on first glance, suggest a dissociation between causal and counterfactual reasoning. Future research is needed to better specify the nature of the relation between causal and counterfactual reasoning early in development. So far, most of the research on this relation has been conducted with adults (e.g., Mandel & Lehman, 1996; Spellman & Mandel, 1999) and the proposals made for younger children are mostly theoretical (e.g., Buchsbaum et al., 2012) and not yet strongly supported by the existing evidence. What is the developmental trajectory of causal and counterfactual reasoning? We do not yet know whether the differences in performance between younger and older children in the present study are best explained by differences in their causal representations, their counterfactual reasoning abilities, processing limitations, or some combination.

6. Conclusion

The current results provide strong support for the claim that children can engage in mature counterfactual reasoning early in development. When presented with a sufficiently simple and clear physical causation task, children as young as 4 demonstrated mature counterfactual reasoning – an ability previous results suggested did not emerge until middle to late childhood. We have considered several possible reasons for children's earlier success on this task than others, and many of these possibilities represent interesting directions for future research.

Acknowledgments

We are grateful to the children and parents in the Greater Toronto Area who participated in this research, and the Ontario Science Centre for facilitating data collection. We thank Hilary Sweatman, Lynn Nguyen, and Noshin Ullah for assistance with participant recruitment, data collection, and coding, Krisma Saini and P. Vaunam Venkadasalam for acting in videos.

Funding

This work was supported by a Postdoctoral Fellowship Award from the Social Sciences and Humanities Research Council of Canada to A. Nyhout, and a grant from the Natural Sciences and Engineering Research Council of Canada to P. A. Ganea.

Appendix A. Supplementary material

The supplementary data available on the Open Science Framework: https://osf.io/j5ht9/?view_only=af761c166d0742f5a894ab207c00097d. Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2018.10.027.

References

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: A summary in eight lessons. In U. InGoswami (Ed.). *Handbook of childhood cognitive development* (pp. 47–83). Oxford, England: Blackwell.

Beck, S. R. (2016). Counterfactuals matter: A reply to weisberg & gopnik. *Cognitive Science, 40*, 260–261. https://doi.org/10.1111/cogs.12242.

Beck, S., & Riggs, K. (2014). Developing thoughts about what might have been. *Child Development Perspectives, 8*(3), 175–179. https://doi.org/10.1111/cdep.12082.

Beck, S., Robinson, E., Carroll, D., & Apperly, I. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77*(2), 413–426. https://doi.org/10.1111/j.1467-8624.2006.00879.x.

Buchsbaum, D., Bridgers, S., Skolnick Weisberg, D., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of The Royal Society B: Biological Sciences, 367*(1599), 2202–2212. https://doi.org/10.1098/rstb.2012.0122.

Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology, 67*, 7.1-7.23. https://doi.org/10.1146/annurev-psych-122414-033249.

Edgington, D. (2011). Causation first: Why causation is prior to counterfactuals. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.). *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology*. Oxford, UK: Oxford University Press.

Epstude, K., & Roese, N. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review, 12*(2), 168–192. https://doi.org/10.1177/1088868308316091.

Flavell, J. H., Green, F. L., & Flavell, E. R. (1995). Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development, 60*.

Frosch, C., McCormack, T., Lagnado, D., & Burns, P. (2012). Are causal structure and intervention judgments inextricably linked? A developmental study. *Cognitive Science, 36*(2), 261–285. https://doi.org/10.1111/j.1551-6709.2011.01208.x.

Gelman, S., & Wellman, H. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38*(3), 213–244. https://doi.org/10.1016/0010-0277(91)90007-q.

German, T., & Nichols, S. (2003). Children's counterfactual inferences about long and short causal chains. *Developmental Science, 6*(5), 514–523. https://doi.org/10.1111/1467-7687.00309.

Glymour, C. (2007). Learning the structure of deterministic systems. In A. Gopnik, & L. Schulz (Eds.). *Causal learning: Psychology, philosophy, and computation*. Oxford: Oxford University Press.

Gopnik, A., & Schulz, L. (2007). *Causal learning: Psychology, philosophy, and computation.* Oxford University Press.

Gopnik, A., & Sobel, D. (2000). Detecting blickets: How young children use knowledge of novel causal powers in categorization and induction. *Child Development, 71*, 1205–1222. https://doi.org/10.1111/1467-8624.00224.

Gopnik, A., Sobel, D., Schulz, L., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology, 37*(5), 620–629. https://doi.org/10.1037//0012-1649.37.5.620.

Harris, P., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition, 61*(3), 233–259. https://doi.org/10.1016/s0010-0277(96)00715-9.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry, 31*(2), 231–270. https://doi.org/10.1162/002438900554352.

Kuczaj, S., & Daly, M. (1979). The development of hypothetical reference in the speech of young children. *Journal of Child Language, 6*(3), 563–579. https://doi.org/10.1017/S0305000900002543.

Leahy, B., Rafetseder, E., & Perner, J. (2014). Basic conditional reasoning: How children mimic counterfactual reasoning. *Studia Logica, 102*, 793–810. https://doi.org/10.1007/s11225-013-9510-7.

Lewis, D. (1973). *Counterfactuals.* Oxford, UK: Basil Blackwell.

Mackie, J. L. (1974). *The cement of the universe: A study of causation.* Oxford: Oxford University Press.

