Developmental Psychology

One- and Two-Year-Olds Grasp That Causes Must Precede Their Effects

Emma C. Tecwyn, Pingki Mazumder, and Daphna Buchsbaum

Online First Publication, May 18, 2023. https://dx.doi.org/10.1037/dev0001551

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

https://doi.org/10.1037/dev0001551

One- and Two-Year-Olds Grasp That Causes Must Precede Their Effects

Emma C. Tecwyn[1], Pingki Mazumder[2], and Daphna Buchsbaum[3]
[1] School of Education, Language and Psychology, York St John University
[2] Department of Psychology, University of Toronto
[3] Cognitive, Linguistic & Psychological Sciences, Brown University

Knowing the temporal direction of causal relations is critical for producing desired outcomes and explaining events. Existing evidence suggests that children start to grasp that causes must precede their effects (the temporal priority principle) by age 3; however, whether younger children also understand this has, to our knowledge, not previously been tested. Given the importance of temporal priority in making sense of the world, we explored when knowledge of this principle develops. In the present study, conducted in a lab or museum in a Canadian city, 1- and 2-year-old children observed an adult perform action *A* on a puzzle box (e.g., spinning a dial), following which an effect *E* occurred (a sticker was dispensed), following which the adult performed action *B* (e.g., pushing a button; *A*–*E*–*B* sequence). In line with the temporal priority principle, toddlers were significantly more likely to manipulate *A* than *B* (Experiment 1, $N = 41$, 22 female), even when *A* was spatially disconnected from the sticker dispenser and further from it than action *B* (Experiment 2, $N = 42$, 25 female). In Experiment 3 ($N = 50$, 25 female), toddlers observed an *A*–*B*–*E* sequence such that both actions *A* and *B* were performed prior to effect *E*. Here, they primarily intervened on *B*, which ruled out that success in Experiments 1–2 was based on a primacy effect. A lack of any age effects across experiments suggests that within the second year of life, children already grasp that causes must precede their effects, providing key insights into causal reasoning in early childhood.

Public Significance Statement
Understanding cause–effect relations—for example, that pressing the switch makes the light come on—is crucial for explaining, predicting, and controlling events in our environment. Our study shows that, within the second year of life, children already grasp a key principle of cause–effect relations—that causes must precede their effects in time. This early-developing causal knowledge enables young children to interact effectively with their environment, and provides a foundation for the development of more sophisticated reasoning skills.

Keywords: causal reasoning, physical causation, spatial contiguity, temporal priority, toddlers

Supplemental materials: https://doi.org/10.1037/dev0001551.supp

The ability to extract and encode causal structure enables us to understand and predict events and control our environment by intervening on it. A large body of research suggests that from a young age, children represent causal structures and use this information to guide their inferences and behavior (see Muentener & Bonawitz, 2017; Sobel & Legare, 2014, for reviews). The concept of causality is fundamental to our understanding of the physical world (Kominsky & Scholl, 2020), enabling us to reason about objects and their interactions (e.g., Baillargeon, 2004), use tools effectively (e.g., Cheke et al., 2012), and solve problems (e.g., Seed & Call, 2014). Beyond this, there is also evidence that causal knowledge contributes to the development of children's cognitive

Emma C. Tecwyn https://orcid.org/0000-0002-2343-2282
We are grateful to the families who participated in this study, and to the Ontario Science Centre and the Royal Ontario Museum for hosting our research. Thank you to Nafisa Mahbub, Nishat Kazi, Samantha Mekhuri, Justine Biado, and Kiah Caneira for providing assistance with data collection.

This work was funded by the Natural Sciences and Engineering Research Council of Canada and the Canada Foundation for Innovation. The authors have no conflict of interest to declare. Data, code, and videos are available at Open Science Framework online at https://osf.io/6m9xr/?view_only=d4494b8c0c134e17abad4864ae8a8077.

This study was not preregistered.

Emma C. Tecwyn served as lead for formal analysis, project administration, visualization, and writing–original draft. Pingki Mazumder served in a supporting role for writing–review and editing. Daphna Buchsbaum served as lead for funding acquisition, resources, supervision, and writing–review and editing and served in a supporting role for formal analysis. Emma C. Tecwyn and Daphna Buchsbaum contributed equally to conceptualization and methodology. Emma C. Tecwyn and Pingki Mazumder contributed equally to data curation and investigation.

Correspondence concerning this article should be addressed to Emma C. Tecwyn, School of Education, Language & Psychology, York St John University, Lord Mayor's Walk, HG/209, Holgate Building, York, YO31 7EX, United Kingdom. Email: emmatecwyn@gmail.com

skills in a variety of less obvious domains, including social reasoning (Buchsbaum et al., 2012), moral reasoning (Hamlin, 2013), and the generation of explanations (Legare, 2012). Thus, it is clear that causality plays a central role in our experience of the world from early in life and is a fundamental building block for more complex cognition.

Despite the central role of causality in our lives, we generally cannot directly sense causal relations and must instead infer causes and effects based on observable evidence (though detection of the causal status of Michottean launching displays provides an exception to this (e.g., Leslie & Keeble, 1987). Although much research—particularly in developmental psychology—has focused on the use of statistical covariation information in causal reasoning (e.g., Gopnik et al., 2001), there are other sources of information that provide evidence for causal relationships, that go beyond what can be inferred on the basis of covariation alone (Bramley et al., 2018). Hume (1739) identified several cues to causality, one of which was the temporal priority principle—that causes must precede their effects in time.

Knowing that causes must precede their effects enables us to understand physical, psychological, and biological causal relations—for example, that pressing the switch makes the light come on; the dog yelped because the person stood on their tail—the person did not stand on the dog's tail because they yelped; the girl cried because she fell and scraped her knee, not because her mom hugged her afterward. Knowledge of the temporal direction of causal relations can also be critical for deciding which variable to manipulate to produce a desired outcome, and for explaining why a particular event occurred (Rottman et al., 2014). Consider, for example, a faucet or mixer tap that is missing the hot and cold labels. When trying to figure out which side is hot you might fiddle with both and try to notice the subsequent change in water temperature. The temporal order of events enables you to figure out which side you need to manipulate to achieve the desired outcome of warmer water.

Given the utility of temporal priority in making sense of the world, there has been significant interest in exploring when knowledge of this principle develops. In an early study by Kuhn and Phelps (1976; Table 1) children were shown pictures of causal events and asked, for example, "did the water spill because the cup fell or did the cup fall because the water spilled?" Results suggested that children might not grasp the temporal priority principle until around 6 years of age; however, the task was linguistically complex, requiring children to choose between the sentences "*A* because *B*" and "*B* because *A*" to describe a series of events. Thus, younger children's failure could be attributed to insufficient verbal knowledge as opposed to a lack of temporal priority understanding. In fact, children as young as 3 years succeeded in a simpler task that required them to point to the first or last picture in a causal chain to answer "why did X happen?" (Kun, 1978; Table 1). However, a limitation of picture tasks is that participants do not see the actual causal events; thus, they are dependent on children's existing knowledge of causal schema.

Another task that has been used in multiple studies is the *A–E–B* paradigm (Bullock & Gelman, 1979; Rankin & McCormack, 2013; Shultz & Mendelson, 1975; Sophian & Huber, 1984). In this task, children are typically shown a sequence of novel mechanical events, consisting of potential cause *A* (e.g., a ball rolling down a ramp at one end of a box), followed by effect *E* (e.g., a jack-in-a-box popping up), followed by potential cause *B* (e.g., a ball rolling down a ramp at the other end of the box). Children are then asked to make a judgment about which potential cause (*A* or *B*) resulted in the effect, either explicitly (e.g., "which ball made the jack come up?") or by asking them to generate the effect themselves (e.g., "can you make the jack come up?").

