WARWICK
THE UNIVERSITY OF WARWICK

Manuscript version: Author's Accepted Manuscript
The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/130728

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

warwick.ac.uk/lib-publications

Causality Influences Children's and Adults' Experience of Temporal Order

Running Title: Development of Causal Reordering

Emma C Tecwyn [1*], Christos Bechlivanidis [2], David A Lagnado [2], Christoph Hoerl [3], Sara Lorimer [4], Emma Blakey [5], Teresa McCormack [4], & Marc J Buehner [1]

1 Cardiff University, School of Psychology

2 University College London, Department of Psychology

3 University of Warwick, Department of Philosophy

4 Queen's University Belfast, School of Psychology

5 The University of Sheffield, Department of Psychology

Author note

This research was funded by Leverhulme Trust Research Project Grant RPG-2015-267

[*] Correspondence: emma.tecwyn@bcu.ac.uk

Emma C Tecwyn is now at Department of Psychology, Birmingham City University

The authors have no conflict of interest to declare.

© 2019, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI: 10.1037/dev0000889

Abstract

Although it has long been known that time is a cue to causation, recent work with adults has demonstrated that causality can also influence the experience of time. In *causal reordering* (Bechlivanidis & Lagnado, 2013, 2016) adults tend to report the causally consistent order of events, rather than the correct temporal order. However, the effect has yet to be demonstrated in children. Across four pre-registered experiments, 4- to 10-year-old children (N=813) and adults (N=178) watched a 3-object Michotte-style ‘pseudocollision’. While in the canonical version of the clip object A collided with B, which then collided with object C (order: ABC), the pseudocollision involved the same spatial array of objects but featured object C moving before object B (order: ACB), with no collision between B and C. Participants were asked to judge the temporal order of events and whether object B collided with C. Across all age groups, participants were significantly more likely to judge that B collided with C in the 3-object pseudocollision than in a 2-object control clip (where clear causal direction was lacking), despite the spatiotemporal relations between B and C being identical in the two clips (Experiments 1—3). Collision judgements and temporal order judgements were not entirely consistent, with some participants—particularly in the younger age range—basing their temporal order judgements on spatial rather than temporal information (Experiment 4). We conclude that in both children and adults, rather than causal impressions being determined only by the basic spatial-temporal properties of object movement, schemata are used in a top-down manner when interpreting perceptual displays.

Keywords: causality, causal perception, cognitive development, Michottean launching, temporal cognition, time perception

Causality Influences Children's and Adults' Experience of Temporal Order

The ability to learn about and represent causal relations is fundamental to our ability to navigate and understand the world as it enables us to interpret, explain and thus predict, events in our environment. A large body of research suggests that from a young age, children represent causal structures and use this information to guide their inferences and behaviour (see Muentener & Bonawitz, 2017; Sobel & Legare, 2014 for recent reviews). There is evidence that causal knowledge contributes to the development of children's cognitive skills in a variety of domains (e.g., physical reasoning, Baillargeon, 2004; moral reasoning, Hamlin, 2013; generating explanations, Legare, 2012), thus demonstrating that causality plays a central role in our experience of the world from early in life.

It has long been known that temporal cues strongly influence people's causal judgements. Both adults' (e.g., Buehner & May, 2003; Lagnado & Sloman, 2006) and children's (e.g., Bullock & Gelman, 1979; McCormack et al., 2015; Mendelson & Shultz, 1976; Rankin & McCormack, 2013; Schlottmann et al., 1999) causal judgements show sensitivity to the principles of temporal priority (causes must precede their effects) and temporal contiguity (causally related events typically occur close together in time). More recently, it has become apparent that the relations between time and causality are in fact bidirectional—just as temporal cues influence our causal judgements, causal beliefs, in turn, influence the experience of time. Empirically, this influence of causal beliefs on temporal experience has been demonstrated in studies of two effects: *causal binding* and *causal reordering*. Studies of causal binding have shown that if one event A is believed to be the cause of another event B, the interval between the two events is perceived as shorter in duration than the same objective interval where the two events are not causally linked (Buehner 2012; 2015; Buehner & Humphreys, 2009). This represents a quantitative shift in the perception of the temporal duration of an interval, such that causally-related events are drawn towards one another, or ‘bound’ together in time.

A small number of recent studies have also demonstrated that causal beliefs can influence not only the subjective interval between events but also the temporal order in which the events are perceived to occur. In causal reordering (Bechlivanidis & Lagnado, 2013; 2016) the temporal order in which events are perceived to have occurred is reversed, so that the experienced order of events is in line with causality. That is, if participants have a background belief that A is a cause of B, they are likely to report that A happened before B even when shown a sequence of events in which B happened first. In the first study to demonstrate causal reordering, participants interacted with an on-screen ‘physics world’ consisting of animated objects with different properties. After learning the properties of the objects and the causal relations between them, participants watched a clip that violated the learned causal order of events (i.e., if they had learned that A caused B, they saw a clip in which B happened before A). Participants were significantly more likely to report that events occurred in the order consistent with their causal beliefs than the objective temporal order (Bechlivanidis & Lagnado, 2013).

Further evidence that causal beliefs influence adults’ experience of the temporal order of events comes from a study by Desantis and colleagues (2016). In this study participants watched a random-dot-kinematogram (RDK) on a computer screen and learned that pressing one key (e.g., left) caused the RDK motion to become briefly coherent in one direction (e.g., upwards), and pressing a different key (e.g., right) led to coherent motion in the opposite direction (e.g., downwards). Having learned this association, in a critical test phase, participants continued to execute keypresses, but sometimes the coherent motion of the RDK occurred *before* the keypress. For these trials, participants were more likely to (incorrectly) report that the motion occurred after their keypress when coherent motion was in the expected (i.e. learnt) direction, compared with when it was in the unexpected, incongruent direction. This finding is indicative of causal reordering because participants apparently perceived events to occur in the order that reflected their learned causal beliefs (Desantis et al., 2016).

The above causal reordering studies were based on causal relations that participants learned in an initial training phase. On the basis of this evidence alone, it is not possible to determine whether the reordering effect is dependent on recently learned rules about unfamiliar causes and effects, or whether it might represent a more general phenomenon that occurs in any situation that evokes an impression of causality. In addition, the Desantis et al. (2016) study involved intentional action by the participant, thus the reordering effect found might not be explained solely by causal beliefs (e.g., illusion of control could also play a role). To address these issues, Bechlivanidis and Lagnado (2016) designed a ‘one shot’ experiment that involved showing participants a single brief clip. The clip was based on a Michottean launching event (i.e. a simple collision between horizontally arranged two-dimensional objects), adapted to involve three objects (ABC) instead of the typical two. Crucially, the third object in line (C) moved before the second object in line (B); i.e., the effect occurred *before* its presumed cause (see e.g., Figure 2a). Participants were significantly more likely to report perceiving that the events happened in an order consistent with causation (ABC) than in the objective temporal order (ACB). Participants also tended to (incorrectly) report that B made C move, suggesting that presumed causality—in the form of a collision between B and C—was the basis on which reordering occurred (Bechlivanidis & Lagnado, 2016).

Taken together, these studies provide compelling evidence that adults temporally reorder events in line with their assumptions about causality, regardless of whether those assumptions are the result of recent learning or are based on perceptual cues. However, nothing is currently known about the developmental origins of this phenomenon, despite the potential for developmental research to enhance our understanding of the nature of the links between causal and temporal cognition. Children's causal cognition has been studied extensively (see Muentener & Bonawitz, 2017; Sobel & Legare, 2014 for recent reviews) and even infants show some sensitivity to causality in Michottean launching displays (e.g., Leslie & Keeble, 1987; Mascalzoni et al., 2013; Oakes, 1994; Schlottmann et al., 2002), but whether children's causal impressions are strong and reliable enough to modulate their temporal order perception, as is true for adults, remains an open question.

Research on whether causal beliefs can affect children's temporal perception has so far been limited to a small number of developmental studies of causal binding—the perceived shortening of duration between two events that are believed to be causally related. Cavazzana and colleagues (2014, 2017) investigated the binding effect in 8- to 11-year-old children and adults. In each trial, participants watched letters of the alphabet rapidly flash up on a screen in a random order, and had to report which letter was on the screen when target events occurred. In some trials participants heard two tones (which were causally unrelated to one another) and in other trials participants pressed a key that resulted in a tone (causally related events), with the duration between the pairs of events identical in both cases. The adults' judgements of which letters were on the screen when these target events occurred revealed the classic binding effect—the causally related keypress and tone were perceived as occurring closer together in time compared to the causally unrelated tones. However, the researchers failed to find evidence of causal binding in the children, leading them to conclude that the effect emerges late in development and may be linked to the development of higher-order cognitive processes (Cavazzana, Begliomini, & Bisiacchi, 2014, 2017).

Although Cavazzana et al. concluded that this type of binding was a late-emerging phenomenon, their findings contrast with those of some recent studies using simplified childfriendly tasks. In these tasks, rather than retrospectively reporting the time at which an event occurred, participants either anticipated when they expected a target event (e.g., a rocket on a screen launching) to occur following an initial event (keypress or non-causal signal, Blakey et al., 2018), or gave a categorical estimation of the interval between the two events (Lorimer et al., under review). Children in both of these studies showed a binding effect—they were more likely to perceive the duration between two events to be shorter when there was a causal connection between them (i.e., when the rocket launch was caused by a keypress as opposed to preceded by an arbitrary signal). These findings suggest that susceptibility to causal binding is present in children as young as four years and that the magnitude of the binding effect does not increase developmentally, even into adulthood (Blakey et al., 2018; Lorimer et al., under review). Thus, it appears that, rather than being a late emerging phenomenon as suggested by the results of Cavazanna et al., causal binding reflects a fundamental way in which cognition shapes perception, and, at least from four years, is not modulated either by increased experience of causal relations or higher-order cognitive/reasoning processes that are known to change developmentally.

Causal binding and reordering effects are both examples of causal beliefs influencing temporal experience, suggesting that the relationship between time and causality is bidirectional. It thus seems intuitively plausible that the emergence of these effects may follow the same developmental trajectory. However, it is difficult to generate developmental predictions about causal reordering effects based on studies of causal binding, because there are no detailed models of these effects that assume they have a common basis (indeed, there is considerable disagreement over the mechanisms underpinning causal binding, e.g., Borhani, Beck, & Haggard, 2017; Buehner, 2012; Faro, McGill, & Hastie, 2013; Merchant & Yarrow, 2016). Nevertheless, the recent studies on causal binding in children help motivate an examination of whether causal reordering is also observable in children. The aim of the present study was to investigate for the first time whether children as young as four years are susceptible to the causal reordering effect, and if so, whether and how this changes across development. If we find evidence of reordering from a young age, this would provide further evidence for an early-developing bidirectional relation between time and causality, where causality already plays a critical role in children's interpretation of the environment, including its temporal features. On the other hand, if children do not reorder, or if susceptibility to reordering increases with age, this would suggest that the role of causal beliefs in interpreting temporal order develops slowly, perhaps as a result of increasing experience with causal systems.

