Journal of Experimental Child Psychology 162 (2017) 58–71

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Learning to make things happen: Infants' observational learning of social and physical causal events

Anna Waismeyer [*], Andrew N. Meltzoff

Institute for Learning & Brain Sciences, University of Washington, Seattle, WA 98195, USA

A R T I C L E I N F O

Article history:
Received 6 October 2016
Revised 20 April 2017

Keywords:
Causal learning
Statistical learning
Imitation
Social cognition
Social learning
Nonverbal communication

A B S T R A C T

Infants learn about cause and effect through hands-on experience; however, they also can learn about causality simply from observation. Such *observational causal learning* is a central mechanism by which infants learn from and about other people. Across three experiments, we tested infants' observational causal learning of both social and physical causal events. Experiment 1 assessed infants' learning of a physical event in the absence of visible spatial contact between the causes and effects. Experiment 2 developed a novel paradigm to assess whether infants could learn about a social causal event from third-party observation of a social interaction between two people. Experiment 3 compared learning of physical and social events when the outcomes occurred probabilistically (happening some, but not all, of the time). Infants demonstrated significant learning in all three experiments, although learning about probabilistic cause–effect relations was most difficult. These findings about infant observational causal learning have implications for children's rapid nonverbal learning about people, things, and their causal relations.

---

[*] Corresponding author.
E-mail address: annawais@u.washington.edu (A. Waismeyer).

http://dx.doi.org/10.1016/j.jecp.2017.04.018
0022-0965/

Introduction

Learning about causality is fundamental to infant and child development. Piaget (1954) studied how infants learn about cause and effect based on their own manipulation of objects, theorizing that self-generated action is the foundation of causal learning. Michotte (1962) championed the idea that direct perception could specify cause and effect when the stimulus parameters were arranged in a particular fashion (e.g., collision and launching events). Modern experiments on causal learning during infancy and early childhood have extended this work to cover a wide range of topics, including but not limited to (a) causal Bayes net accounts of causal learning and knowledge (e.g., Gopnik et al., 2004; Gopnik & Schulz, 2007), (b) infants' parsing of Michottean events (Leslie & Keeble, 1987), (c) infants' expectations about basic physical causal relations such as support and containment (Baillargeon et al., 2012), and (d) the role of language in binding causal event sequences (Bonawitz et al., 2010; Hickling & Wellman, 2001).

Here we focused on a variant of causal learning during infancy termed *observational causal learning* (Meltzoff, Waismeyer, & Gopnik, 2012). Observational causal learning is characterized by children learning novel cause–effect relations by observing others' causal acts and then designing their own intervention to bring about the same outcomes (acting to bring about an effect is commonly referred to as an "intervention" in philosophical and causal learning literatures). For example, children may watch as their parents struggle to turn on the television using multiple remote controls; one works to turn on the television and the other does not (e.g., Barr & Hayne, 2003). The *same* act of button pushing is causally effective when applied to one remote control but not when applied to another. Children need not engage in hands-on trial-and-error learning (Piaget, 1954), the television and remote control are not spatially connected (Michotte, 1962), and there need not be any verbal description binding the cause and the effect. Children readily learn how to make things happen by watching how others make them happen.

Early in development, infants imitate simple behaviors they see (e.g., Barr, Dowden, & Hayne, 1996; Elsner & Aschersleben, 2003; Meltzoff, 1988, 2007; Paulus, 2014; Rovee-Collier & Barr, 2010). However, based on this literature alone, it is unclear how much weight infants give to duplicating the precise actions they see versus the outcomes achieved or the degree to which they observationally learn to use a specific object to cause an outcome at a distance. The literature on over-imitation and emulation indicates that toddlers and young children sometimes reproduce actions for their own sake and reenact the goals of a demonstrator without adopting the specific means/actions that adults used (e.g., Buchsbaum, Gopnik, Griffiths, & Shafto, 2011; Carpenter & Call, 2002; Horner & Whiten, 2005; Paulus, Hunnius, Vissers, & Bekkering, 2011; Williamson, Meltzoff, & Markman, 2008). In most of the cases in this literature, the child's manipulation of an object causes an outcome in that same object, not on a different remote object located at a distance.

Meltzoff et al. (2012) showed that 2- to 4-year-olds can learn novel cause–effect relations through observation alone, without prior hands-on experience with the objects, in the absence of linguistic support binding the cause and effect and even when the outcome is spatially displaced from the causal action. The participants observed an adult performing the same action on two different objects. The children selectively chose to use the object that, when acted on, had produced a distal effect. Moreover, the study showed that infants did not behave in this way when the cause and effect were temporally reversed such that the events were simply associated in time but not causally related. This finding and other related ones in the causal learning literature (Bonawitz et al., 2010; Waismeyer, Meltzoff, & Gopnik, 2015) have expanded our understanding of how young children learn cause–effect relations through observation alone, but this work has focused on *physical outcomes* (e.g., the activation of an inanimate object motion or a light turning on). An important remaining question is whether, and if so to what degree, infants might be able to learn about *social* causal relations from observation alone. The current study extended the scope of observational causal learning by testing whether infants can learn novel social causal acts simply by being bystanders or third-party observers who "eavesdrop" on the social interactions between two other people.