Shultz and Mendelson (1975) reported that children only reliably selected the preceding event (*A*) as causal from 4 years of age, whereas 3-year-olds were more likely to attribute causality to the event that followed the effect (*B*), suggesting that children of this age were perhaps responding based on a recency effect. However, this task also used some linguistically complex questions, for instance asking children to predictively judge "If *A* is present would *X* happen." Subsequently, the majority of studies using the *A–E–B* paradigm have concluded, like Kun (1978), that children grasp temporal priority from 3 years, though performance has typically been found to improve with age. Poorer performance by 3-year-olds compared with older children has been attributed to

Table 1

Summary of Previous Studies Investigating the Development of Knowledge of the Temporal Priority Principle (Arranged in Chronological Order of Publication Date)

Study	Ages tested	Task	Measure	Evidence for grasp of temporal priority?
Shultz and Mendelson (1975, temporal-order problem)	3- to 4-, 6- to 7-, and 9- to 11-year-olds	*A–E–B* (3 different mechanical tasks)	Three verbal judgments (e.g., "What makes the bell ring?" "If A is present will X happen?") and one action (e.g., "Make the bell ring")	From 4 years
Kuhn and Phelps (1976)	5- to 8-year-olds	Sentence choice	Verbal judgment/pointing—point to the sentence that "goes best with the picture" (of a causal event)	From 6 years; improved performance across age groups
Kun (1978)	3- to 8-year-olds	Causal chain picture sequence (ABC)	Pointing—why did B happen? Point to picture A or C	From 3 years
Bullock and Gelman (1979)	3- to 5-year-olds	*A–E–B* (balls and ramps)	Verbal judgment ("Which ball made the jack come up?") and action ("Make the jack come up")	From 3 years; improved performance across age groups
Sophian and Huber (1984, Experiment 2, order-only problems)	3- and 5-year-olds	*A–E–B* (social task and mechanical task)	Verbal judgment (e.g., "Which light made the dog turn?")	From 3 years; improved performance across age groups
Rankin and McCormack (2013)	3- and 4-year-olds	*A–E–B* (balls and ramps)	Verbal judgment ("Which ball made the teddy come up?")	From 3 years; improved performance across age groups

inferior memory or language comprehension skills, or a weaker grasp of temporal priority; for example, younger individuals may have some awareness of the role of temporal order in cause–effect relations but lack understanding that it provides a hard constraint, and thus their knowledge is insufficient to appropriately guide their behavior (Bullock & Gelman, 1979; Rankin & McCormack, 2013; Sophian & Huber, 1984; Table 1).

Based on the existing evidence we can conclude that children already have at least an emerging grasp of temporal priority by age three. However, whether children under age three have an understanding of temporal priority has, to our knowledge, not previously been tested. There is reason to expect that children younger than 3 years might be capable of identifying the direction of causal relations on the basis of temporal information. From infancy, children receive rich environmental input in terms of examples of cause–effect relationships, particularly by observing others interacting with causal systems (e.g., pressing a switch makes the light come on; pressing the remote makes the TV come on, e.g., Sim & Xu, 2017). As soon as infants are able to act on the world, for example, via reaching and grasping from around 6 months of age, they experience the causal effects of their own actions (e.g., pressing the button on the toy makes it play music). It has been argued by some (e.g., Gopnik et al., 2004) that children might even have some innate assumptions about causal structure, or "starting-state" theories (e.g., Gopnik & Wellman, 1994). For this to be true, we would expect basic aspects of causality like the temporal priority principle to be present early in life if not from birth, but based on current evidence this is actually somewhat controversial.

There is growing experimental evidence that toddlers possess a sophisticated ability to learn and reason about causal relations in a variety of tasks. By 2 years of age, children can make appropriate causal inferences on the basis of statistical contingency, and design appropriate causal interventions to bring about desired effects (e.g., Gopnik et al., 2001; Meltzoff et al., 2012). They can reason about conditional independence (Sobel & Kirkham, 2006); learn causal rules based on abstract relations (Walker & Gopnik, 2014) and higher-order generalizations (Sim & Xu, 2017); and apply learned causal functions to solve novel problems (Goddu & Gopnik, 2020). In some specific cases, toddlers even outperform preschoolers at learning causal rules (Walker et al., 2016). However, whether toddlers make causal inferences based on the temporal order of events has not, to our knowledge, been addressed explicitly. This study aims to address this gap and add to our understanding of the developmental emergence of children's ability to reason about causal systems; specifically in relation to how temporal order information constraints causal inferences.

Notably, these previous studies that have revealed toddlers' impressive causal reasoning skills used observational learning paradigms and action-based measures, rather than measuring children's explicit judgments. They also involved minimal verbal instruction, thus circumventing issues of language comprehension and the cognitive demands of verbalizing a response. Therefore, in the present study, we investigated 1- and 2-year-olds' grasp of the temporal priority principle using an *A–E–B* paradigm. Although the *A–E–B* paradigm requires minimal verbal ability compared to some other methods (e.g., Kuhn & Phelps, 1976), children in previous *A–E–B* tasks have still been asked to make explicit judgments about which event was the causal one. Thus, the main measure used in previous studies still depended on causal language comprehension, which may be inappropriate for toddlers, and could explain the reported poorer performance of 3-year-olds compared with older children in earlier research. Therefore, in the present study, we used an action measure. An additional advantage of an action measure is that it tells us whether children can use their knowledge to generate appropriate interventions—if you know that *A* causes *E*, you should be able to act on *A* to bring about *E* (Walker & Gopnik, 2014).

Across three experiments, toddlers watched an adult manipulate two actions (*A* and *B*) on a puzzle box, and saw a sticker being dispensed from the box (effect *E*). The spatial relationship between these three events, or their temporal order, differed between experiments. Toddlers were then allowed to interact with the box to retrieve stickers for themselves. As an initial test of their grasp of temporal priority, in Experiment 1 toddlers saw a demonstration of *A–E–B*, such that *A* was performed temporally prior to the sticker being dispensed, and *B* was performed after this effect occurred. Actions *A* and *B* were both on the main puzzle box, equidistant from and spatially contiguous with the sticker dispenser (Figure 1A) so that potential spatial and physical cues to causality were controlled for. In Experiment 2, as a stronger test of temporal priority understanding, toddlers again saw a demonstration of *A–E–B*, but unlike in Experiment 1, action *A* was physically disconnected from and spatially discontiguous with the sticker dispenser, and a greater distance away from it compared with *B* (Figure 1B). This meant that temporal order information was pitted against another cue to causality—spatial contiguity. Finally, to examine whether toddlers might be acting on the basis of a primacy effect (performing the first action they saw the demonstrator interact with, i.e., *A*) rather than acting in accordance with temporal priority, in Experiment 3, they saw a demonstration of *A–B–E* on the Experiment 1 puzzle box. Here, both actions *A* and *B* were temporally prior to the sticker being dispensed and *A* was still the first action performed by the demonstrator, but temporal contiguity information pointed to action *B* as the more likely cause.

Experiment 1: Temporal Priority

In Experiment 1, toddlers observed an adult perform action *A* on a puzzle box (e.g., spinning a dial), following which an effect *E* occurred (a sticker was dispensed), following which the adult performed action *B* (e.g., pulling a lever; Figure 1A). Toddlers could then interact with the puzzle box and their interventions (manipulations of *A* and *B*) were recorded. Actions *A* and *B* were both on the same box that dispensed the sticker and were equidistant from the sticker dispenser (Figure 1A).

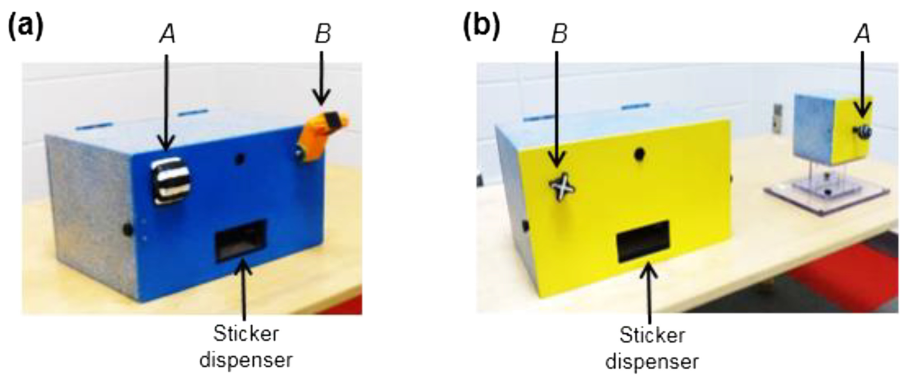
If toddlers behave randomly—that is, if they are equally likely to manipulate *A* and *B*—then this could suggest that they do not grasp temporal priority, or that they have difficulty remembering the sequence of events demonstrated. If they primarily manipulate *B*, this might indicate a recency effect, as was suggested to be the case for younger 3-year-olds by Shultz and Mendelson (1975). Finally, if toddlers primarily manipulate *A*—that is, the action that precedes the effect, this would be indicative of a grasp of temporal priority, and the ability to intervene appropriately based on this knowledge.

Materials and Method

Approval for this research (Experiments 1–3) was granted by the Institutional Research Ethics Board for Human Subjects at the

Figure 1

Examples of the Puzzle Boxes Used in (a) Experiments 1 and 3 Where A and B Were Equidistant and Both Spatially Contiguous With the Sticker Dispenser and (b) Experiment 2 Where A Was Physically Disconnected From and Spatially Discontiguous With the Sticker Dispenser and a Greater Distance Away From It Compared With B



Note. See the online article for the color version of this figure.