The Michottean launching paradigm used by Bechlivanidis and Lagnado (2016) provides a very useful context in which to examine this issue, because the task does not involve children having to acquire familiarity with a new set of causal relations or make effortful causal inferences. While there is long-standing debate over how best to interpret the infancy data which has used Michottean-type tasks (Saxe & Carey, 2006; Cohen & Amsell, 1998; Schlottmann, 2000; White, 2017), we can be confident that even preschoolers have a distinctive impression of physical causation when they see prototypical launch events (Schlottmann, Cole, Watts, & White, 2013; Schlottmann, Allan, Linderoth, & Hesketh, 2002). Although in some circumstances young children are somewhat more tolerant than adults in ascribing causation to launching events that deviate from the prototypical launching sequence in most respects their explicit causal judgements are remarkably similar to those of adults (Schlottmann et al., 2013; see also Bechlivanidis, Schlottmann & Lagnado (2019) for recent evidence that adults are in fact more tolerant of deviation than previously assumed).

General Method

Approval for this study (Experiments 1—4) was granted by Cardiff University School of Psychology Ethics Committee, EC.16.02.09.4448R, 'Time and Causality in Cognitive Development'. All studies were pre-registered and are available at the following links: Experiment1: https://osf.io/nqbtm/, Experiment 2: https://osf.io/vcesk/register/565fb3678c5e4a66b5582f67, Experiment 3: http://aspredicted.org/blind.php?x=z7e5xr; Experiment 4: http://aspredicted.org/blind.php?x=ip226r.

Participants

For each experiment we initially aimed to recruit approximately 30 participants per age group and use a within-subjects design (for the sake of economic use of participants), with participants viewing both of the critical clips (there were two in each experiment, the 3-object pseudocollision and the control clip) in a counterbalanced order, yielding two conditions (pseudocollision first or second). Once we reached this sample size we tested for order effects; specifically, for each age group we tested whether the order in which participants saw the two critical clips influenced their responses for either of our measures (TOJ and CJ). For all four experiments, critical clip order influenced performance for at least one age group on at least one measure (see supplementary Table S1 Figure S1); thus, in each case we switched to a between-subjects design, whereby we proceeded to collect additional data to give approximately 30 participants per age group per condition, and only analysed the first of the two critical clips participants watched. That is, in the analyses reported below, participants contributed data points for either the pseudocollision clip or the control clip.

The exact number of participants per experiment was determined by availability in schools and museums. Specifically, we did not turn away anyone who wanted to participate while we were in a given setting. To enable us to examine performance differences across development and compare children and adults within the same model the child sample for each experiment was divided into multiple age groups.

All participants were tested individually. Adults were either tested in a room at a university (undergraduate students) or at a local science museum (museum visitors). The adults tested at a university received course credit for participating. Children were either tested in a room at their school or at a local science museum and received a sticker for participating.

Materials

All experiments were programmed in Adobe Flex 4.6 and presented to participants on an Acer TravelMate P236 13.3" laptop. Examples of the clips presented in Experiment 1 are depicted in Figures 1 and 2.

Design

All Participants only took part in one of the four experiments. The following variables were randomized across participants: direction of object motion in clips (left to right, right to left); practice clip order; colour of the shapes (which varied between experiments).

Coding and preliminary analyses

For each critical clip we coded participants' responses to (a) the TOJ question (shape selected (A, B, C) and whether it was correct/incorrect) and (b) the CJ question (yes/no and whether it was correct/incorrect). For each experiment we ran preliminary analyses to check for an effect of direction of motion (left-right or right-left) on either of our response variables. As we found no significant influence of motion direction, data were collapsed across this variable for all subsequent analyses.

Experiment 1

In Experiment 1, we modified Bechlivanidis and Lagnado's (2016) Experiment 1 to make it more appropriate for young children. The critical clips were identical in terms of their spatiotemporal features to those used in the original study. However, whereas participants in Bechlivanidis and Lagnado's (2016) experiment were required to order all of the events that occurred via drag and drop, we greatly simplified the response variables to reduce task demands. In the critical clips for our task, participants were asked a single temporal order judgement (TOJ) question ("Which square started moving last?") and a single collision judgement (CJ) question ("Did square B bump into square C, yes or no?" see Method for further details). We also introduced 4 non-causal practice clips (two involving two objects and two involving three objects; Figure 1a—b) that participants watched before viewing the critical clips, to familiarize participants with the type of clip they would be watching and what they should be attending to.

Method

Participants. Our final sample consisted of 61 adults (41 female, 3-object: N = 31, M_{age} =29 years; 2-object: N = 30, M_{age} = 23 years) and 282 children (164 female). An additional four children were tested but excluded because they were inattentive (N = 3) or did not understand the task instructions (N = 1). The child sample was divided into 4 age groups per condition: 4- to 6-year-olds (3-object: N = 35, M_{age} = 5 years 8 months; 2-object: N = 35, M_{age} = 5 years 4 months), 6- to 7-year-olds (3-object: N = 36, M_{age} = 7 years 2 months; 2-object: N = 35, M_{age} = 7 years 0 months), 7- to 9-year-olds (3-object: N = 35, M_{age} = 8 years 8 months; 2-object: N = 35, M_{age} = 8 years 5 months) and 9- to 10-year-olds (3-object: N = 36, M_{age} = 9 years 11 months; 2-object: N = 35, M_{age} = 9 years 9 months).

Procedure. Participants were told that they would watch some short clips of squares moving around on the screen and answer some questions about what they saw. They were told that they would only get to see each clip once so they should make sure to pay attention, and that they would know when each clip was going to start because they would see a 'clock' fill in from white to black (Figures 1 and 2), after which the squares would start to move, which was then demonstrated to them once.

Practice clips. Participants first watched 4 non-causal practice clips (see Figure 1a), and were asked a TOJ question after each clip. At the start of each practice clip the squares were aligned vertically in columns at one side of the screen and they started to move horizontally one at a time, so there was no implied causal connection between the motion onsets of the squares.[1] After each practice clip, participants saw a screen with the squares in their final configuration (i.e., where they ended up after the motion), and were asked a single TOJ question: either, "Which square started moving first?" or "Which square started moving last?" to establish their experience of the motion onset of the squares. These questions were asked in an alternating order across the four practice clips. The rationale for asking both of these questions was to encourage participants to attend to the motion of all of the squares. Given that children may not always accurately interpret the words "before" and "after" until at least 5 years of age (e.g., Blything & Cain 2016; Blything, Davies & Cain, 2015) we deliberately avoided the use of these terms.

Figure 1 about here

Critical clips. The critical clips consisted of a 2-object control clip and a 3-object "pseudocollision" clip (Figure 2) presented in a counterbalanced order. The shapes in the critical clips – which were all squares in Experiment 1 – will henceforth be labelled A, B, and C. At the start of each critical clip the shapes were aligned horizontally. In the 3-object

[1] White (2017) reported strong impressions of causality for an array of four vertically aligned objects that were simultaneously 'launched'. However, the displays used in his study were very different from our practice clips where the objects moved separately and there was no 'launcher' object.

pseudocollision (Figure 2a), square A moved towards square B and stopped adjacent to it; immediately after this, square C started moving away from square B, and after 350 ms, square B started moving away from square A; at no stage did square B make contact with square C. All shapes moved at a speed of 30 mm/s. The 2-object control clip was identical to the 3-object pseudocollision, except that square A was not present (Figure 2b). Critically, the relative onset of motion of squares B and C was exactly the same in both clips.

As in the practice clips the shapes remained in their final positions after each critical clip, and participants were asked a TOJ: “Which square started moving last?” This form of words was used rather than the more straightforward “Which square moved last?” because squares B and C stopped moving simultaneously (and so technically they both moved last). Participants were also asked a collision judgement (CJ) question about shapes B and C: “Did the (e.g.) black square (B) bump into the (e.g.) red square (C), yes or no?” and the experimenter pointed at the relevant squares on the screen as they asked this question. The aim of asking this was to establish whether children had the impression that B had collided with C.

Figure 2 about here

Pre-registered confirmatory analyses. To establish which of the age groups tested were susceptible to causal reordering, for each age group we used Chi-square tests to compare participants’ TOJ and CJ responses in the 2-object control clip and the 3-object pseudocollision (as a reminder, these clips were identical except for the inclusion/exclusion of object A). Where the assumptions for using the chi-square test were not met (i.e., expected values of < 5 in one or more cells) we used Fisher’s Exact Test. If participants were reordering events in line with an impression of causality, we would expect a significantly greater proportion of participants' TOJs and CJs to be accurate in the 2-object control clip than in the 3-object pseudocollision.

Exploratory analyses. To further examine developmental changes in reordering we used binomial logistic regression conducted in R (R Core Team, 2017) to ascertain the effect of age group on the likelihood of responding correctly to (a) the TOJ question and (b) the CJ question for the 3-object pseudocollision. If the models revealed a significant effect of age group, planned pairwise comparisons were conducted with Tukey-adjusted p-values for multiple comparisons, to establish which age groups differed from one another. Correlation between our two measures (TOJs and CJs) was assessed by calculating Phi coefficients, which is a measure of association between two binary variables. Specifically, we were interested to know whether participants who reordered events B and C were more likely to report perceiving a collision between these two objects (and vice versa).

Results

Following Bechlivanidis and Lagnado (2016) and our pre-registered analysis plan, for the following analyses we excluded participants who, following the TOJ question, gave the nonsensical response that square A started moving last. This resulted in the exclusion of 28/132 children (14 4- to 6-year-olds; seven 6- to 7-year-olds; six 7- to 9-year-olds; one 9- to 10-year-old) from the group who contributed data on the 3-object pseudocollision clip. No adults needed to be excluded on this basis.

Practice clips. Performance in the 2-object practice clips ranged from 69% correct responses (4- to 6-year-olds) to 93% correct responses (adults). Performance in the 3-object practice clips ranged from 60% correct responses (4- to 6-year-olds) to 94% correct responses (adults, see Table S2 for full details).

Pre-registered confirmatory analyses. Across all age groups, the majority of participants responded correctly to the TOJ question (that B moved last) in the 2-object control clip (Figure 3a). Participants in all age groups were significantly more likely to respond correctly (say B started moving last) in the 2-object control clip than the 3-object pseudocollision (Chi-square tests: $p < 0.001$ for all, Table 1). Participants in all age groups were also significantly more likely to respond correctly (no) to the CJ question (e.g., "did the green (B) square bump into the red (C) square, yes or no?", see Figure 3b) in the 2-object control clip than the 3-object pseudocollision (Chi-square tests: $p \leq 0.001$ for all, Table 1).

Figure 3 about here

Table 1 about here

Exploratory analyses. Logistic regression revealed that participants' tendency to report the correct order of events (TOJ question) in the pseudocollision was significantly influenced by age group (Wald $\chi^2 = 10.68$, df = 4, $p = 0.030$). Posthoc contrasts with Tukey adjusted *p*-values for multiple comparisons revealed a significant difference between adults and 9- to 10-year-olds (log odds ratio = 1.54, $p = 0.036$), with adults being more likely to respond correctly/less likely to reorder. There were no other significant differences between groups after adjusting for multiple comparisons ($p \geq 0.124$ for all other pairs of age groups, Table S3). Participants' tendency to report perceiving a collision between objects B and C (CJ question) in the pseudocollision was also significantly influenced by age group (Wald $\chi^2 = 10.43$, df = 4, $p = 0.034$). Posthoc contrasts with Tukey adjusted p-values for multiple comparisons revealed a significant difference between 9- to 10-year-olds and 7- to 9-year-olds (log odds ratio = 1.72, $p = 0.038$), with the older children being more likely to perceive a collision. There were no other significant differences between age groups in responses to the CJ question after adjusting for multiple comparisons ($p \geq 0.470$ for all other pairwise comparisons). These patterns of responding with age group as a categorical predictor were in keeping with analyses of child data only when age in years was included as a continuous predictor (see Table S6). TOJs and CJs were significantly associated for the 3-object pseudocollision—participants who reordered events B and C were more likely to report perceiving a collision between those objects (Phi = 0.26, $p = 0.002$, see Table S7 for details per age group).