Prior work using "eavesdropping" paradigms has shown that infants can both learn new words (Akhtar, Jipson, & Callanan, 2001) and learn about an adult's emotional proclivities (Repacholi &

Meltzoff, 2007) from observing social–communicative interactions between other people. In one study, 15-month-old infants learned about a person's emotional disposition from observation of third-party interactions (e.g., Repacholi, Meltzoff, Toub, & Ruba, 2016). The infants predicted an adult's future emotional reaction after observing the statistical pattern of emotional responses exhibited by that person. This in turn influenced whether the infants would imitate certain actions in the presence of the adult. In other work, infants were able to learn words, such as object labels, when overhearing those words being used in a third-party social interaction (e.g., Akhtar, 2005; Hauf, 2009).

In the current experiments, we had infants eavesdrop on a social interaction between two people, but we explored a different question—whether infants learn a causally instrumental act from eavesdropping on other people's social interactions. Do infants learn how to bring about physical and social outcomes from watching the statistical pattern of evidence demonstrated by others and then act on this information in their own subsequent behavior to bring about the same outcome?

Overview for the current experiments

To test infants' observational causal learning of social events, we modified the causal learning task used by Waismeyer et al. (2015). In the original experiment, infants observed a causal display in which placing an object on a box caused a spatially remote machine to dispense marbles. This is similar to inserting one's parking ticket into an exit machine to cause a remote barrier to be lifted. The causal relation between the ticket insertion and the barrier movement is at a distance, but physical contact between the ticket and the machine triggers the causal chain of events. In Waismeyer et al., physical contact between the object and the box was necessary for the spatially remote marble dispenser to dispense the marbles.

We modified the original task because children have a strong bias toward contact causality (Bullock, Gelman, & Baillargeon, 1982; Kushnir & Gopnik, 2007; Sobel & Buchanan, 2009). It is possible that the initial physical contact in Waismeyer et al. (2015) was a necessary component for the infants to succeed in that test. In Experiment 1, we tested infants' learning of causality at a distance by presenting them with a novel causal display that did not involve spatial contact. It appeared as if a machine dispensed a marble when the experimenter shook one of two freestanding test objects located at a distance from the marble dispenser. Shaking one of the objects caused a marble to dispense, but performing the same shaking action with the other object did not (counterbalanced). The key question was whether infants would be more likely to reproduce the experimenter's act on the causally effective object than on the causally ineffective object. This procedure also allowed us to test social causality in the second experiment.

In Experiment 2, we used the causal displays from Experiment 1 except that we replaced the marble dispenser with a *social* outcome—that of an adult dispensing the marble instead of the machine doing so. The infants were third-party observers who watched two adults play a novel social game. When one adult shook a particular wooden object, a second adult produced a marble. The experimental question was what infants would do when given their own chance to get the highly desirable marble. Would they perform the correct intervention and selectively shake the effective object?

In Experiment 3, we tested observational causal learning of *probabilistic* causal relations. Many causal relations in the real world occur probabilistically (occurring some, but not all, of the time). This may be the case for causal relations between inanimate objects (a wire or connection may be broken), but it is especially true of causal relations between social agents. For example, one may verbally request that someone open the door, and that other person may do so one day but not the next day. To examine the question of probabilistic observational causal learning, we presented two independent groups of infants with the same causal displays as in Experiments 1 and 2 except that the potential causes were now probabilistically effective. The question in Experiment 3 was how well infants would perform in an observational causal learning paradigm when the cause-and-effect relation was a social event and the result occurred only probabilistically and not deterministically.

Experiment 1: Observational learning of physical causality at a distance

In the first experiment, participants watched an adult experimenter act on two objects that served as potential causes of a desirable distal event. Shaking one object (*effective object*) always caused the desirable event, whereas shaking the other object (*ineffective object*) never did. The distal event was a marble dispensing from a machine located 30 cm away (Fig. 1A). After observing this causal chain, participants were presented with the two objects side by side. The key question was whether they chose to take action (the intervention) on the effective object to cause the desirable outcome (marble dispensing). We hypothesized that 24-month-olds would intervene on the causally effective object more often than expected by chance (two objects, counterbalanced, chance = .50).

Method

Participants

The participants were 32 24-month-olds within ±14 days of their birthdays (*M* = 24.09 months, *SD* = 6.80 days). Equal numbers of boys and girls were tested. An additional 3 participants began testing but were excluded due to experimenter error ($n = 2$) or for refusal to touch the test objects ($n = 1$). Participants were recruited by telephone from the university's computerized participant pool. Preestablished criteria for admission into the study were that infants be full-term, have normal birth weight, and have no known developmental concerns. The sample was primarily middle to upper-middle class with 77% White, 0% Asian, and 23% "other" (10% were of Hispanic ethnicity) according to parental report.