University of Toronto (Protocol 30,903, "The Development of Casual and Social Learning in Children").

Participants

We aimed to collect data from at least 20 1-year-olds and 20 2-year-olds. The exact number of participants (when we reached the minimum) was determined by availability in testing locations; specifically, we did not turn away children who wanted to participate while we were in a given setting, even once we reached our threshold. Forty-one 12- to 35-month-olds completed Experiment 1; twenty 12- to 23-month-olds (10 male, mean age = 18.6 months) and twenty-one 24- to 35-month-olds (nine males, mean age = 28.8 months). Participants were recruited from the university's database or at local museums. The caregivers of 26 of the participants completed an optional demographic questionnaire. Ten identified their children as White, one East Asian, four South Asian, two Middle Eastern, and nine multiracial. An additional 16 toddlers were tested but their data were excluded for the following reasons: failed to start interacting with puzzle box during predetermined timeframe of 2 min (three); acted but did not reach criterion of activating puzzle box five times (five; two of these also had caregiver interference and one was noted as being distracted/losing interest); puzzle box malfunctioned (three); caregiver interference (three); sibling interference (one); and experimenter error (one).

Stimuli and Testing Setup

We used a custom-made puzzle box (36 cm × 20 cm × 23 cm) that dispensed stickers. The puzzle box had three different-colored (blue, red, yellow) interchangeable front panels, each of which had a different pair of actions on the front (Figure 1A; see Table S1 in the online supplemental materials for details of puzzle-box actions). The assignment of panel, actions as *A* and *B*, and the side on which action *A* appeared (left or right) were counterbalanced. The box contained a remotely operated sticker dispenser, which meant that either action could be the "causal" one, because the experimenter covertly triggered sticker release when the participant manipulated the necessary action. Toddlers were tested in a museum in the Toronto area or a lab at the University of Toronto. The caregiver was present during testing but remained seated behind the participant during the session.

Procedure

Demonstration Phase. Two experimenters were involved in running the experiment: one who acted as the demonstrator (E1), and one who accompanied the toddler (E2). At the start of the experiment, E2 and the child entered the testing area to find E1 "busy" looking at a clipboard. E2 said to the subject "Hmmm, it looks like she [E1] is still getting ready, but we can wait here, and we can watch" and the child was encouraged to sit on a chair which faced toward the puzzle box (Figure 1A) at a distance of approximately 1 m. E2 directed the child's attention to E1 and the puzzle box as necessary—for example, if they became distracted or stopped attending, E2 pointed to the setup and said "shall we watch?" E1 stood behind the puzzle box and performed an intentional *A–E–B* demonstration as follows: without engaging in eye contact with the participant or E2, E1 intentionally performed action *A*, following which a tube containing a sticker was dispensed from the center of the box (effect *E*), following which they intentionally performed action *B*. E1 saw the outcome of their actions, but did not verbally acknowledge it or pick up the sticker (Video S1 in the online supplemental materials, available at https://osf.io/6m9xr/?view_only=d4494b8c0c134e17abad4864ae8a8077). We chose to have the demonstrator act intentionally but not pedagogically (i.e., the demonstrator did not use the child's name or make eye contact with them), because we did not want ostensive cues to lead children to faithfully copy all actions, or "overimitate" (see Hoehl et al., 2019; Keupp et al., 2018 for recent reviews of the overimitation literature). This intentional demonstration was repeated until the subject had attended to two full demonstrations (defined as the child's gaze being oriented toward the apparatus during the demonstration, as confirmed by E2). E1 then looked to the child and said "Oh hey, I'm all done here! You can have a turn, and you can have these stickers!" (indicating stickers dispensed from the puzzle box). E1 then left the testing area.

Response Phase. E2 approached the puzzle box with the child. Each time the child manipulated *A*, a sticker was dispensed (which constituted an activation of the puzzle box), up to a maximum of five times. Nothing happened if they manipulated *B*. If the child did not spontaneously interact with the puzzle box, E2 provided neutral encouragement such as "It's your turn, you can try anything!" E2 commented on the outcome (e.g., "oh, a sticker!") but deliberately avoided any causal language (e.g., "you made the sticker come out"). After the fifth sticker had been dispensed, E2 said "I think that's the end of your turn for today!" and moved the puzzle box out of reach. If the child did not complete five activations or failed to interact with the puzzle box for >2 min the session was ended.

Scoring and Analysis

Data and code are available at https://osf.io/6m9xr/?view_only=d4494b8c0c134e17abad4864ae8a8077. All sessions were scored live. Where consent was given, sessions were also videotaped and behavior was re-coded from the footage. We first coded all of the participant's manipulations of *A* and *B* (e.g., *A–A–A–B–B–A–A*). From this raw data, we coded:

1. First Intervention. We looked at whether children's very first action in the response phase was on *A* or *B*. This enabled us to assess children's causal inferences before they received any feedback based on the outcomes of their own actions. Binomial tests were used to test whether the first manipulation (*A* or *B*) differed significantly from random responding (chance = 0.5). Binary logistic regression with age in months as a continuous predictor was used to test whether children's very first manipulation of the puzzle box (*A* or *B*) varied across the age range tested.

2. First Two Interventions. Children often copy causally irrelevant actions ("overimitation"; see Hoehl et al., 2019; Keupp et al., 2018 for recent reviews), including those performed after an outcome is achieved (Nielsen et al., 2015). Toddlers are less likely to overimitate than older children (Chudek et al., 2016; McGuigan & Whiten, 2009; Tecwyn et al., 2020), and we deliberately chose for the demonstrator to act intentionally rather than pedagogically to reduce the likelihood of faithful copying. However, if we are to claim that the children in the present study are acting on the basis of temporal priority it is important to rule out that toddlers who act on *A* first are doing so simply because they are faithfully copying all of the demonstrator's actions.

Specifically, if toddlers' behavior is guided by the temporal priority principle, then toddlers should be more likely to perform *A* followed by another *A* (*A–A* sequence) than *A* followed by *B* (*A–B* sequence), because after performing *A* and getting a sticker, they must perform *A* again to get another sticker. In contrast, imitative behavior would lead to *A–B,* as toddlers would perform the sequence just as the demonstrator did. We used chi-square goodness of fit tests to see whether the performance of the four possible sequences differed from chance-level, followed by binomial tests to establish whether the subset of toddlers who manipulated *A* first went on to manipulate *A* or *B* next (i.e., the dependent variable was this subset of toddlers' second action performed). Binary logistic regression with age in months as a continuous predictor was used to test whether the tendency to manipulate *A* again versus *B*, after an initial manipulation of *A*, varied across the age range tested. This was of interest because "overimitation" is known to increase across early childhood (Chudek et al., 2016; McGuigan & Whiten, 2009).

3. First Intervention Per Activation. A puzzle-box activation occurred each time the participant manipulated *A,* regardless of what actions had preceded it, and resulted in the box dispensing a sticker. We recognize that examining performance on trial 1—in this case, toddlers' very first intervention—is often considered most informative in relation to participants' "naïve" inferences (though see e.g., Barker & Povinelli, 2019 for a recent argument against privileging trial 1 data). However, based on the previous literature showing variable performance of 3-year-olds in temporal priority tasks, we did not know whether toddlers would succeed on this task at all, let alone on their first intervention. Looking at whether children first acted on the causal action *A* or on the noncausal action *B* across all five activations of the puzzle box enabled us to (a) increase power for detecting an overall ability to act in accordance with temporal priority and (b) detect learning, to help establish whether the causal relationship (or association) was learnable for this age group.

To analyze these data, we used mixed effects logistic regression with a random intercept per child, and assumed a fixed slope across participants. The dependent variable was whether participants manipulated *A* or *B* first on each activation, gender (male or female) was included as a fixed factor, age in months and puzzle-box activation (1–5) as centered continuous predictors, and participant ID as a random factor. Where factors and predictors were found not to be significant, they were removed from the model, and simplified versions of the model were compared with more complex versions using the difference in Akaike information criterion (AIC) scores. AIC scores allow selection of the model that best explains the data, while accounting for model complexity, with $\Delta\text{AIC} > 2$ generally considered strong support for the higher-scoring model. We also ran an intercept-only model to compare overall performance to chance (likelihood of acting on *A* vs. *B* first).