Discussion

Across all of the age groups tested, participants were significantly more likely to report the correct order of events (say that square B started moving last) in the 2-object control clip than the 3-object pseudocollision clip, despite the relative onset of motion of squares B and C being identical in both clips. The results for the 2-object clip provide evidence that participants of all ages were able to perceptually distinguish the relative onset of motion of squares B and C, as they almost always gave the correct response to the TOJ question in this case. This suggests that participants' TOJs were influenced by the inclusion of square A, which gave the clip clear causal direction. In addition, all participants were significantly less likely to report perceiving contact between objects B and C in the 2-object control clip than the 3-object pseudocollision (i.e, they were more likely to correctly respond "no" to the CJ question in the former), which indicates that the causal impression generated by the pseudocollision was the basis for reordering.

Adults in the present experiment were less likely to reorder than in Bechlivanidis and Lagnado's (2016, Experiment 1) original one-shot study (42% vs. 83% reordering). This difference in performance is probably due to the inclusion of practice trials in the present task. Asking a TOJ question after each practice trial presumably causes participants to focus more on the temporal order of events, so when they get to the critical clips they have a good idea what they should be attending to. In fact, given the long temporal interval (350 ms) between the motion of two objects and the fact that adults were expecting to be asked about the temporal order of events, it is perhaps surprising that we nevertheless still find evidence for reordering in almost half of the adults tested (in contrast, only 6% of adults responses were incorrect in the 3-object practice trials). While 9- to -10-year-olds were more likely to reorder events than adults in the 3-object pseudocollision, and more likely to report perceiving a collision between objects B and C than 7- to 9-year-olds, there was no clear developmental pattern in performance according to either of our measures.

Although the data from Experiment 1 provided some initial evidence that children as young as four years reorder events in line with causal impressions, the fact that a large proportion of participants in the younger age groups gave the response that object A started moving last (41% in our youngest age group) and thus had to be excluded is unsatisfactory. This high level of exclusions makes it impossible to properly determine the developmental trajectory of the reordering phenomenon, as this hangs on how the A-responders would re-distribute between B and C if they did not give the nonsensical A response. Why might participants—specifically, young children—say that A started moving last? Two features of Experiment 1 may have led children to respond in this way. First, while we deliberately avoided the use of the terms “before” and “after” given young children’s well-established difficulties with these terms, it is possible that the question “which square started moving last?” is also rather complex for young children—particularly the combination of “started” and “last”. Second, because we alternated the TOJ question between practice trials, either asking which square moved first *or* which square moved last, it is possible that in some cases children were expecting to be asked about which square moved first (rather than last) in the critical clip, and gave a response to that question instead (though note that if this were true we would expect the same issue to affect the 2-object control clip). In Experiment 2 we addressed both these issues, with the aim of getting a clearer picture of the developmental trajectory of susceptibility to causal reordering.

Experiment 2

In Experiment 2 we again presented participants with a 3-object pseudocollision and a 2-object control clip. However, to prevent participants from responding "A" in the critical TOJ question, object A was a circle, whereas B and C were both squares, and we explicitly asked about the squares (Figure 2a[ii]). Participants were introduced to the different shapes at the start of the task, and they saw a practice clip involving a circle and two squares. To address the other issues that might have contributed to the high levels of A-responding in Experiment 1, we changed the TOJ so that for all clips (practice and critical) participants were asked "Which square moved *first*?" We also reduced the number of practice clips from four to two, as we suspected the extensive practice phase could have contributed to the decreased prevalence of reordering in adults compared to the level reported by Bechlivanidis and Lagnado (2016).

Method

Participants. Our final sample consisted of 63 adults (56 female; 3-object: N = 30, M_{age} = 20 years; 2-object: N = 33, M_{age} = 20 years) and 207 children (127 female), none of whom had participated in Experiment 1. An additional four children were tested but excluded because of a lack of attention (N = 3) or insufficient English language skills (N = 1). The child sample was divided into 3 age groups per condition: 4- to 6-year-olds (3-object: N = 33, M_{age} = 5 years 5 months; 2-object: N = 32, M_{age} = 5 years 4 months), 6- to 8-year-olds (3-object: N = 33, M_{age} = 7 years 4 months; 2-object: N = 32, M_{age} = 7 years 1 month) and 8- to 10-year-olds (3-object: N = 33, M_{age} = 9 years 8 months; 2-object: N = 32, M_{age} = 9 years 1 month).

Materials. The materials were the same as in Experiment 1 except that object A was a circle and we changed the colour of the shapes to blue, orange and grey, as it occurred to us that red-green colour-blindness could have been an issue in Experiment 1.

Procedure. The task instructions were the same as for Experiment 1, with the addition that before viewing the practice clips participants were introduced to the different shapes (square and circle), and children in the youngest age group were asked to name the shapes (their data were excluded if they were unable to).

Practice clips. Participants watched two non-causal practice clips (Figure 1b) in a random order and were asked the same TOJ question after each one: "Which square moved first?"

Critical clips. The 2-object control clip was identical to the clip used in Experiment 1. The 3-object test clip was identical except that object A was a circle instead of a square (Figure 2a[ii]).

Results

Practice clips. Performance in the 2-object practice clip ranged from 71% of participants responding correctly (4- to 6-year-olds) to 87% of participants responding correctly (adults). Performance in the 3-object practice clip ranged from 66% of participants responding correctly (4- to 6-year-olds and 6- to 8-year-olds) to 90% of participants responding correctly (adults, see Table S2 for full details).

Pre-registered confirmatory analyses. Across all age groups, the majority of participants responded correctly to the TOJ question (that C moved first) in the 2-object control clip (Figure 4a). In contrast to Experiment 1, in Experiment 2 there was a clear pattern of decreasing response accuracy to the TOJ question for the 3-object pseudocollision (blue bars of Figure 4a): younger children were more likely to respond correctly than older children and adults when asked "Which square moved first?" Comparisons of TOJ responses between the 2-object and 3-object clips revealed that while 8- to 10-year-olds and adults were significantly more likely to respond correctly in the 2-object clip than the 3-object clip (chi-square tests, *ps* $\leq$ 0.003, Table 1), the 4- to 6- and 6- to 8-year-olds' performance did not differ significantly between the two critical clips (Fisher's Exact Test, *ps* > 0.082). Participants in all age groups were significantly more likely to say square B collided with square C in the 3-object pseudocollision than the 2-object control clip (Figure 4b, Chi-square tests: *ps* $\leq$ 0.002 for all, Table 1).

Figure 4 about here

Exploratory analyses. Logistic regression revealed that participants' tendency to report the correct order of events (TOJ question) in the pseudocollision was significantly influenced by age group (Wald χ^2 = 10.52, df = 3, p = 0.015). After correcting p-values for multiple comparisons (Tukey adjustment) the youngest children were significantly more likely to respond correctly/less likely to reorder than adults (log odds ratio = 1.90, p = 0.038). There were no other significant differences between groups after adjusting for multiple comparisons ($p \geq 0.065$ for all other pairs of age groups, Table S4). Participants' tendency to report perceiving a collision between objects B and C (CJ question) in the 3-object pseudocollision was not significantly influenced by age group (Wald χ^2 = 4.97, df = 3, p = 0.172). These patterns of responding with age group as a categorical predictor were in keeping with analyses of child data only when age in years was included as a continuous predictor (see Table S6). TOJs and CJs were significantly associated for the 3-object pseudocollision—participants who reordered events B and C were more likely to report perceiving a collision between those objects (Phi = 0.19, p = 0.029, see Table S7 for details per age group).

Discussion

Our Experiment 2 adult data closely replicates the results of Experiment 1—we again found evidence for the reordering of events in line with causality, according to both the TOJ data and the CJ data. Interestingly, reducing the number of practice clips appeared to have little impact on adults' susceptibility to reordering (we had speculated that including fewer practice clips might lead to more adults reordering), though we did make additional task modifications that could have reduced susceptibility (e.g., asking the same TOJ question throughout; only ever asking about the squares). However, by contrast to the findings of Experiment 1, children's TOJs in Experiment 2 suggest that it is only from around 8 years of age that reordering of events in line with causal impressions emerges (as 8- to 10-year-olds was the youngest age group in which we found a significant difference in TOJ performance between the 2-object and 3-object clips, see Table 1), and that susceptibility to this effect increases with age. Somewhat surprisingly, the two youngest groups of children (4- to 6- and 6- to 8-year-olds) were equally likely to correctly report the identity of the square that moved first (C) in the 2-object and 3-object clips and were highly accurate in both cases, providing no evidence that the inclusion of object A led them to reorder events in this version of the task. Furthermore, 4- to 6-year-olds were significantly more likely to report the correct order of events in the pseudocollision than adults.

The child CJ data, on the other hand, largely mirror what we found in Experiment 1—all age groups were significantly more likely to incorrectly report perceiving a collision in the 3-object pseudocollision than the 2-object control clip, and responses did not differ significantly across age groups. Thus, we see an intriguing difference in the pattern of performance across our two measures for the youngest children—their CJs suggest that they viewed B as bumping into C in the 3-object clip, but they do not report reordering in their TOJs. Specifically, while almost all children in the youngest group provided the correct response to the TOJ question for both clips (providing no evidence for reordering), around

60% of them incorrectly reported perceiving a collision between B and C in the 3-object clip, which suggests that the inclusion of object A *did* generate an impression of causality for them.

The results of Experiment 2 raise two distinct questions: (1) what might explain the difference in children's TOJ responses between Experiments 1 and 2, and (2) how can we reconcile the difference between young children's TOJ data and CJ data in Experiment 2? We will start by addressing the first question. One possibility is that young children really do experience the correct order of events in the 3-object clip (i.e., the increasing susceptibility to reordering with age result of Experiment 2 is valid) but something about the procedure in Experiment 1 led them to give answers that misleadingly suggested they reordered the events. Alternatively, perhaps children really do reorder events in line with causality (i.e., the Experiment 1 TOJ result is valid), but something about the procedure in Experiment 2 leads them to give an answer that misleadingly suggests they did not reorder the events. Finally, it seems feasible that the results of both experiments are valid, but the modifications we made to the procedure in Experiment 2 led young children to ignore object A (circle) and focus solely on the two squares; thus they performed comparably in the 2-object and 3-object clips.

To elaborate on this potential 'ignore object A' explanation for the Experiment 2 TOJ data: in Experiment 1 the practice trials encouraged participants to attend to the entire display because all shapes were squares, and the TOJ question differed between clips—sometimes participants were asked about which square moved first, and sometimes about which moved last. Thus, when they saw the critical clip they were likely attending to the entire display, including object A, which is presumably critical for the reordering effect to occur given that without attending to object A, the 3-object clip is identical to the 2-object control clip. During the practice trials of Experiment 2, on the other hand, participants were primed to attend only to the 2 squares (B and C), as they were only ever asked about these shapes, and furthermore they were only ever asked which one moved first. Thus, when they saw the 3-object pseudocollision they may have completely ignored the circle and focussed their attention only on the two squares (B and C), and specifically on which one moved first (anecdotally, some children reported that they were using this strategy).