Mandel, D., & Lehman, D. (1996). Counterfactual thinking and ascriptions of cause and preventability. *Journal of Personality And Social Psychology, 71*(3), 450–463. https://doi.org/10.1037//0022-3514.71.3.450.

McCormack, T., Butterfill, S., Hoerl, C., & Burns, P. (2009). Cue competition effects and young children's causal and counterfactual inferences. *Developmental Psychology, 45*(6), 1563–1575. https://doi.org/10.1037/a0017408.

McCormack, T., Frosch, C., & Burns, P. (2011). The relationship between children's causal and counterfactual judgments. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.). *Understanding counterfactuals, understanding causation: issues in philosophy and psychology*. Oxford, UK: Oxford University Press.

McCormack, T., Frosch, C., Patrick, F., & Lagnado, D. (2015). Temporal and statistical information in causal structure learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(2), 395–416. https://doi.org/10.1037/a0038385.

McCormack, T., Ho, M., Gribben, C., O'Connor, E., & Hoerl, C. (2018). The development of counterfactual reasoning about doubly-determined events. *Cognitive Development, 45*, 1–9. https://doi.org/10.1016/j.cogdev.2017.10.001.

McCormack, T., Simms, V., McGourty, J., & Beckers, T. (2013). Encouraging children to think counterfactually enhances blocking in a causal learning task. *Quarterly Journal of Experimental Psychology, 66*(10), 1910–1926. https://doi.org/10.1080/17470218.2013.767847.

Nyhout, A., Henke, L., & Ganea, P. (2017). Children's counterfactual reasoning about causally overdetermined events. *Child Development*. https://doi.org/10.1111/cdev.12913.

Pearl, J. (2000). Comment. *Journal of the American Statistical Association, 95*(450), 428–431. https://doi.org/10.1080/01621459.2000.10474213.

Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: developing a sense of "Nearest Possible World". *Child Development, 81*(1), 376–389. https://doi.org/10.1111/j.1467-8624.2009.01401.

Rafetseder, E., & Perner, J. (2010). Is reasoning from counterfactual antecedents evidence for counterfactual reasoning? *Thinking & Reasoning, 16*, 131–155. https://doi.org/10.1080/13546783.2010.488074.

Rafetseder, E., & Perner, J. (2014). Counterfactual reasoning: Sharpening conceptual distinctions in developmental studies. *Child Development Perspectives, 8*(1), 54–58. https://doi.org/10.1111/cdep.12061.

Rafetseder, E., & Perner, J. (2018). Belief and counterfactuality. *Zeitschrift für Psychologie, 226*, 110–121. https://doi.org/10.1027/2151-2604/a000327.

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From

childhood to adulthood. *Journal of Experimental Child Psychology, 114*(3), 389–404. https://doi.org/10.1016/j.jecp.2012.10.010.
Riggs, K., Peterson, D., Robinson, E., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality?. *Cognitive Development, 13*(1), 73–90. https://doi.org/10.1016/s0885-2014(98)90021-1.
Robinson, E. J., & Beck, S. R. (2000). What is difficult about counterfactual reasoning? In P. Mitchell, & K. Riggs (Eds.). *Children's reasoning and the mind* (pp. 101–120). Hove: Psychology Press.
Schulz, L., Gopnik, A., & Glymour, C. (2007). Preschool children learn about causal structure from conditional interventions. *Developmental Science, 10*(3), 322–332. https://doi.org/10.1111/j.1467-7687.2007.00587.x.
Schulz, L., & Sommerville, J. (2006). God does not play dice: causal determinism and preschoolers' causal inferences. *Child Development, 77*(2), 427–442. https://doi.org/10.1111/j.1467-8624.2006.00880.x.
Sobel, D. M. (2004). Exploring the coherence of young children's explanatory abilities: Evidence from generating counterfactuals. *British Journal of Developmental Psychology, 22*, 37–58. https://doi.org/10.1348/026151004772901104.
Sobel, D. M. (2011). Domain-specific causal knowledge and children's reasoning about possibility. In C. Hoerl, T. McCormack, & S. R. Beck (Eds.). *Understanding counterfactuals, understanding causation: Issues in philosophy and psychology* (pp. 123–146). Oxford, UK: Oxford University Press.
Sobel, D., Tenenbaum, J., & Gopnik, A. (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science, 28*(3), 303–333. https://doi.org/10.1207/s15516709cog2803_1.
Spellman, B., & Mandel, D. (1999). When possibility informs reality. *Current Directions in Psychological Science, 8*(4), 120–123. https://doi.org/10.1111/1467-8721.00028.
Strickland, B., Silver, I., & Keil, F. C. (2017). The texture of causal construals: Domain-specific biases shape causal inferences from discourse. *Memory & Cognition, 45*(3), 442–455. https://doi.org/10.3758/s13421-016-0668-x.
Weisberg, D., & Gopnik, A. (2013). Pretense, counterfactuals, and bayesian causal models: Why what is not real really matters. *Cognitive Science, 37*(7), 1368–1381. https://doi.org/10.1111/cogs.12069.
Woodward, J. (2006). Sensitive and insensitive causation. *Philosophical Review, 115*(1), 1–50. https://doi.org/10.1215/00318108-115-1-1.
Woodward, J. (2003). *Making things happen: A theory of causal explanation.* Oxford: Oxford University Press.