Results

Interobserver Reliability

Consent for videotaping was given for 35 out of 41 participants in Experiment 1. Following initial coding of all videos, to assess interobserver reliability, 30% of these recorded sessions were coded by a second observer (an undergraduate research assistant) who was blind to the purpose of the study. The observers agreed perfectly for all three of our measures (first intervention per activation; first intervention; first two interventions (Cohen's kappa = 1.0 for all).

Preliminary Analyses

Toddlers did not show a preference for manipulating any particular actions, one-way analysis of variance (ANOVA) with average number of puzzle-box activations where *A* was intervened on first as the dependent variable, and demonstrated action type (e.g., pull-knob, turn-dial) as a between-subjects factor, $F(5, 35) = 2.39$, $p = .81$, and the side of the puzzle box on which action *A* appeared (left or right) did not affect toddlers' tendency to intervene on it (independent samples *t*-test: $t = 0.71$, $df = 39$, $p = .48$).

To rule out that differences in temporal contiguity between events *A–E* versus *E–B* could have influenced toddlers' behavior (specifically, if action *A* was closer to effect *E* in time than *B* was, then toddlers could have chosen the "correct" action [*A*] on the basis of temporal contiguity rather than temporal priority). To examine this, 50% of the recorded sessions were recoded for event timings by an observer who was blind to the study hypotheses (an undergraduate research assistant). Specifically, the timing of the following events was recorded: (a) start of demonstrator contact with action *A*; (b) the sticker being dispensed into the tray; and (c) start of demonstrator contact with action *B*. From these timings, the duration between *A*–*E* and the duration between *E*–*B* were calculated. Each participant watched two demonstrations, so mean durations for the two intervals were calculated. Based on these values, mean *A*–*E* and *E*–*B* durations were calculated. The *E*–*B* interval was significantly shorter than the *A*–*E* interval ($M_{A-E} = 1{,}508$ ms, $M_{E-B} = 871$ ms, $t = 4.01$, $df = 20$, $p < .001$), which suggests that in preferentially intervening on *A* toddlers were not acting on the basis of temporal contiguity; otherwise, they should have shown a preference for action *B*.

First Intervention

The very first-time toddlers acted on the box, they were significantly more likely to manipulate the action that preceded the effect (*A*) than the action that followed the effect (*B*; 29 out of 41 manipulated *A* first, binomial test: $p = .01$; see Table S2 in the online supplemental materials for the first intervention by participants who were excluded for not completing five puzzle-box activations). Logistic regression with age in months as a centered continuous predictor showed that first intervention was not influenced by toddlers' age ($\beta_{age} = 0.24$, $SE = 0.35$, $p = .48$).

First Two Interventions

The majority of toddlers' first two interventions were *A*–*A* (26 out of 41), and performing *A* followed by *B* was the least common response (three out of 41 instances; Table 2). This was significantly different from chance, chi-square goodness of fit test: $\chi^2(3) = 32.85$, $p < .001$. Of those toddlers who manipulated *A* first, they were significantly more likely to manipulate *A* again next (first two interventions = *A*–*A*, 26 out of 29) than to manipulate *B* next (first two interventions = *A*–*B*, three out of 29, binomial test: $p < .001$) and this behavior was not significantly influenced by age ($\beta_{age} = -0.78$, $SE = 0.72$, $p = .27$).

First Intervention Per Activation

Figure 2 shows toddlers' interventions across all five activations of the puzzle box. Mixed effects logistic regression revealed no effect of gender ($\beta_{gender} = -0.01$, $SE = 0.44$, $p = .98$) or age in months ($\beta_{age} = -0.20$, $SE = 0.44$, $p = .65$, Figure S1 in the online supplemental materials) on performance and the model AIC was larger with these factors included (AIC = 177.1) so they were removed and not considered further. The simplified model (AIC = 173.4) showed that the likelihood of intervening on *A* first increased across activations ($\beta_{activation} = 0.57$, $SE = 0.24$, $p = .02$; Figure 2) and the intercept-only model revealed that toddlers were significantly more likely to intervene on *A* than *B* ($\beta_{intercept} = 2.34$, $SE = 0.50$, $p < .01$).

Table 2
Toddlers' First Two Interventions on the Puzzle Box in Experiments 1–3

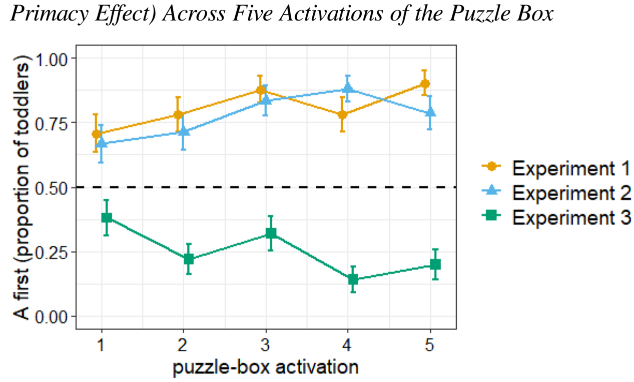
First two interventions	Experiment 1 (*A–E–B* connected)	Experiment 2 (*A–E–B* disconnected)	Experiment 3 (*A–B–E* connected)
A–A	26	21	6
A–B	3	7	13
B–A	6	9	7
B–B	6	5	24

Discussion

The results of Experiment 1 provide evidence that 1- and 2-year-olds act in accordance with temporal priority. Toddlers were more likely to intervene on action *A* than *B* on their very first interaction with the puzzle box, as well as across all five activations. By examining toddler's first two interventions we can also be confident that they were not simply copying the actions of the demonstrator (otherwise we would have expected *A*–*B* to be the most common first two interventions, as opposed to *A*–*A*). Given that toddlers did not intervene randomly or preferentially manipulate *B*, we can also rule out that their memory skills are insufficient to recall what they observed, or that they acted on the basis of a recency effect, both of which have been offered as explanations for 3-year-olds' poorer performance in previous studies (Bullock & Gelman, 1979; Rankin & McCormack, 2013; Sophian & Huber, 1984). Furthermore, as there was a larger temporal interval between *A* and *E* than between *E* and *B*, it was not the case that toddlers were using temporal contiguity as a cue to causality.

We found no evidence that performance was influenced by age, suggesting that toddlers already grasp temporal priority and act in accordance with this principle within the second year of life. This result contrasts with the findings of previous studies suggesting that knowledge of temporal priority improves between the ages of 3 and 4 years (Rankin & McCormack, 2013, Shultz & Mendelson, 1975). We propose that these findings can be reconciled based on the different measures used. Whereas previous studies investigating temporal priority knowledge asked children to make explicit judgments about which event was causal, which requires verbal

Figure 2
Mean Proportion of Toddlers Who Intervened on A *First (as Opposed to* B*), in Experiment 1 (Temporal Priority), 2 (Temporal Priority vs. Spatial Contiguity), and 3 (Temporal Priority vs. Primacy Effect) Across Five Activations of the Puzzle Box*



Note. Error bars are standard error of measurement and the dashed line represents choosing randomly between *A* and *B*. See the online article for the color version of this figure.

knowledge and poses additional cognitive demands, the present study exploited toddlers' spontaneous behavioral interventions, as was the case for other studies that have revealed impressive causal reasoning skills in this age group (e.g., Goddu & Gopnik, 2020; Walker et al., 2016).

Although the findings of Experiment 1 provide compelling evidence that 1- and 2-year-olds use temporal priority to identify a cause, of additional interest is how robust their understanding of temporal priority is, and whether they continue to act in accordance with temporal priority when other conflicting cues to causal structure are available. One additional cue to causal structure, also identified by Hume (1739), is spatial contiguity, which dictates that a physical cause and its effect should be connected by a spatially continuous path. Children appear to be sensitive to spatial contiguity from a young age, as 2-year-olds' causal inferences are sensitive to contact relations between cause and effect (Bonawitz et al., 2010) and 1-year-olds are able to solve simple means-end problems by attending to the physical connectedness of objects (Brown, 1990; Willatts & Rosie, 1989). Furthermore, 18- to 30-month-olds were more likely to copy an action that preceded an outcome in an observational causal learning task when the events were spatially contiguous, compared with when they were spatially discontiguous (Tecwyn et al., 2020).

In Experiment 2, we provided toddlers with a more stringent test of their grasp of temporal priority by presenting a scenario in which temporal priority was pitted against spatial contiguity. Toddlers saw a demonstration where action *A* preceded the sticker being dispensed, but was performed on a physically disconnected box, whereas action *B* followed the effect in time but was spatially contiguous with it on the main box (see Figure 1B). Previous work with a similar paradigm found that when 3- to 5-year-olds were presented with a conflict between spatial contiguity and temporal priority, they attributed the spatially discontiguous preceding event as the cause, as opposed to the spatially contiguous event that followed the effect (see e.g., Figure 1B; Bullock & Gelman, 1979). Thus, while preschool-age children appear to understand that temporal order provides a hard constraint regarding what could have caused what because causes *must* precede their effects (Bramley et al., 2018), and this, therefore, overrides spatial contiguity in their causal inferences, it is unclear how toddlers may behave when given additional conflicting cues to causality.