If this explanation is correct, then why were younger children's TOJs more affected by the changes to the task (and adults apparently unaffected)? One possibility is that the causal impression generated by the clip is more irresistible to older children and adults because of their more extensive experience of a variety of causal systems and, hence, stronger priors—perhaps we become less able to 'escape' the impression of causality as we get older (Bechlivanidis, 2015).

Turning to the second question of how to reconcile the difference between young children's TOJ data and CJ data in Experiment 2, we see two possibilities. First, perhaps young children's CJ data, which in both experiments suggests they had a causal impression, could be explained by children glossing the test question as a question about whether there was a collision in the clip rather than interpreting it as a question about B and C. Specifically, perhaps these young children incorrectly say "yes" because they do perceive *a* collision (between objects A and B), but they do not actually perceive contact between objects B and C. (We note that one difficulty with this interpretation is that it seems inconsistent with the 'ignore A' explanation of the young children's TOJ data, because it suggests that children paid sufficient attention to A to perceive it making contact with B). The second possibility is that both TOJ and CJ data are valid in Experiment 2, i.e., there is a genuine difference between how collision perception and temporal order perception are affected by the causality manipulation in the youngest group. That is, perhaps in this youngest group, participants have the impression that B collided with C, but their temporal order judgements are not affected by the causality manipulation in the way that older participants' judgements are.

In Experiment 3 we attempted to reduce the likelihood of participants engaging in an 'ignore A' strategy by presenting a series of practice clips that encouraged them to attend to all three shapes. If only attending to objects B and C was driving the pattern of TOJ responses in Experiment 2, then young children should revert to reordering (replicating the results of Experiment 1). If on the other hand younger children really are less susceptible to causal reordering then we should replicate the results of Experiment 2.

Experiment 3

The critical clips and questions that followed were the same as in Experiment 2 (Figure 2a[ii] and 2b). However, to encourage participants to attend to all of the shapes (which may not have been the case in Experiment 2 and could explain the lack of reordering in young children compared to in Experiment 1) we made some changes to the practice clips. Specifically, we aimed to create a situation in which, by the time the critical clips were viewed, participants did not know which shape they would be asked about. We did this by varying which object we asked about between practice trials: on some trials we asked which *shape* moved first, and in others we asked which *circle* moved first. Then, on the critical trials we asked which *square* moved first (Figure 1c).

Method

Participants. Our final sample consisted of 54 adults (40 female, 3-object: N = 28, M_{age} = 19 years; 2-object: N = 26, M_{age} = 19 years) and 197 children (119 female), none of whom had participated in Experiments 1—2. An additional two children were tested but excluded because they were inattentive (N=1), or because they repeatedly responded "don't know" to the questions (N=1). The child sample was divided into 3 age groups per condition: 4- to 6-year-olds (3-object: N = 34, M_{age} = 5 years 1 month; 2-object: N = 32, M_{age} = 5 years 5 months), 6- to 8-year-olds (3-object: N = 34, M_{age} = 7 years 1 month; 2-object: N = 31, M_{age}

= 7 years 0 months) and 8- to 10-year-olds (3-object: N = 34, M_{age} = 9 years 7 months; 2-object: N = 31, M_{age} = 9 years 1 month).

Materials. The materials were the same as in Experiments 1 and 2 but we again changed the colours of the shapes to red, blue and yellow (because a few of the youngest children were unsure of the colour grey in Experiment 2).

Procedure. Participants saw three non-causal practice clips (Figure 1 c): two clips with one square and one circle, and one clip with two circles and a square. After the 2-object practice clips participants were asked “which *shape* moved first?” and the correct answer was the circle for one clip, and the square for the other clip. After the 3-object practice clip participants were asked “which *circle* moved first?” The critical clips (2-object control clip and 3-object pseudocollision) were the same as in Experiment 2 (Figure 2a[ii] and 2b).

Results

Practice clips. Performance in the 2-object practice clips ranged from 76% of participants responding correctly (4- to 6-year-olds) to 95% of participants responding correctly (adults). Performance in the 3-object practice clip ranged from 55% of participants responding correctly (4- to 6-year-olds) to 94% of participants responding correctly (adults, see Table S2 for full details).

Pre-registered confirmatory analyses. Across all age groups, the majority of participants responded correctly to the TOJ question (that C moved first) in the 2-object control clip (Figure 5a). As in Experiment 2, there was a pattern of decreasing response accuracy in the TOJ question for the 3-object pseudocollision (blue bars of Figure 5a): younger children were again more likely to respond correctly than older children and adults when asked “Which square moved first?” Comparisons of TOJ responses between the 2-object and 3-object clips revealed that while 6- to 8-year-olds, 8- to 10-year-olds and adults were significantly more likely to respond correctly in the 2-object clip (Chi square tests, *ps* ≤ 0.002, Table 1), the 4- to 6-year-olds' performance did not differ significantly between the two critical clips (Fisher's Exact Test, $p = 0.108$, Table 1). As in Experiments 1 and 2, participants in all age groups were significantly more likely to say square B collided with square C in the 3-object pseudocollision than the 2-object control clip (Figure 5b, Chi-square tests: *ps* ≤ 0.017 for all, Table 1).

Figure 5 about here

Exploratory analyses. Logistic regression revealed that participants' tendency to report the correct order of events (TOJ question) in the pseudocollision was significantly influenced by age group (Wald $\chi^2 = 11.32$, df = 3, $p = 0.010$). Posthoc contrasts with Tukey adjusted p-values for multiple comparisons revealed a significant difference between 4- to 6-year-olds and 8- to 10-year-olds (log odds ratio = 1.69, $p = 0.015$), with the youngest children being more likely to respond correctly/less likely to reorder than the oldest children. There were no other significant differences between groups after adjusting for multiple comparisons (*ps* ≥ 0.124 for all other pairs of age groups, Table S5). Participants' tendency to report perceiving a collision between objects B and C (CJ question) in the 3-object pseudocollision was not significantly influenced by age group (Wald $\chi^2 = 1.20$, df = 3, $p = 0.754$). These patterns of responding with age group as a categorical predictor were in keeping with analyses of child data only when age in years was included as a continuous predictor (see Table S6). TOJs and CJs were significantly associated for the 3-object pseudocollision—participants who reordered events B and C were more likely to report perceiving a collision between those objects (Phi = 0.23, $p = 0.010$, see Table S7 for details per age group).

Discussion

In Experiment 3, we once again replicated our adult results. Thus, while including practice clips (and potentially simplifying the response measures) reduces susceptibility to causal reordering compared with in a 'one-shot' experiment where participants only see the critical clip, it seems that the number and nature of the practice clips does not influence adults' performance. Even using our simplified paradigm, around 40% of adults reorder the events, and 40-60% incorrectly report perceiving contact between objects B and C.

The child data from Experiment 3 is largely comparable to that obtained in Experiment 2—TOJ accuracy for the 3-object pseudocollision decreases with age (8- to -10-year-olds were significantly less accurate than 4- to 6-year-olds), and once again there is a discrepancy between the youngest children's TOJ responses and their CJ responses. Thus, we did not find any evidence that encouraging young children to attend to all of the objects in the display made them more likely to reorder events in line with causality. It is therefore tempting to conclude that young children really are less susceptible to causal reordering than older children and adults. This conclusion, though, still leaves us to explain why the youngest children's CJ responses resembled those of adults—there was no significant difference between age groups for the pseudocollision CJ responses. As we pointed out above, there are two possible reasons for this: i) either it is the case that these children's CJ data is explained by a tendency to interpret the test question as being about whether there was a collision (as opposed to where the collision occurred) or, ii) more radically, children's perception of collision are affected by the causality manipulation but their temporal order judgements are not.

However, a further possible explanation for the observed data remains, which was raised by some anecdotal observations while running Experiment 3 with the younger children. First, a handful of children spontaneously gave a response to the TOJ question for the 3-object pseudocollision (responding that square C moved first) before the experimenter had asked the question. This was despite the fact that, based on the practice trials, the experimenter might feasibly have asked “which *shape* moved first?”, or “which *circle* moved first?” to which the correct answer would have been object A/the circle in both cases. This suggests that these participants may have been responding to something other than the question being asked. Second, one 4-year-old correctly gave the response ‘C’, and then spontaneously said “because it’s in the lead!” This raises the possibility that some children, rather than reporting the motion onset, may be reporting the final spatial position of the objects, taking into account the direction of movement, and this misinterpretation may be more common for younger children. That is, when asked “Which square moved first?” they respond to the question “Which *came* first”, or which went furthest to the right (if motion direction is left-to-right), which is object C. In addition, spontaneous verbalizations by some children also suggested that the TOJ question was being misinterpreted—for example, some children responded that C moved first, but then went on to describe events along the lines of “A moved and hit B, and then that moved and hit C”, which was incompatible with the TOJ response they gave. Finally, it seems unlikely that 4- to 6-year-olds would only respond correctly 52% of the time in the 3-object practice trial, but 83% of the time in the 3-object pseudocollision given that the two clips were similar in terms of their complexity (they both involved three objects, and the relative motion onsets of the objects were identical in the two clip types).

If some children are inappropriately responding in this way (i.e., giving their answer on the basis of spatial position on the screen rather reporting temporal order), this could also explain the high levels of A-responding in Experiment 1. Recall that around 40% of the youngest age group gave the response “A” when asked “Which square started moving last?” This seemed baffling as square A was quite clearly the first object to move, but makes sense if some children are responding on the basis of the objects’ final positions (considering direction of movement), as outlined above. Under this account, object A “came last”—it finished spatially “behind” squares B and C. If we assume a similar proportion of the youngest children also responded along these lines in Experiments 2 and 3, that would explain a large chunk of the C-responses (because C “won/came first”), which in these two experiments happened to correspond to the correct answer about which object moved first. A reduction in the proportion of children responding on this “winner/loser” basis across age groups could explain the apparent developmental pattern of younger children appearing to give more accurate TOJs in the 3-object pseudocollision than we observed in Experiments 2 and 3. This account could also explain the differential way in which the causality manipulation affected TOJs and CJs—if the aforementioned hypothesis is correct (i.e., some proportion of young children are responding on the basis of which object came first/last), then it seems likely that the CJ data are valid, and younger children’s TOJ data are being influenced by the nature of the TOJ question being asked and do not reflect their actual perception of temporal order.

Experiment 4

In Experiment 4 we replicated Experiment 3, but replaced the 2-object control clip with a 3-object canonical collision where A was a circle and B and C were squares (just like the pseudocollisions in Experiments 2 and 3), so the veridical order of motion was ABC. As in Experiments 2 and 3, we asked participants “which square moved first?” If younger children are making a genuine TOJ, and are as accurate as they appear to be in Experiments 2 and 3, then in the canonical clip they should respond “B”. If they still respond “C” then this will provide support for the “winner/loser” spatially-based response outlined above.

To address whether the CJ results in the previous experiments might be explained by a tendency to respond “yes” when asked about the 3-object pseudocollision because of the presence of a collision between objects A and B, instead of only asking whether square B bumped into square C, for the critical clips we asked about all pairs of squares in a random order (i.e., Did A bump into B? Did B bump into C? Did A bump into C?). If participants are responding to this question in the way it is intended, for both critical clips participants should respond "yes" for A-B and "no" for A-C. They should also respond "yes" when asked about B-C in the canonical collision; if they also respond "yes" in the pseudocollision then this will provide evidence that participants do indeed perceive the movement of C as caused by B.