If toddlers grasp that temporal order places a hard constraint on cause–effect relations, then they should primarily intervene on *A* in our spatially discontiguous setup (*A–E–B*, with *A* performed on a separate box; see Figure 1B). If they have some knowledge of temporal priority but are perhaps still incorporating the inviolability of this constraint into their causal reasoning (Rankin & McCormack, 2013), then they might be influenced by spatial cues and be more likely to intervene on *B* in Experiment 2 than in Experiment 1. Finally, if toddlers give more weight to the cue of spatial contiguity than temporal priority in this paradigm then they should be more likely to intervene on *B* than *A*.

Experiment 2: Temporal Priority Versus Spatial Contiguity

Materials and Method

Participants

Forty-two 12–35-month-olds who had not participated in Experiment 1 completed Experiment 2; twenty 12- to 23-month-olds (seven male, mean age = 18.8 months) and twenty-two 24- to 35-month-olds (10 males, mean age = 29.4 months). The caregivers of 35 of these participants completed an optional demographic questionnaire. Seventeen identified their children as White, one Black, four East Asian, one South East Asian, two South Asian, one Middle Eastern, eight multiracial, and one other (West Indies). An additional 14 toddlers were tested but their data excluded for the following reasons: failed to start interacting with puzzle box during predetermined timeframe of 2 min (five); acted but did not reach criterion of activating puzzle box five times (five; four of these were also noted as being distracted/disinterested); caregiver interference (four).

Stimuli and Testing Setup

Experiment 2 involved two puzzle boxes: the main puzzle box (the same box that was used in Experiment 1) but with a single action on it, which served as action *B*, and an additional small puzzle box, which also had a single action on it (*A*). There was a gap of ~20 cm between the two boxes (Figure 1B, Video S2 in the online supplemental materials, available at https://osf.io/6m9xr/?view_only=d4494b8c0c134e17abad4864ae8a8077). Thus, *A* was now spatially discontiguous with and physically disconnected from the sticker dispenser, whereas *B* was spatially contiguous with and closer to it.

Procedure

The procedure was the same as for Experiment 1.

Results

Interobserver Reliability

Consent for videotaping was given for 39 out of 42 participants in Experiment 2. Following full coding of these recorded sessions, 30% were coded by a second observer (an undergraduate research assistant) who was blind to the purpose of the study. The observers agreed perfectly for all three of our measures (first intervention per activation; first intervention; first two interventions (Cohen's kappa = 1.0 for all).

Preliminary Analyses

Toddlers did not show a preference for manipulating any particular actions, one-way ANOVA: $F(5, 36) = 0.42$, $p = .83$, and the side of the puzzle box on which action *A* appeared (left or right) did not affect toddlers' tendency to intervene on it (independent samples *t*-test: $t = 0.15$, $df = 40$, $p = .88$).

As for Experiment 1, based on the re-coding of 50% of the Experiment 2 videos, the *E–B* interval was significantly shorter than the *A–E* interval ($M_{A-E} = 1{,}673$ ms, $M_{E-B} = 1{,}097$ ms, $t = 3.77$, $df = 22$, $p = .001$). This suggests that in preferentially intervening on *A* toddlers were not acting on the basis of temporal contiguity; otherwise, they should have shown a preference for action *B*.

First Intervention

The very first-time toddlers intervened on the spatially discontiguous puzzle box they were significantly more likely to manipulate *A* than *B* (28 out of 42 manipulated *A* first, binomial test: $p = .04$; see

Table S2 in the online supplemental materials for the first intervention by participants who were excluded for not completing five puzzle-box activations). Logistic regression with age as a continuous predictor showed that initial behavior was not influenced by toddlers' age ($\beta_{age} = -0.04$, $SE = 0.33$, $p = .89$).

First Two Interventions

Half of the toddlers' first two interventions were *A–A* (21 out of 42, Table 2), which was significantly different from chance, chi-square goodness of fit test: $\chi^2(3) = 14.76$, $p = .002$. Of the 28 toddlers who manipulated *A* first, it was significantly more likely that they manipulated *A* again next (first two manipulations = *A–A*, 21 out of 28) as opposed to manipulating *B* next (first two manipulations = *A–B*, seven out of 28, binomial test: $p = .01$). This behavior was not significantly influenced by age ($\beta_{age} = -0.36$, $SE = 0.46$, $p = .42$).

First Intervention per Activation

Figure 2 shows how toddlers intervened across all five activations of the puzzle box. Mixed effects logistic regression showed no effect of gender ($\beta_{gender} = -0.01$, $SE = 0.26$, $p = .96$) or age in months ($\beta_{age} = -0.18$, $SE = 0.26$, $p = .49$, Figure S1 in the online supplemental materials) on performance and model AIC was larger with these factors included (220.6) so these factors were removed and not considered further. The simplified model (AIC = 217.1) showed that there was an increasing tendency to manipulate *A* first across activations ($\beta_{activation} = 0.40$, $SE = 0.19$, $p = .03$; Figure 2) and the intercept-only model revealed clear evidence that toddlers primarily intervened on the temporally prior action: they were significantly more likely to manipulate *A* than *B* ($\beta_{intercept} = 1.51$, $SE = 0.28$, $p < .001$).

Discussion

As in Experiment 1, toddlers in Experiment 2 demonstrated a grasp of temporal priority, even under more stringent conditions where temporal priority was pitted against salient spatial cues. Specifically, despite *A* being on a spatially discontiguous box, and a greater distance from the sticker dispenser than *B*, toddlers were significantly more likely to intervene on *A* than *B* on their very first intervention, as well as across their five activations of the puzzle box. This is in spite of the fact that previous research has demonstrated sensitivity to spatial contiguity as an indicator of causal plausibility in this age group (Tecwyn et al., 2020). Furthermore, there was a larger temporal interval between *A* and *E* than between *E* and *B*, so it was not the case that toddlers were using temporal contiguity as a cue to causality.

As in Experiment 1, we found no evidence for behavior changing across the age range tested. Thus, these data extend the findings of Bullock and Gelman (1979) with 3- to 5-year-olds to 1- to 2-year-olds, suggesting that already within the second year of life temporal priority trumps another strong cue to causality—spatial contiguity. This is also in keeping with studies with older children and adults showing that when temporal order is pitted against other information (e.g., statistical covariation), these groups also tend to prioritize temporal information when making causal judgments (e.g., Burns & McCormack, 2009; Lagnado & Sloman, 2006; Schlottmann, 1999).

While Experiments 1–2 provide evidence that toddlers have a comparable grasp of temporal priority to that found in older children, an alternative explanation for the observed results remains: namely, that the data could be explained by a primacy effect, where participants reproduce the first action of the demonstrator (i.e., *A*). More subtly, it is possible that children imitate whatever actions the demonstrator does up until the effect occurs, which would also result in copying just action *A*. To address these concerns, in Experiment 3, toddlers watched an adult perform *A* followed by *B* followed by a sticker being dispensed (effect *E*), so that *A* was still the first action, but both actions *A* and *B* were temporally prior to the effect, and action *B* immediately preceded the effect (*B* was temporally contiguous with the effect) and was the causally necessary action.

If toddlers are acting on the basis of a primacy effect in this paradigm, then they should primarily intervene on *A* as in Experiments 1 and 2, because this is the first action performed by the demonstrator. If they are instead using temporal information, then in Experiment 3, they should be more likely to intervene on action *B* first than in Experiment 1, which used the same puzzle-box setup, because in Experiment 3 both temporal priority and temporal contiguity (the closeness of two events in time, Hume, 1739) cues point to action *B* as the most likely single cause. Finally, it is possible that toddlers will copy the sequence of both actions (*A–B*) in Experiment 3, imitating all of the demonstrator's actions that occur prior to the effect. The inference that both *A* and *B* are causally necessary would be valid, as this would be in line with temporal priority information (both actions precede the effect), and it is ambiguous whether *A* is causally necessary in this case. However, a previous study showed that toddlers tend not to reproduce sequences of two actions in a comparable setup (Tecwyn et al., 2020), and they may in fact be limited in their ability to learn causal action sequences, even when they observe unambiguous evidence that a sequence is in fact causally necessary (Tecwyn et al., 2021). Nonetheless, if toddlers copy the sequence *A–B* this would lead to them acting on *A* first, working against our prediction that they will prefer action *B* in this case.