Method

Participants. Our final sample consisted of 127 children (65 female); 65 4- to 6-year-olds, none of whom had participated in Experiments 1—3 (pseudocollision: N = 35, M_{age} = 5 years 10 months; canonical collision: N = 30, M_{age} = 6 years 1 month) and 62 8- to 10-year-olds (pseudocollision: N = 32, M_{age} = 8 years 10 months; canonical collision: N = 30, M_{age} = 8 years 9 months). An additional 4 children were tested but excluded because they were inattentive (N=2), because they could not name the shapes (N=1), or because of experimenter error (N=1).

Procedure. The practice clips were the same as for Experiment 3 (Figure 1c). The critical clips consisted of the 3-object pseudocollision (ACB, Figure 2a[ii]) from Experiments 2 and 3, and a 3-object canonical collision (ABC, Figure 2c). In the canonical collision, object A moved towards object B and stopped adjacent to it, following which B started moving towards object C. B stopped adjacent to C, and C started moving away from B. As for the pseudocollision, all objects moved at a speed of 30 mm/s.

Results.

Practice clips. Performance in the 2-object practice clips was 72% correct responses for 4- to 6-year-olds and 92% correct responses for 8- to 10-year-olds. Performance in the 3-

object practice clip was 58% correct responses for 4- to 6-year-olds and 84% correct responses for 8- to 10-year-olds (see Table S1 for full details).

Pre-registered confirmatory analyses. Four- to six-year-olds' TOJs were significantly less accurate for the canonical collision where the correct response was 'B' (23% correct), than for the reordered pseudocollision where the correct response was 'C' (80% correct, $\chi^2 = 20.87$, $p < 0.001$); in fact, they were equally likely to say that C moved first for the pseudocollision and the canonical clip (Figure 6). The 8- to 10-year-olds on the other hand mostly gave the (correct) response that B moved first in the canonical clip, though 30% of participants in this age group still erroneously claimed that C moved first in the canonical clip (Figure 6). The older children were more likely to respond correctly in the canonical clip than in the pseudocollision, but not significantly so (canonical collision: 70% correct, pseudocollision: 59% correct, $\chi^2 = 0.76$, $p = 0.382$).

Figure 6 about here

Participants in both age groups were significantly more likely so respond 'yes' when asked whether A bumped into B (which it did) compared with when asked whether A bumped into C (which it did not), and this was true for both clip types (canonical and reordered, *ps* < 0.001 for all, Figure 7).

Figure 7 about here

In both age groups and for both types of clip the majority of participants (>80%) responded 'yes' when asked whether B bumped into C (Figure 7). There was no significant difference between the responses children in either age group gave for the canonical collision and the reordered collision when asked whether square B bumped into square C (4- to 6-year-olds: $\chi^2 = 0.03$, $p = 0.959$; 8- to 10-year-olds: $\chi^2 = 0.336$, $p = 0.562$).

Exploratory analyses. TOJs and CJs were significantly associated for the 3-object pseudocollision—participants who reordered events B and C were more likely to report perceiving a collision between those objects (Phi = 0.31, p = 0.013, see Table S2 for details per age group).

Discussion

Experiment 4 again replicated the developmental pattern of TOJ responses from Experiments 2 and 3, with younger children appearing to give more accurate TOJs (saying C moved first) than older children for the reordered pseudocollision clip. However, the results for the canonical collision strongly suggest that this does not reflect a better ability to perceive the veridical order of events in early childhood. When shown a canonical collision, older children gave more accurate TOJs than younger children. Specifically, the majority of children in the younger age group responded incorrectly to the TOJ question when presented with a canonical collision where the correct answer was ‘B’, which strongly suggests that they tend to give the response ‘C’ regardless of clip type. Eight- to 10-year-olds on the other hand mostly gave the correct response ‘B’ for the canonical collision, though almost 1/3 still responded ‘C’, suggesting that the TOJ question may also cause problems for some older children. Thus it appears that the majority of young children and some older children may not be interpreting the TOJ question (“which square moved first?”) as it was intended; instead they appear to respond on the basis of which square ‘came first’, choosing a square on the basis of spatial position. Furthermore, as in the previous experiments we did not find the expected association between TOJs and CJs for the youngest group of children.

In addition to asking whether square B bumped into square C as in Experiments 1–3, in Experiment 4 we also asked participants for their collision judgements about the other pairs of shapes. This enabled us to establish that children of all of the ages tested do indeed understand the collision question and interpret it correctly (i.e., they are able to correctly identify the presence/absence of a ‘bump’ between object pairs) – they typically say ‘yes’ when asked whether A bumped into B, and ‘no’ when asked whether A bumped into C. Interestingly, > 80 % of participants in both age groups reported (incorrectly) that B did bump into C in the pseudocollision. Given that a comparable percentage of participants gave this response for the canonical collision, this provides strong evidence that the causal impression generated by the pseudocollision is similar to that generated by the canonical collision.

General Discussion

Across four experiments we investigated whether children, like adults, reorder events in line with causality. We modified an existing adult paradigm (Bechlivanidis & Lagnado, 2016) for this purpose: in each experiment participants watched a 3-object pseudocollision in which the order of events was manipulated so that, unlike in a canonical collision, the third object in line (C) moved *before* the middle object (B) (i.e., the order of motion onset was ACB, and object B never collided with object C). They were then asked (a) a temporal order judgement (TOJ) question and (b) a collision judgement (CJ) question (three in Experiment 4). If participants reorder events in line with causality, then they should incorrectly report that B moved before C. If the introduction of A affects whether they perceive a collision between B and C, they should also incorrectly report that B bumped into C.

Overall, we found evidence that the causality manipulation affected children’s perception of the order of events in the sequence. Across all four experiments participants in all age groups (including adults) were significantly more likely to report perceiving a collision between objects B and C in the 3-object pseudocollision than in the 2-object control clip, despite the spatiotemporal relations between B and C being identical in the two clips.

Furthermore, CJs did not differ significantly between age groups (apart from in Experiment 1, where 9- to 10-year-olds were more likely to report a collision than 7- to 9-year-olds). We also found evidence for reordering according to our TOJ measure in the majority of age groups: from 4 years in Experiment 1, from 8 years in Experiment 2, and from 6 years in Experiment 3. However, our two measures were not consistently associated with one another (see supplementary Table S7) and the TOJ data from the younger children showed an interesting pattern of results that warrants further discussion.

Although TOJ responses in Experiment 1 provided evidence for reordering in all age groups, taken at face value the subsequent TOJ results from Experiments 2 and 3 suggested that younger children did not reorder events, and may in fact have been more accurate than older children and adults in their perception of the order of events. However, Experiment 4 demonstrated that some children—particularly in the younger age range—had a systematic tendency to respond based on spatial rather than temporal information when asked "Which square moved first?" Specifically, when shown a canonical collision where the order of motion onset was ABC, the majority of young children still reported that C moved first (i.e., before B). Thus, it appears that some children respond on the basis of which square 'came first', rather than which started to move first. This basis for responding can also explain the large proportion of young children saying that object A started moving last in Experiment 1—in this case, A 'came last'.

Despite deliberately avoiding use of the terms 'before' or 'after' in our TOJ questions, our results demonstrate that, at least under these circumstances, asking which object moved first/last is also not an appropriate measure of very young children's temporal order perception in this context (i.e., when there is a possible spatial interpretation of the question). The general idea that young children are likely to (erroneously) focus on spatial rather than temporal cues has a long history within developmental psychology (Piaget, 1969; see

McCormack, 2015, for historical review). The current findings add to the body of evidence that suggests that young children may privilege spatial information, perhaps because of the more concrete nature of spatial cues (Casasanto & Boroditsky, 2007; Casasanto, Fotakopoulou, & Boroditsky, 2010).

However, Experiment 4 also confirmed that young children's collision judgements were valid: following the canonical clip, they were able to accurately identify the presence (between A and B) and absence (between A and C) of a 'bump' between objects. Taken together with the CJ results for Experiments 1-3, this suggests that the inclusion of object A generates a causal impression that modulates children's experience of the subsequent motion of B and C. In Experiment 4, children in both age groups were equally likely to report perceiving a collision between B and C in the pseudocollision (where there was no collision between these objects) and in a 3-object canonical collision (where there actually was a collision between B and C). This suggests that for 4- to 10-year-olds, as for adults, the pseudocollision generates the same impression of causality as a genuine collision.

What then should we conclude about the developmental profile of the reordering effect? Setting aside the data from the youngest age group (4- to 6-year-olds), there was no evidence across Experiments 1—3 that susceptibility to the causal reordering effect increases with age. This suggests that causal reordering is present in children, as it is in adults, and that it remains stable over development. The key issue is whether we should conclude that this effect is also present in early childhood, in 4- to 6-year-olds. As we have pointed out, across four experiments the CJ data from this age group consistently suggested that they are as likely as older children and adults to mistakenly report that B collided with C in the 3-object clip. The data from Experiment 4 indicate that there is no reason to assume that the causality manipulation genuinely had a differential effect on young children's collision perception and their temporal order perception; rather, their temporal order judgements were unreliable. The

4- to 6-year-olds' performance in the 3-object practice clips—where it was not possible to respond on the basis of a spatial strategy—were poor compared with other age groups, suggesting that children in this age group may have difficulties tracking and remembering the order of motion onset of three objects. Thus, the most conservative conclusion is that we do not yet know whether 4- to 6-year-olds show the causal reordering effect. However, taken alongside children's CJ data, we believe that the findings of Experiment 1 provide a good reason for believing that causal reordering is indeed evident in this age group. Unlike in Experiments 2—4, we can exclude children in Experiment 1 who responded to the TOJ question on the basis of spatial position: these are the children who reported that A started moving last. Indeed, our existing analysis excluded these children (based on our pre-registered confirmatory analysis plan), and a substantial majority of the remaining children in this group (76%) reported that C was the last object to move in the 3-object pseudocollision clip (but not in the 2-object clip). Thus, the findings of Experiment 1 suggest that causal reordering is present even in 4- to 6-year-olds.

In sum, we believe that our findings provide evidence for an early-developing role of causality in interpreting the environment. While infants' causal perception has previously been shown to be influenced by bottom-up visual factors in a comparable way to adults' (e.g., the grouping effect, Choi & Scholl, 2004; Newman et al., 2008), the present study demonstrates that children's causal perception can also exert top-down effects on their temporal perception, as is the case for adults (Bechlivanidis & Lagnado, 2016). This evidence that causality can influence children's experience of time is in keeping with recent research showing that children as young as four years are susceptible to temporal binding—with children predicting that events will occur earlier if they are causally connected to a preceding event, compared to when it is preceded by an arbitrary predictive signal (Blakey et al., 2018). Thus, it appears that not only do children use temporal cues to make causal judgements (e.g., Bullock & Gelman, 1979; McCormack et al., 2015; Mendelson & Shultz, 1976; Rankin & McCormack, 2013; Schlottmann et al., 1999); they also use causal cues to make temporal judgements—about the duration between events, and about the order in which events occurred.

Although the results presented in the current study are illuminating with respect to the developmental trajectory of causal reordering, important questions remain regarding the mechanism underpinning the effect. Properly answering these questions is beyond the scope of the present study, and will require developing new paradigms to distinguish between possible explanations of the reordering effect. Nevertheless, in what follows we outline these different potential explanations, discuss what has been established to date, and describe our ongoing work with adults that aims to generate new evidence to definitively distinguish between these alternative explanations.

There are three distinct types of explanation that might account for the reordering effect, which are set out by Bechlivanidis and Lagnado (2016). First, it is possible that when viewing the 3-object pseudocollision participants fail to see all of the events and so they do not actually perceive their order (inattention). Specifically, it is plausible that the motion of object B could be missed, as attention is diverted by the motion onset of object C. On such an explanation, reordering occurs because participants 'fill in' the missing information by making a *post hoc* inference on the basis of the most likely order of events, given their causal impression. Arguably this is the least interesting explanation of the effect, because it suggests that participants simply speculate about what might have happened, rather than their judgments being based on processing the events that they were presented with. Second, the reordering effect could occur if participants do attend to and accurately perceive the order of all events, but because of the causal impression generated by the clip, the memory of events they ultimately retrieve is of the more plausible causal order (misremembering). Finally, it may be the case that participants' original representation of the temporal order of events matches the causal order rather than the objective order—i.e., they actually perceive events happening in an order that does not reflect reality (misperceiving). This last possibility is particularly interesting, because it challenges what might be seen as the intuitive view of perception, namely that events are perceived in the order in which they occur, so that the temporal structure of experience simply mirrors the temporal structure of events in the world (Hoerl, 2013; Phillips, 2014).

Previous findings with adults speak against the inattention account of reordering (that participants do not attend to all of the objects in the pseudocollision). When participants first watch a pseudocollision, and are subsequently presented with a pseudocollison and a canonical collision side by side, they tend to mistake the pseudocollision they initially saw for the canonical collision. In contrast, when they are first presented with a slightly modified pseudocollision clip in which B does not move at all, this is detected by most people and they are able to identify it as the clip they saw, rather than mistaking it for a canonical collision (Experiment 2, Bechlivanidis & Lagnado, 2016). This suggests that participants apparently do attend to the behaviour of object B—they are not simply filling in missing information *post hoc* because they did not see what happened. However, this study could not distinguish between 'misremembering' and 'misperceiving' accounts of the reordering effect. Distinguishing between these two accounts is difficult because in the studies to date participants have made their judgments after the events have happened. Ideally, in order to examine what participants perceive (rather than what they construct in memory), a paradigm would be used that taps into the processes that occur while the events themselves unfold. However, given the very short time scales over which the events happen, such a paradigm could not involve participants making explicit verbal judgments, as such judgments are by necessity *post-hoc*. We are currently testing a paradigm with adults that we believe taps into the processes that occur as the events unfold, in which participants have to synchronize the occurrence of another unrelated event with the onset of movement of B or C. In this task, participants are given multiple opportunities to view the pseudocollision and adjust the timing of the unrelated event so that they perceive it as occurring simultaneously either with the movement of B or the movement of C. If causal reordering stems from a genuine perceptual effect (participants perceive B moving before C), then the temporal location of events should be shifted to match causal assumptions—when synching with B, participants should place the unrelated event earlier than the actual onset of B's motion, and when synching with C they should place the unrelated event later than the actual onset of motion. If instead participants accurately perceive the order of events (they perceive C moving before B) and it is only later that their causal impression interferes with their temporal order judgement, then their placements of the unrelated event should reflect the veridical timing of B's and C's motion onset.

Depending on our adult findings, we hope to subsequently explore whether this task can also be adapted for use with children, although the task is likely to be more challenging than the one used in the current study because of the need for multiple trials in which millisecond timing adjustments are made (though see Blakey et al., 2018). We should emphasize, though, that in our view the developmental profile of the reordering effect is interesting regardless of whether a misremembering or misperceiving explanation of it is correct. This is because, regardless of which of these explanations is correct, reordering serves as a novel demonstration of how causal assumptions have top-down effects on basic processes. Establishing whether such assumptions play a similar role in children sheds light on the extent to which causal cognition plays a similar fundamental role from early in development.

Thus, the current findings are informative with regards to children's causal reasoning abilities more broadly. First, our results add to the small body of work suggesting that children's perception of physical causation is largely similar to that of adults (Schlottmann, Allan, et al., 2002; Schlottmann, Cole, et al., 2013). Previous research has used simple two-object displays and indicated that the introduction of delays or spatial gaps reduces the likelihood that children perceive physical causation (Schlottmann et al., 2013); in this respect children largely resemble adults. However, the pseudocollision presented to children in the present study apparently generated a causal impression (as participants reported that B bumped into C), even though no contact was made and C moved before B. As with adult findings (Bechlivanidis & Lagnado, 2016), these results suggest that, rather than causal impressions being determined only by the basic spatial-temporal properties of object movement, schemata—in this case, a series of collisions—are used in a top-down manner in the interpretation of perceptual displays. Such schemata appear to be used in the same way in young children as in adults. Second, a large body of previous work has demonstrated that young children are able to use the causal structure of events in the world to make inferences and guide their behaviour (e.g., Muentener & Schulz, 2016; Sobel & Legare, 2014). Causal reasoning has been proposed to play an important role in diverse domains, including children's understanding of the physical world (e.g., Baillargeon, 2004), the development of morality (e.g., Hamlin, 2013), and the generation of explanations (e.g., Legare, 2012). The present study extends the evidence on the influence of causality on children's experience of the world to another domain: their experience of time. Thus, the current results add to a growing body of evidence that causality plays a fundamental role in our experience of the world from early in development.

On the assumption that the present study has demonstrated that children as young as four years reorder events to match a causal interpretation, further work is needed to establish the developmental origins of this temporal illusion. For example, a habituation paradigm could be used to test whether or not infants discriminate between a canonical 3-object collision and the reordered pseuodocollision. There would also be value in developing a paradigm appropriate for comparative studies to enable investigation of the evolutionary origins of causal reordering. While 'higher' causal knowledge and inference has been reasonably widely explored in non-human animals (e.g., Seed & Call, 2009), there have been relatively few studies of causal perception. Recent research has demonstrated that chimpanzees are susceptible to causal capture, in which a causal impression can induce perceptual alteration of the spatiotemporal properties of co-occurring events (Matsuno & Tomonaga, 2017; Scholl & Nakamaya, 2002). This provides initial evidence that causality also influences the visual perception of our closest ape relatives, but just how phylogenetically widespread susceptibility to causality-based temporal illusions might be remains an open question.

To conclude, the findings reported in the present study add to a small but growing body of evidence demonstrating an early-developing bidirectional relation between time and causality (Blakey et al., 2018; Lorimer et al., 2017). The current study extends this research by showing that children's causal impressions can qualitatively alter their temporal experience—through the reordering of events to match a causal interpretation.

References

Baillargeon, R. (2004). Infants' Physical World. *Current Directions in Psychological Science*, 13(3), 89–94. https://doi.org/10.1111/j.0963-7214.2004.00281.x

Bechlivanidis, C. (2015). *The arrow of time through the causal lens: When causal beliefs determine temporal order.* Doctoral dissertation, UCL (University College London).

Bechlivanidis, C. & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order, *Cognition, 146,* 58-66. https://doi.org/10.1016/j.cognition.2015.09.001

Bechlivanidis, C., & Lagnado, D. A. (2013). Does the “why” tell us the ‘‘when”? *Psychological Science, 24*(8), 1563-1572. https://doi.org/10.1177/0956797613476046

Bechlivanidis, C., Schlottmann, A., Lagnado, D.A. (2019) Causation Without Realism *Journal Of Experimental Psychology: General*, 148(5), 785-804. doi: 10.1037/xge0000602

Blakey, E., Tecwyn, E. C., McCormack, T., Lagnado, D. A., Hoerl, C., Lorimer, S. & Buehner, M. J. (2018) When Causality Shapes the Experience of Time: Evidence for Temporal Binding in Young Children. *Developmental Science,* e12769. https://doi.org/10.1111/desc.12769

Blything, L. P., & Cain, K. (2016). Children’s processing and comprehension of complex sentences containing temporal connectives: The influence of memory on the time course of accurate responses. *Developmental Psychology*, *52*(10), 1517. http://dx.doi.org/10.1037/dev0000201

Blything, L. P., Davies, R., & Cain, K. (2015). Young children's comprehension of temporal relations in complex sentences: The influence of memory on performance. *Child Development*, *86*(6), 1922-1934. DOI: 10.1111/cdev.12412

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, doing, and controlling: implicit sense of agency over somatosensory events. *Psychological Science*, *28*(7), 882-893. 10.1177/0956797617697693

Buehner, M. J., (2012). Understanding the past, predicting the future: Causation, not intentional action, is the root of temporal binding. *Psychological Science, 23*, 1490-1497. https://doi.org/10.1177/0956797612444612

Buehner, M. J. (2015). Awareness of voluntary and involuntary causal actions and their outcomes. *Psychology of Consciousness: Theory, Research, and Practice, 2*(3), 237-252. DOI:10.1037/cns0000068

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*, 1221-1228. https://doi.org/10.1111/j.1467-9280.2009.02435.x

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology Section A*, *56*(5), 865-890. DOI:10.1080/02724980244000675

Bullock, M., & Gelman, R. (1979). Preschool children’s assumptions about cause and effect: temporal ordering. *Child Development, 50*, 89–96. DOI: 10.2307/1129045.

Casasanto, D., & Boroditsky, L. (2007). Time in the mind: Using space to think about time. *Cognition, 106,* 579-593. https://doi.org/10.1016/j.cognition.2007.03.004

Casansanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and time in the child’s mind: Evidence for a cross-dimensional asymmetry. *Cognitive Science*, *34,* 387-405. https://doi.org/10.1111/j.1551-6709.2010.01094.x

Cavazzana, A., Begliomini, C., & Bisiacchi, P. S. (2014). Intentional binding effect in children: insights from a new paradigm. *Frontiers in Human Neuroscience*, *8*, 651. https://doi.org/10.3389/fnhum.2014.00651

Cavazzana, A., Begliomini, C., & Bisiacchi, P. S. (2017). Intentional binding as a marker of agency across the lifespan. *Consciousness and Cognition*, *52*, 104-114. https://doi.org/10.1016/j.concog.2017.04.016

Choi, H., & Scholl, B. J. (2004). Effects of grouping and attention on the perception of causality. *Perception & Psychophysics*, *66*(6), 926-942. DOI: 10.3758/BF03194985

Cohen, L. B., & Amsel, G. (1998). Precursors to infants' perception of the causality of a simple event. *Infant Behavior and Development*, *21*(4), 713-731. DOI: 10.1016/S0163-6383(98)90040-6

Desantis, A., Waszak, F., Moutsopoulou, K., & Haggard, P. (2016). How action structures time: about the perceived temporal order of action and predicted outcomes. *Cognition*, *146*, 100-109. DOI: 10.1016/j.cognition.2015.08.011

Faro, D., McGill, A. L., & Hastie, R. (2013). The influence of perceived causation on judgements of time: an integrative review and implications for decision-making. *Frontiers in Psychology*, *4*, 217. DOI: 10.3389/fpsyg.2013.00217

Hamlin, J. K. (2013). Moral Judgment and Action in Preverbal Infants and Toddlers: Evidence for an Innate Moral Core. *Current Directions in Psychological Science*, 22(3), 186–193. https://doi.org/10.1177/0963721412470687

Hoerl, C. (2013). A succession of feelings, in and of itself, is not a feeling of succession. *Mind*, 122 (486) (2013), pp. 373-417, 10.1093/mind/fzt070

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 451. DOI: 10.1037/0278-7393.32.3.451

Legare, C.H. (2012). Exploring explanation: explaining inconsistent evidence informs exploratory, hypothesis-testing behavior in young children. *Child Development*, 83 1, 173-85. DOI:10.1111/j.1467-8624.2011.01691.x

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition*, *25*(3), 265-288. DOI: 10.1016/S0010-0277(87)80006-9

Lorimer, S., McCormack, T., Blakey, E., Lagnado, D. A., Hoerl, C., Tecwyn, E.C., & Buehner, M. J. (under review). The Developmental Profile of Temporal Binding: From Childhood to Adulthood.