Experiment 3: Temporal Priority Versus Primacy Effect

Materials and Method

Participants

Fifty 12- to 35-month-olds who had not participated in Experiment 1 or 2 completed Experiment 3; twenty-four 12- to 23-month-olds (14 males, mean age = 18.5 months) and twenty-six 24- to 35-month-olds (11 males, mean age = 29.0 months). The caregivers of 26 of these participants completed an optional demographic questionnaire. Fourteen identified their children as White, one East Asian, one South East Asian, five South Asian, two Middle Eastern, and three Latin American. An additional 21 toddlers were tested but their data excluded for the following reasons: failed to start interacting with puzzle box during predetermined timeframe of 2 min (three); acted but did not reach criterion of activating puzzle box five times (six; one of these also had caregiver interference and three were noted as distracted/disinterested); technical issues (six); caregiver interference (three); and experimenter error (three).

Stimuli and Testing Setup

The Experiment 1 puzzle box was used.

Procedure

The procedure was the same as for Experiments 1–2, except that participants watched demonstrations of *A–B–E* (Video S3 in the online supplemental materials, available at https://osf.io/6m9xr/?view_only=d4494b8c0c134e17abad4864ae8a8077), and each time *B* was manipulated a sticker dispensed, which constituted an activation. Nothing happened when *A* was manipulated.

Results

Interobserver Reliability

Consent for videotaping was given for all 50 participants in Experiment 3. Following initial coding of the full set, 30% of these sessions were coded by a second observer (an undergraduate research assistant) who was blind to the purpose of the study. The observers agreed perfectly for all three of our measures (first intervention per activation; first intervention; first two interventions (Cohen's kappa = 1.0 for all).

Preliminary Analyses

Toddlers did not show a preference for manipulating any particular actions, $F(5, 44) = 1.07$, $p = .39$, and the side of the puzzle box on which action *A* appeared (left or right) did not affect toddlers' tendency to intervene on it ($t = 1.89$, $df = 48$, $p = .06$). Event timings were not coded for Experiment 3 given that both actions *A* and *B* occurred prior to effect *E*, and *B* was clearly more temporally contiguous with *E*.

First Intervention

Toddlers were more likely to first manipulate *B* than *A* on their very first attempt, but their behavior did not differ significantly from chance (31 out of 50 manipulated *B* first, binomial test: $p = .12$; see Table S2 in the online supplemental materials for the first intervention by participants who were excluded for not completing five puzzle-box activations). Logistic regression with age as a continuous predictor showed that initial behavior was not significantly influenced by toddlers' age ($\beta_{age} = 0.54$, $SE = 0.32$, $p = .09$).

First Two Interventions

B–B were the most frequent first two interventions (24 out of 50), followed by *A–B* (13 out of 50, Table 2), which significantly differed from chance, chi-square goodness of fit test: $\chi^2(3) = 16.40$, $p < .001$. The conditional analysis from Experiments 1 and 2 (to see what the subset of children who performed *A* first did next) was not done for Experiment 3, as the aim previously was to rule out that toddlers might be "overimitating" following the *A–E–B* demonstration rather than acting on the basis of temporal priority. In Experiment 3, most toddlers manipulated *B* first, plus performing *A* followed by *B* would be an appropriate behavioral response here given that both actions preceded the outcome (*A–B–E*) and the causal necessity of *A* was ambiguous.

First Intervention per Activation

Figure 2 shows how toddlers intervened across all five activations of the puzzle box. Mixed effects logistic regression showed no effect of gender ($\beta_{gender} = -0.34$, $SE = 0.22$, $p = .13$) or age in months ($\beta_{age} = 0.39$, $SE = 0.22$, $p = .08$, Figure S1 in the online supplemental materials) on performance and removing these factors did not substantially reduce AIC (change in AIC 0.7), so the simplified model was selected. The simplified model (AIC = 270.5) revealed a decreasing tendency to manipulate *A* across activations ($\beta_{activation} = -0.41$, $SE = 0.17$, $p = .01$; Figure 2) and the intercept-only model provided clear evidence that toddlers intervened on the action immediately preceding the effect: they were significantly more likely to manipulate *B* than *A* ($\beta_{intercept} = -1.33$, $SE = 0.24$, $p < .001$).

Comparison of Performance in Experiments 1 and 3

Experiments 1 and 3 used an identical puzzle-box setup (a single box with actions *A* and *B* equidistant from the sticker dispenser; Figure 1A); the only difference was the temporal order of the events toddlers observed (*A–E–B* vs. *A–B–E*). If performance differs significantly between these experiments, then this provides good evidence that toddlers are using temporal information to make causal inferences and guide their behavior, rather than, for example, acting on the basis of a primacy effect.

Mixed effects generalized linear regression with experiment (1 or 3) as a fixed factor and participant ID as a random factor showed a significant effect of experiment on the likelihood of toddlers touching *A* versus *B* first across their five puzzle-box activations. Specifi-cally, toddlers were significantly less likely to intervene on *A* in Experiment 3 than Experiment 1 ($\beta_{experiment} = -3.45$, $SE = 0.46$, $p < .001$; Figure 2). This difference between experiments was also evident when focussing on toddlers' very first manipulation of the puzzle box—they were significantly less likely to intervene on *A* first in Experiment 3 versus Experiment 1, chi-square test for association: $\chi^2(1) = 9.68$, $p = .002$; Figure 2.

Discussion

In contrast to Experiments 1 and 2, in Experiment 3 toddlers were more likely to manipulate *B* than *A*. This was despite the fact that the causal necessity of *A* was ambiguous given that both actions preceded the sticker being dispensed (*A–B–E*); thus, acting on *A* was not incorrect with respect to temporal priority information. If toddlers had been acting based on a primacy effect in Experiments 1 and 2, we would have once again expected them to preferentially intervene on *A* when they saw a demonstration of *A–B–E* in Experiment 3. Instead, our findings suggest that toddlers use temporal order information to guide their causal inferences and behavioral interventions.

Although toddlers were more likely to manipulate *B* rather than *A* first on their very first intervention in Experiment 3, their behavior on this initial interaction with the puzzle box did not differ significantly from chance. Based on temporal order information alone, *A* and *B* were equally likely to be causally necessary, and it was also feasible that both actions were required, that is, that a two-action sequence of *A* followed by *B* was causally necessary. Given that the two actions were both spatially contiguous with the sticker dispenser and equidistant from it, temporal contiguity between action and outcome was the only cue pointing to *B* as the more likely cause. However, although the relative timings of events provide additional information over and above that of temporal order, adults primarily rely on temporal order when making causal structure judgments, with little sensitivity to the exact timing of events (Bramley et al., 2018).

Crucially, though, toddlers' behavior differed significantly between Experiments 1 and 3—even according to our first intervention measure. These two experiments used the same puzzle-box setup; all that differed was the sequence of events that toddlers observed. This demonstrates a sensitivity to the temporal order of events in causal reasoning, though further work is needed to examine how temporal contiguity factors into toddlers' causal inferences.

General Discussion

We investigated whether toddlers' actions reflect the principle that causes must precede their effects—a grasp of which enables us to distinguish cause from effect, explain events, produce desired outcomes, and is a prerequisite for more sophisticated causal inferences. One- and 2-year-olds who observed *A–E–B* demonstrations preferentially intervened on *A*, regardless of whether *A* was spatially contiguous with the sticker dispenser (Experiment 1) or on a separate, spatially discontiguous box (Experiment 2). In both experiments, toddlers were more likely to intervene on *A* on their very first manipulation, as well as across five activations of the puzzle box. The most common first two manipulations in both of these experiments were *A–A*, ruling out that toddlers were "overimitating" the sequence of actions they had seen demonstrated, without understanding the puzzle box's causal structure.

Experiment 3 ruled out that toddlers acted on the basis of a primacy effect in Experiments 1–2: after observing an *A–B–E* demonstration, 1- and 2-year-olds preferentially intervened on *B*, despite the fact that *A* was still the first action they saw the demonstrator perform. Although toddlers' choice of action did not differ significantly from chance the very first time they acted on the puzzle box in Experiment 3 (though *B* was chosen more often than *A*), both actions were temporally prior to the outcome, making them both plausible causes according to temporal order information. Furthermore, toddlers were significantly more likely to act on *B* first in Experiment 3 than in Experiment 1, which would not be expected if they were acting on the basis of a primacy effect (which would predict that they would primarily act on *A* in both cases). These findings contribute to our understanding of the developmental emergence of children's ability to reason about causal systems; specifically in relation to how temporal order information constrains causal inferences from the second year of life.