Mascalzoni, E., Regolin, L., Vallortigara, G., & Simion, F. (2013). The cradle of causal reasoning: newborns' preference for physical causality. *Developmental Science*, *16*(3), 327-335. https://doi.org/10.1111/desc.12018

Matsuno, T., & Tomonaga, M. (2017). Causal capture effects in chimpanzees (*Pan troglodytes*). *Cognition*, *158*, 153-164. DOI: 10.1016/j.cognition.2016.10.023

McCormack, T. (2015). The development of temporal cognition. *Handbook of Child Psychology and Developmental Science*, *2*, 624-670.

Mendelson, R., & Shultz, T. R. (1976). Covariation and temporal contiguity as principles of causal inference in young children. *Journal of Experimental Child Psychology*, *22*(3), 408-412. https://doi.org/10.1016/0022-0965(76)90104-1

Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences*, *8*, 22-27. DOI: 10.1016/j.cobeha.2016.01.006

Muentener, P., & Bonawitz, E. (2017). The Development of Causal Reasoning. In Waldman, M. R.. (Ed.) *The Oxford Handbook of Causal Reasoning*. Oxford University Press. DOI: 10.1093/oxfordhb/9780199399550.013.40

Newman, G. E., Choi, H., Wynn, K., & Scholl, B. J. (2008). The origins of causal perception: Evidence from postdictive processing in infancy. *Cognitive Psychology*, *57*(3), 262-291. DOI: 10.1016/j.cogpsych.2008.02.003

Oakes, L. M. (1994). Development of infants' use of continuity cues in their perception of causality. *Developmental Psychology*, *30*(6), 869. DOI: 10.1037/0012-1649.30.6.869

Phillips, I. (2014). Experience of and in Time. *Philosophy Compass*, 9(2), 131–144.

Piaget, J. (1969). *The child's conception of time.* London: Routledge and Keegan Paul.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rankin, M. L., & McCormack, T. (2013). The temporal priority principle: at what age does this develop? *Frontiers in Psychology, 4.* https://doi.org/10.3389/fpsyg.2013.00178.

Saxe, R., & Carey, S. (2006). The perception of causality in infancy. *Acta Psychologica*, *123*(1-2), 144-165. DOI: 10.1016/j.actpsy.2006.05.005

Schlottmann, A. (1999). Seeing in happen and knowing how it works: how children understand the relation between perceptual causality and underlying mechanism. *Developmental Psychology, 35,* 303-317. DOI: 10.1037/0012-1649.35.1.303

Schlottmann, A., Allen, D., Linderoth, C., & Hesketh, S. (2002). Perceptual causality in children. *Child Development, 73*, 1656-1677. https://doi.org/10.1111/1467-8624.00497

Schlottmann, A., Cole, K., Watts, R., & White, M. (2013). Domain-specific perceptual causality in children depends on the spatio-temporal configuration, not motion onset. *Frontiers in Psychology*, *4*, 365. DOI: 10.3389/fpsyg.2013.00365

Scholl, B. J., & Nakayama, K. (2002). Causal capture: Contextual effects on the perception of collision events. *Psychological Science*, 13(6), 493-498. DOI: 10.1111/1467-9280.00487

Seed, A., & Call, J. (2009). Causal knowledge for events and objects in animals. In *Rational Animals, Irrational Humans* (pp. 173-188). Keio University.

Sobel, D. M., & Legare, C. H. (2014). Causal learning in children. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(4), 413-427.

White, P. A. (2018). Perceptual impressions of causality are affected by common fate. *Psychological Research*, 82(4), 652-664. DOI: 10.1007/s00426-017-0853-y

Table 1. Summary of results comparing performance in the 2-object control clip and the 3-object pseudocollision for all age groups in Experiments 1—3 for the temporal order judgement (TOJ) and collision judgement (CJ) measures.

		Age Group				
	Measure	**4 to 6**	**6 to 7**	**7 to 9**	**9 to 10**	**Adult**
Exp. 1	**TOJ**	$\chi^2 = 29.89$ $p < 0.001$	$\chi^2 = 32.61$ $p < 0.001$	$\chi^2 = 28.13$ $p < 0.001$	$\chi^2 = 40.24$ $p < 0.001$	$\chi^2 = 15.99$ $p < 0.001$
	CJ	$\chi^2 = 10.56$ $p = 0.001$	$\chi^2 = 15.59$ $p < 0.001$	$\chi^2 = 17.21$ $p < 0.001$	$\chi^2 = 32.94$ $p < 0.001$	$\chi^2 = 18.28$ $p < 0.001$

		Age Group				
	Measure	**4 to 6**	**6 to 8**	**8 to 10**		**Adults**
Exp. 2	**TOJ**	$p = 0.238$ [a]	$p = 0.082$ [a]	$\chi^2 = 8.72$ $p = 0.003$		$\chi^2 = 16.31$ $p < 0.001$
	CJ	$\chi^2 = 13.89$ $p < 0.001$	$\chi^2 = 9.67$ $p = 0.002$	$\chi^2 = 7.33$ $p = 0.007$		$\chi^2 = 13.12$ $p < 0.001$
Exp. 3	**TOJ**	$p = 0.108$ [a]	$p = 0.002$[a]	$\chi^2 = 22.70$ $p < 0.001$		$\chi^2 = 12.83$ $p < 0.001$
	CJ	$\chi^2 = 5.73$ $p = 0.017$	$\chi^2 = 22.71$ $p < 0.001$	$\chi^2 = 20.75$ $p < 0.001$		$\chi^2 = 14.84$ $p < 0.001$

[a] Fisher's Exact Test

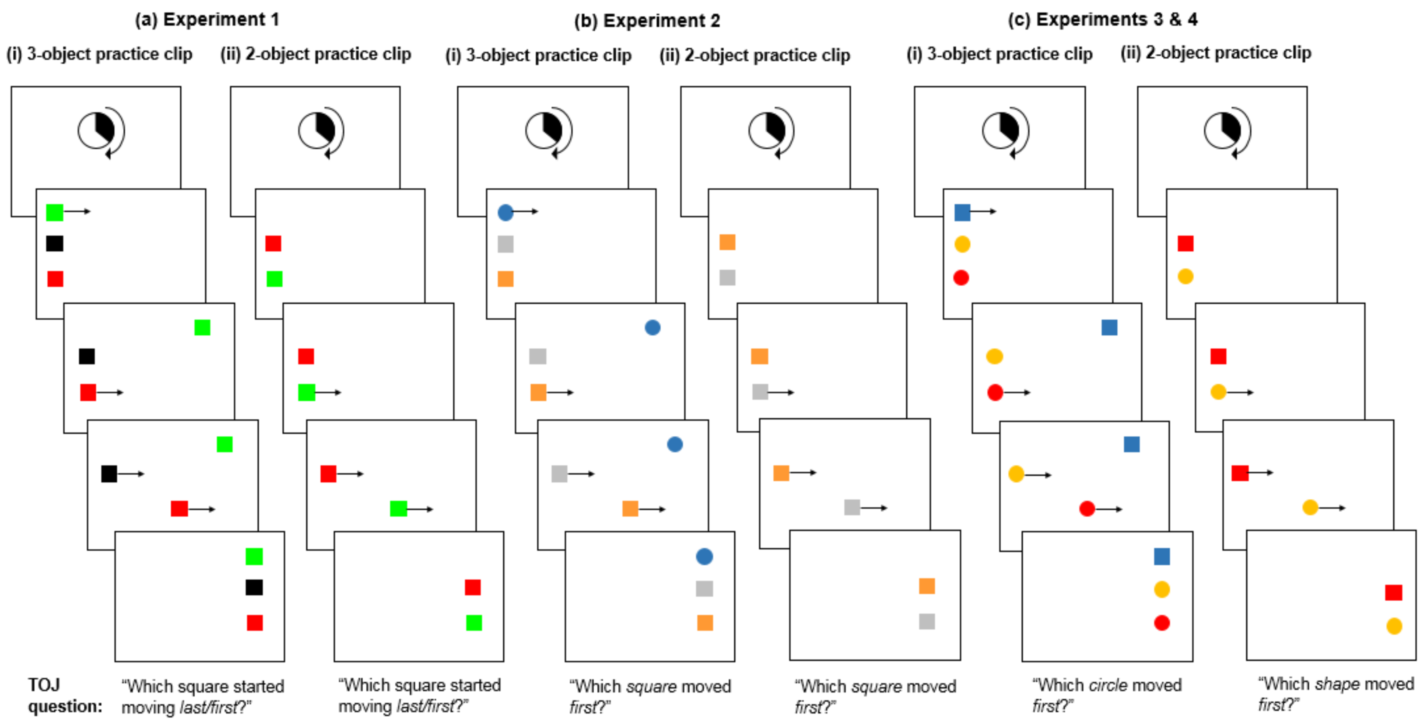

Figure 1. Schematic representations of example practice clips seen by participants in (a) Experiment 1, (b) Experiment 2 and (c) Experiments 3 and 4, and the TOJ question they were asked after each clip. Direction of motion shown is left-to-right, but could also be right-to-left. The colours of the objects were randomized between participants. Clips were presented in a random order. In Experiment 1 participants saw two clips of each type (3-object and 2-object; 4 in total) and motion onset order of the shapes was random. They were either asked about which square started moving last or first, with the order alternating between clips. In Experiment 2 participants saw one clip of each type and the circle always moved first in the 3-object clip. In Experiments 3 and 4 participants saw one 3-object clip where the square always moved first, and two 2-object clips: one where the circle moved first and one where the square moved first (not shown).