Our first intervention data support the idea that toddlers as a group grasp that causes must precede their effects, as most individuals in Experiments 1 and 2 intervened on the action temporally prior to the effect (*A*) the very first time they acted on the puzzle box, before receiving any reinforcement. However, there was also evidence of learning—toddlers became even more likely to intervene on the causal action across five activations. It is possible that some toddlers learned the causal relation based on an action–reward association; however, their rapid ability to learn this within five trials could also suggest a form of reinforcement learning supported by underlying causal learning mechanisms, instead of an entirely learned action–reward association. While we cannot rule out that some individuals succeeded overall via reinforcement-based learning of the causal relationship, we believe that the high rate of correct initial interventions suggests that, at the very least, some (implicit) knowledge of the relevance of the temporal order of events supports toddlers' behavior in this task, prior to them gaining experience with the puzzle box and forming action–reward associations (see Espinosa et al., 2022 for a similar argument in relation to physical reasoning in dogs).

Across experiments, we found no evidence for developmental change in performance over the age range tested. These results suggest that 1- and 2-year-olds grasp that causes must precede their effects, can use this knowledge to intervene appropriately, and adhere to it even when it is pitted against salient spatial cues to causality, which this age group is known to be sensitive to (Tecwyn et al., 2020). We conjecture that the poorer performance of 3-year-olds compared with older children in some previous *A–E–B* paradigms (Rankin & McCormack, 2013; Shultz & Mendelson, 1975) was likely due to the response measure used—specifically, having to make an explicit causal judgment. This likely taxed the verbal comprehension skills of the youngest children and increased the cognitive load of the task. In contrast, measuring toddlers' behavioral interventions suggests that a grasp of the causal temporal priority principle is already present in the second year of life.

We found no evidence for age difference within the age range tested for any of our experiments. If, as has been argued by some (e.g., Gopnik et al., 2004) children have some innate assumptions about causal structure, then we might expect basic aspects of causality like the temporal priority principle to be present even in early infancy. Given the success of toddlers in our task, future work could extend the age range below 12 months, and/or collect a larger sample of children at the bottom end of our age range, to determine whether a firm grasp of temporal priority is already present by 12 months. Establishing just how early in development an appreciation of temporal priority is present would shed light on the extent to which this knowledge is innate versus dependent on environmental input, either via observation of others' interactions with causal systems, or infants experiencing the causal effects of their own actions. Whether infants have intuitions about temporal priority could be investigated by adapting the paradigm from the present study so that expectations about causal temporal directionality could be measured based on looking behaviors.

We believe that our results provide a compelling case that 1- and 2-year-olds grasp that causes must precede their effects, and the patterns of behavior observed were largely consistent with those seen in earlier *A–E–B* paradigm studies with older children. However, there is another potential explanation for our findings (which is not necessarily incompatible with our interpretation and could also explain the results of previous *A–E–B* studies with older children) that warrants discussion. It is possible that toddlers ignored events that occurred after the effect (i.e., *B* in Experiments 1 and 2)—either because these were not encoded as causally relevant ("default encoding process," McCormack & Hoerl, 2007), or because toddlers failed to attend to *B* as their attention was captured by the salience of the sticker being dispensed from the puzzle box. On this account, toddlers primarily acted on *A* in Experiments 1 and 2 not because they grasped that *A* must be causal because *A* preceded *E* whereas *B* followed *E*, but because they did not encode *B*. Critically, this account does not require any explicit understanding about the role of temporal order in determining causal structure.

It is important to note that these two accounts (failing to encode the action after the effect vs. explicit reasoning about temporal order) are not necessarily incompatible with one another and would be challenging to tease apart behaviorally. It is feasible that young children may begin with simpler heuristics (e.g., "the action immediately preceding the effect is causal") which reflect implicit knowledge of the causal relevance of the temporal order of events, and over time, with the accumulation of evidence and exposure to/understanding of causal language and explicit teaching (which will vary among 1- to 3-year-olds) they come to explicitly understand that "event *A* preceded effect *E* whereas event *B* followed it, therefore *A* must be causal and *B* cannot be." Along similar lines, it has recently been suggested that children below the age of around 4 years are incapable of genuine temporal reasoning, instead relying on a simpler "temporal updating" system (Hoerl & McCormack, 2019).

Regarding the possibility that toddlers may have failed to attend to events occurring after the effect: we did code toddlers' attention to the demonstrations live, and toddlers only progressed to the response phase of the experiment once they appeared to have actively attended to two full demonstrations, so their observable attention was equivalent across all parts of the demonstration (*A*, *E*, and *B*). Future work could use eye-tracking to more definitively ascertain whether children attend to all events in the sequence (including those that occur after the attention-grabbing effect) and whether/how attentional allocation to events in a causal sequence may change across development.

If it was firmly established that toddlers do attend to all events in the demonstration, then a suggestion put forward by Rankin and McCormack (2013) for disentangling a failure to encode *B* as causally relevant from explicit reasoning about temporal order would be to occlude the event sequence, then tell children about order in which events occurred. If children do not see events unfold in front of them but still preferentially intervene on *A*, then this would rule out that they simply fail to encode *B* because it occurs after effect *E*. However, although this adaptation of the task may be appropriate for 3- to 4-year-olds (the age group tested by Rankin & McCormack, 2013), the increased language and cognitive demands relative to the task used in the present study may impede toddlers' performance, potentially making results difficult to interpret, and it is not clear that a described sequence of events would be encoded in the same way as an observed one.

Although most toddlers acted in accordance with temporal priority the majority of the time, there was some (nonage-related) individual variation in behavior, with some toddlers interacting with action *B* which came after effect *E* in Experiments 1 and 2. While the present study cannot distinguish between alternate explanations for these individual differences, it is of interest to consider what might underpin them. One possibility, which fits with the "simple heuristic" account set out above, is that children who did not act in accordance with this principle all of the time might lack a full appreciation of why *A* is a cause rather than *B* (Rankin & McCormack, 2013). Instead, they may rely on simpler heuristics that do not require consideration of the entire sequence of events (McCormack & Hoerl, 2007), and this lack of explicit knowledge may result in errors.

Another possibility for why a minority of toddlers might have intervened on action *B* in Experiments 1 and 2—perhaps particularly those who did this after first acting on *A*—is that they were exploring or testing causal hypotheses—for instance, investigating whether *B* might have an additional, previously unobserved, effect. There is some evidence that toddlers prioritize exploration and causal learning over gaining a reward more than older children and adults, which may facilitate broad learning and enable them to understand more about the world (Gopnik, 2020). This broad exploration can lead them to the discovery of unexpected causal rules (Walker & Gopnik, 2014) and higher-order generalizations (Sim & Xu, 2017). Thus, what might be viewed as a lack of concrete understanding, inefficient behavior, or heuristics gone wrong (in our case, acting on *B* in Experiments 1 and 2 despite having observed it following the outcome) might in fact be a result of exploration or hypothesis testing, which has the potential to improve a child's knowledge of their environment (Sumner et al., 2019).

Future research could use the same puzzle-box paradigm to investigate how toddlers (and older children) weight and integrate other cues to causality. Experiment 3 suggests that temporal contiguity may factor into toddlers' causal inferences, but further work is required to establish how much the closeness of two events in time constrains young children's causal inferences. Adults tend only to make use of the relative timing of events when temporal order is not a useful cue (Bramley et al., 2018), and previous research with 4- to 7-year-olds has resulted in mixed conclusions regarding how heavily temporal contiguity is weighted. Some evidence suggests children base their causal judgments on temporal contiguity over and above their knowledge of causal mechanisms (Schlottmann, 1999), whereas other findings indicate an ability to accept a temporal delay between cause and effect when a physical rationale is provided (Mendelson & Shultz, 1976). To assess toddlers' sensitivity to temporal contiguity, this causal cue could be pitted against temporal priority by presenting an *A–Delay–E–B* sequence, so that *A* precedes the effect, but is less temporally contiguous with it.

In conclusion, by measuring toddlers' behavioral interventions in an *A–E–B* paradigm we have shown that 1- and 2-year-olds act in accordance with the temporal priority principle. This suggests that a grasp of this fundamental cue to causality emerges early in development, enabling young children to understand, predict, and intervene on the world around them, as well as providing a foundation for the development of more sophisticated causal inferences.