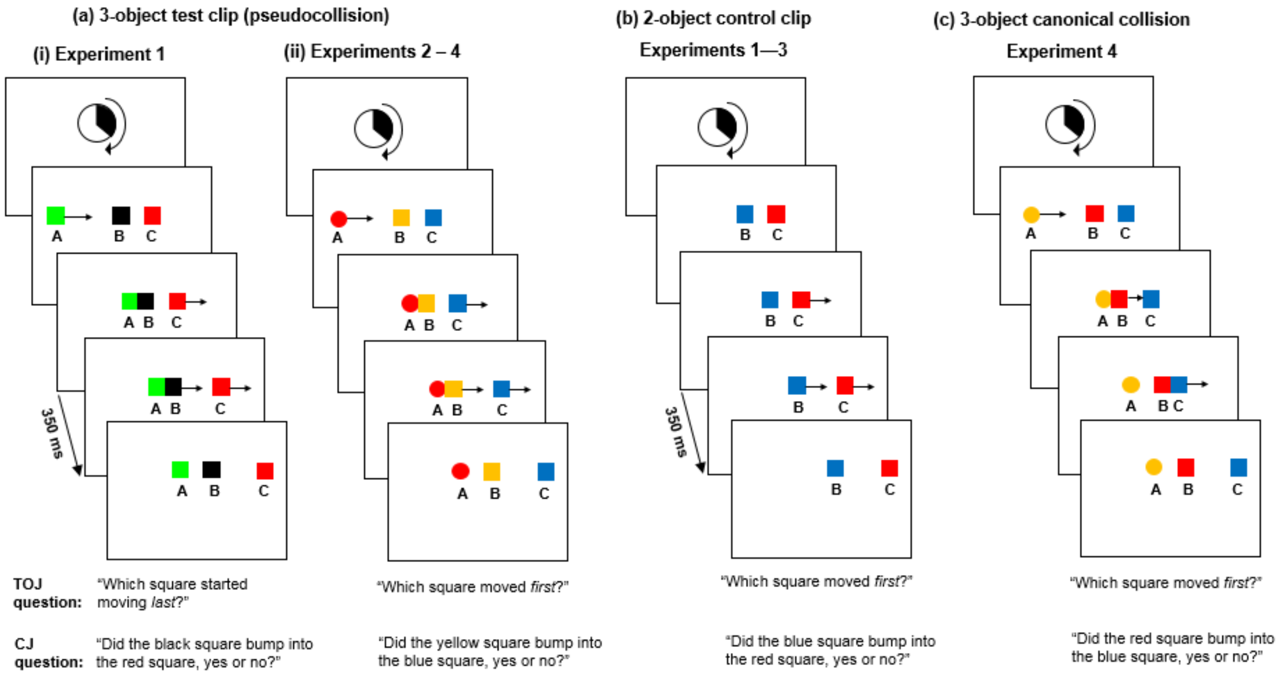

Figure 2. Schematic representations of (a) the 3-object pseudocollision clip used in [i] Experiment 1 and [ii] Experiments 2—4; (b) the 2-object control clip used in Experiments 1—3; and (c) the 3-object canonical collision used in Experiment 4, and the TOJ and CJ questions participants were asked after each clip. Direction of motion shown is left-to-right, but could also be right-to-left. The colours of the objects were randomised between participants. In Experiment 2 the colours used were orange, blue and grey (not shown). In Experiment 4, participants were asked a CJ question about each pair of shapes (in a random order) for the pseudocollision and the canonical collision, so for the example shown for the latter they would also have been asked whether the yellow circle bumped into the red square, and whether the yellow circle bumped into the blue square.

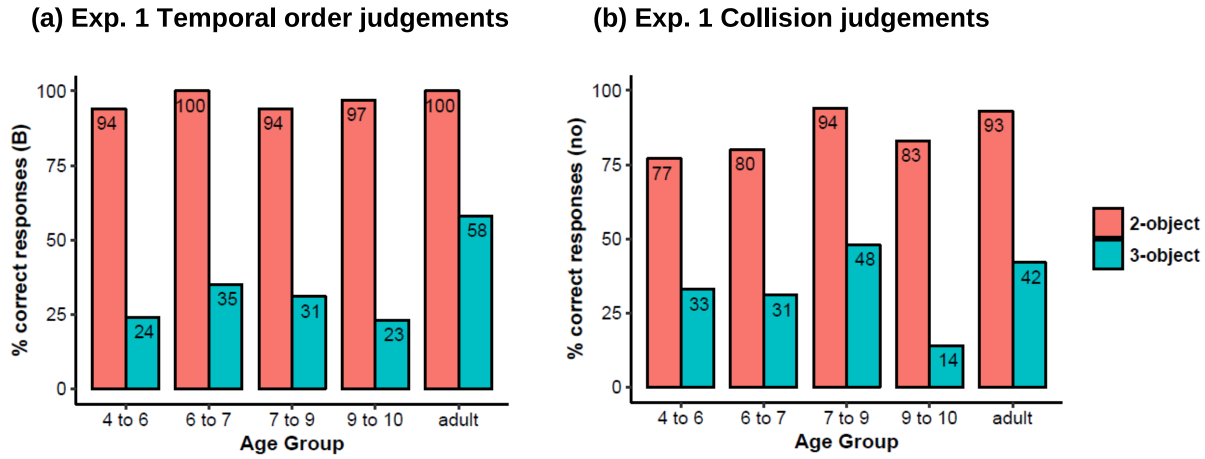

Figure 3. Percentage of participants in each age group who gave the correct response in (a) the temporal order judgement question (square B); and (b) the collision judgement question (no), in the 2-object control clip (red bars/left-hand bar for each age group) and 3-object pseudocollision (blue bars/right-hand bar for each age group) of Experiment 1.

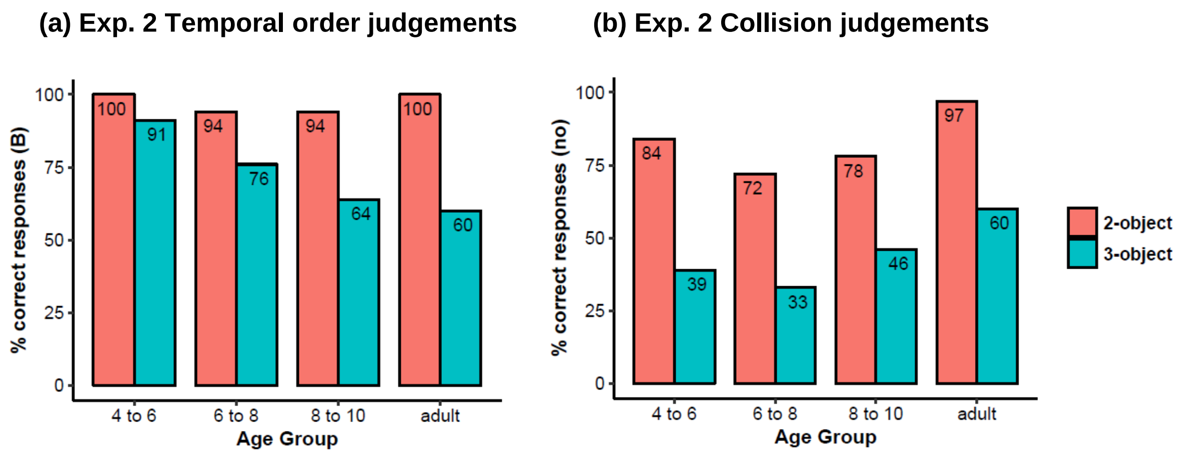

Figure 4. Percentage of participants in each age group who gave the correct response in (a) the temporal order judgement question (square C); and (b) the collision judgement question (no) in the 2-object control clip (red bars/left-hand bar for each age group) and 3-object pseudocollision (blue bars/right-hand bar per age group) of Experiment 2.

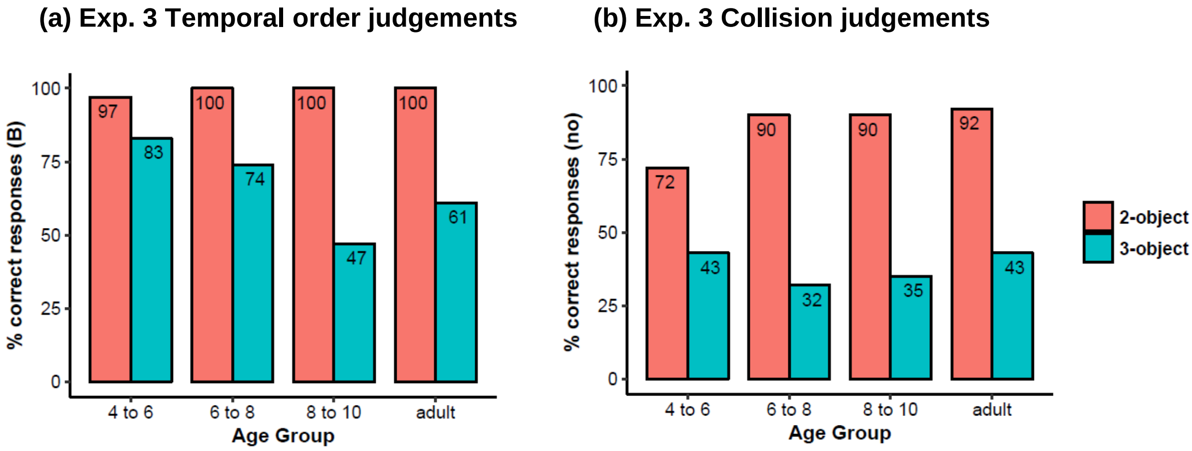

Figure 5. Percentage of participants in each age group who gave the correct response in (a) the temporal order judgement question (square C); and (b) the collision judgement question (no) in the 2-object control clip (red bars/left-hand bar for each age group) and 3-object pseudocollision (blue bars/right-hand bar for each age group) of Experiment 3.

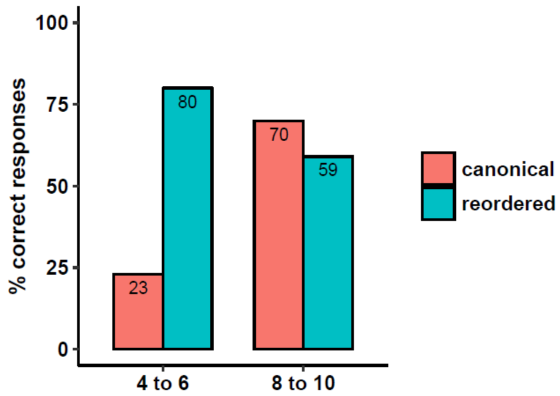

Figure 6. Percentage of participants in each age group of Experiment 4 who gave the correct response for the temporal order judgement question for the canonical collision (red bars/left-hand bar for each age group, correct answer was B) and the reordered collision (blue bars/right-hand bar for each age group, correct answer was C).

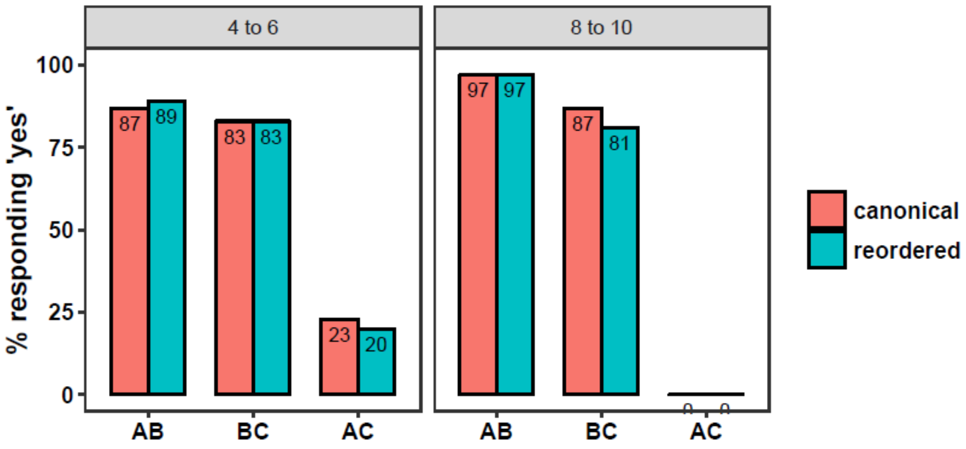

Figure 7. Percentage of participants in each age group who responded 'yes' to each of the three causal impression questions for the canonical collision (red bars/left-hand bar for each age group) and the reordered pseudocollision (blue bars/right-hand bar for each age group).