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

Barker, K. B., & Povinelli, D. J. (2019). Anthropomorphomania and the rise of the animal mind: A conversation. *Journal of Folklore Research*, *56*(2–3), 71–90. https://doi.org/10.2979/jfolkrese.56.2_3.05

Bonawitz, E. B., Ferranti, D., Saxe, R., Gopnik, A., Meltzoff, A. N., Woodward, J., & Schulz, L. E. (2010). Just do it? Investigating the gap between prediction and action in toddlers' causal inferences. *Cognition*, *115*(1), 104–117. https://doi.org/10.1016/j.cognition.2009.12.001

Bramley, N. R., Gerstenberg, T., Mayrhofer, R., & Lagnado, D. A. (2018). Time in causal structure learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1880–1910. https://doi.org/10.1037/xlm0000548

Brown, A. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, *14*(1), 107–133. https://doi.org/10.1016/0364-0213(90)90028-U

Buchsbaum, D., Seiver, E., Bridgers, S., & Gopnik, A. (2012). Learning about causes from people, and about people from causes: Statistical inference and social causal reasoning. In F. Xu & T. Kushnir (Eds.), *Rational constructivism* (pp. 125–160). Academic Press.

Bullock, M., & Gelman, R. (1979). Preschool children's assumptions about cause and effect: Temporal ordering. *Child Development*, *50*(1), 89–96. https://doi.org/10.2307/1129045

Burns, P., & McCormack, T. (2009). Temporal information and children's and adults' causal inferences. *Thinking & Reasoning*, *15*(2), 167–196. https://doi.org/10.1080/13546780902743609

Cheke, L. G., Loissel, E., Clayton, N. S., & Burne, T. (2012). How do children solve Aesop's Fable?. *PLoS ONE*, *7*(7), e40574. https://doi.org/10.1371/journal.pone.0040574

Chudek, M., Baron, A. S., & Birch, S. (2016). Unselective overimitators: The evolutionary implications of children's indiscriminate copying of successful and prestigious models. *Child Development*, *87*(3), 782–794. https://doi.org/10.1111/cdev.12529

Espinosa, J., Tecwyn, E. C., & Buchsbaum, D. (2022). Searching high and low: Domestic dogs' understanding of solidity. *Animal Cognition*, *25*(3), 555–570. https://doi.org/10.1007/s10071-021-01568-3

Goddu, M., & Gopnik, A. (2020). Toddlers and preschoolers use relational concepts to solve problems. *Poster presented at the Annual Meeting of the Cognitive Science Society*.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1803), Article 20190502. https://doi.org/10.1098/rstb.2019.0502

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*(1), 3–32. https://doi.org/10.1037/0033-295X.111.1.3

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

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. Hirschfield & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). Cambridge University Press.

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

Hoehl, S., Keupp, S., Schleihauf, H., McGuigan, N., Buttelmann, D., & Whiten, A. (2019). 'Over-imitation': A review and appraisal of a decade of research. *Developmental Review*, *51*, 90–108. https://doi.org/10.1016/j.dr.2018.12.002

Hoerl, C., & McCormack, T. (2019). Thinking in and about time: A dual systems perspective on temporal cognition. *Behavioral and Brain Sciences*, *42*, Article e244. https://doi.org/10.1017/S0140525X18002157

Hume, D. (1739). *A treatise of human nature*. Oxford University Press.

Keupp, S., Behne, T., & Rakoczy, H. (2018). The rationality of (over)imitation. *Perspectives on Psychological Science*, *13*(6), 678–687. https://doi.org/10.1177/1745691618794921

Kominsky, J. F., & Scholl, B. J. (2020). Retinotopic adaptation reveals distinct categories of causal perception. *Cognition*, *203*, Article 104339. https://doi.org/10.1016/j.cognition.2020.104339

Kuhn, D., & Phelps, H. (1976). The development of children's comprehension of causal direction. *Child Development*, *47*(1), 248–251. https://doi.org/10.2307/1128307

Kun, A. (1978). Evidence for preschoolers' understanding of causal direction in extended causal sequences. *Child Development*, *49*(1), 218–222. https://doi.org/10.2307/1128612

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 451–460. https://doi.org/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), 173–185. https://doi.org/10.1111/j.1467-8624.2011.01691.x

Leslie, A. M., & Keeble, S. (1987). Do 6-month-old infants perceive causality? *Cognition*, *25*(3), 265–288. https://doi.org/10.1016/S0010-0277(87)80006-9

McCormack, T., & Hoerl, C. (2007). Young children's reasoning about the order of past events. *Journal of Experimental Child Psychology*, *98*(3), 168–183. https://doi.org/10.1016/j.jecp.2007.06.001

McGuigan, N., & Whiten, A. (2009). Emulation and "overemulation" in the social learning of causally opaque versus causally transparent tool use by 23- and 30-month-olds. *Journal of Experimental Child Psychology*, *104*(4), 367–381. https://doi.org/10.1016/j.jecp.2009.07.001

Meltzoff, A. N., Waismeyer, A., & Gopnik, A. (2012). Learning about causes from people: Observational causal learning in 24-month-old infants. *Developmental Psychology*, *48*(5), 1215–1228. https://doi.org/10.1037/a0027440

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

Muentener, P., & Bonawitz, E. (2017). Development of causal reasoning. In M. Waldmann (Ed.), *Oxford handbook of causal reasoning* (pp. 677–698). Oxford University Press.

Nielsen, M., Kapitány, R., & Elkins, R. (2015). The perpetuation of ritualistic actions as revealed by young children's transmission of normative behavior. *Evolution and Human Behavior*, *36*(3), 191–198. https://doi.org/10.1016/j.evolhumbehav.2014.11.002

Rankin, M. L., & McCormack, T. (2013). The temporal priority principle: At what age does this develop? *Frontiers in Psychology*, *4*, Article 178. https://doi.org/10.3389/fpsyg.2013.00178

Rottman, B. M., Kominsky, J. F., & Keil, F. C. (2014). Children use temporal cues to learn causal directionality. *Cognitive Science*, *38*(3), 489–513. https://doi.org/10.1111/cogs.12070

Schlottmann, A. (1999). Seeing it happen and knowing how it works: How children understand the relation between perceptual causality and underlying mechanism. *Developmental Psychology*, *35*(1), 303–317. https://doi.org/10.1037/0012-1649.35.1.303

Seed, A. M, & Call, Josep. (2014). Space or physics? Children use physical reasoning to solve the trap problem from 2.5 years of age. *Developmental Psychology*, *50*(7), 1951–1962. https://doi.org/10.1037/a0036695

Shultz, T. R., & Mendelson, R. (1975). The use of covariation as a principle of causal analysis. *Child Development*, *46*(2), 394–399. https://doi.org/10.2307/1128133

Sim, Z. L., & Xu, F. (2017). Learning higher-order generalizations through free play: Evidence from 2- and 3-year-old children. *Developmental Psychology*, *53*(4), 642–651. https://doi.org/10.1037/dev0000278

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: the development of causal reasoning in toddlers and infants. *Developmental Psychology*, *42*(6), 1103–1115. https://doi.org/10.1037/0012-1649.42.6.1103

Sobel, D. M., & Legare, C. H. (2014). Causal learning in children. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(4), 413–427. https://doi.org/10.1002/wcs.1291

Sophian, C., & Huber, A. (1984). Early developments in children's causal judgments. *Child Development*, *55*(2), 512–526. https://doi.org/10.2307/1129962

Sumner, E., Li, A. X., Perfors, A., Hayes, B., Navarro, D., & Sarnecka, B. W. (2019). *The exploration advantage: Children's instinct to explore allows them to find information that adults miss*. PsyArXiv. https://doi.org/10.31234/osf.io/h437v

Tecwyn, E. C., Mahbub, N., Kazi, N., & Buchsbaum, D. (2021, July 26–29). Can 1- and 2-year-old toddlers learn causal action sequences? *Proceedings of the 43rd Annual Conference of the Cognitive Science Society, Online/Virtual*.

Tecwyn, E. C., Seed, A. M., & Buchsbaum, D. (2020). Sensitivity to ostension is not sufficient for pedagogical reasoning by toddlers. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society*.

Walker, C. M., Bridgers, S., & Gopnik, A. (2016). The early emergence and puzzling decline of relational reasoning: Effects of knowledge and search on inferring abstract concepts. *Cognition*, *156*, 30–40. https://doi.org/10.1016/j.cognition.2016.07.008

Walker, C. M., & Gopnik, A. (2014). Toddlers infer higher-order relational principles in causal learning. *Psychological Science*, *25*(1), 161–169. https://doi.org/10.1177/0956797613502983

Willatts, P., & Rosie, K. (1989, April 27–30). *Planning by 12-month-old infants*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Kansas City, MO.

Received August 19, 2022
Revision received February 22, 2023
Accepted February 22, 2023 ■