Stimuli

Two sets of wooden objects were used and differed from each other in both shape and color. A green egg (6×4 cm) and yellow square (7×7 cm) were used in the first test trial, and a red cylinder (7.5×3.25 cm) and blue hemisphere (4.75×9.75 cm) were used in the second test trial. The objects were at either side of a rectangular platform (69×20 cm). For the response period, the platform supporting the objects was slid toward the infant. The experimenter was seated directly across the table from the infant, and a marble-dispensing machine was located to the infant's right, near the far corner of the table (Fig. 1A).

Procedure

Before testing, the families went to a waiting room and completed consent forms, and the child was acclimated to the adult who would conduct the study. After the child appeared to be comfortable, the family was escorted to the testing room. During test trials, the infant was seated on the parent's lap at a black table (72×120 cm) on which the sliding platform and test objects were displayed. All responses were video-recorded.

The infant was presented with two test trials. On each test trial, the infant first observed a series of causal events (the *stimulus presentation period*). During this period, the objects were on the experimenter's side of the table and out of reach of the child (~15 cm from the table edge on the experimenter's side). After the stimulus presentation period, the platform containing the test objects was slid across the table to bring the objects within the infant's reach. The infant was given a 30-s period to produce a response (*response period*). Next the experimenter put away the two test objects from the first test trial and placed the second set of test objects on the platform. The second test trial then proceeded in the same way as the first.

Stimulus presentation period. The infant observed as an experimenter, the *initiator*, shook the two test objects. The initiator first shook one object three times in a row and then shook the second object three times in a row. The initiator then repeated this sequence (see Table 1). The marble dispensation occurred deterministically. Each time the initiator shook the effective object (object and side counterbalanced), a marble immediately dispensed from the marble machine 30 cm away (six of six times). When the initiator shook the ineffective object, the marble never dispensed (zero of six times). The temporal contingencies were such that adults observing this series of events guessed that the effective

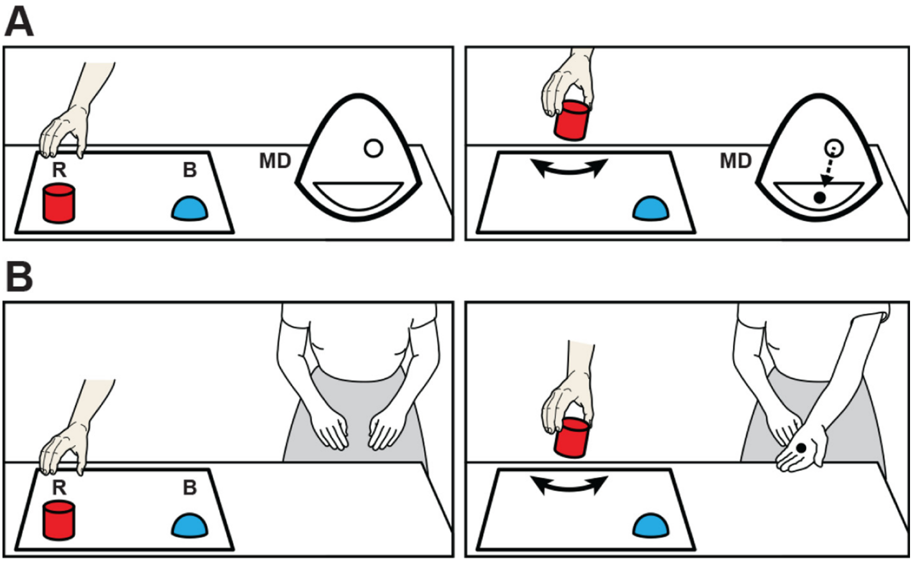

Fig. 1. (A) Experiment 1 used the physical event shown in the top row. (B) Experiment 2 used the social event shown in the bottom row. The red cylinder (R) and the blue hemisphere (B) served as potential causes in a chain of events (counterbalanced). MD indicates the marble dispenser. The elements in the figure are not drawn to scale; see text for measurements. (The web version of this article is in color.)

Table 1
Activation patterns shown to 24-month-olds in Experiments 1 and 2.

Activation pattern	Object 1 →	Object 2 →	Object 1 →	Object 2 →	Object 1: # of effective actions	Object 2: # of effective actions
1	1–1–1	0–0–0	1–1–1	0–0–0	6 of 6	0 of 6
2	0–0–0	1–1–1	0–0–0	1–1–1	0 of 6	6 of 6

Note. “1” indicates an effective event, and “0” indicates an ineffective event. There are two different activation patterns (rows). The time course for each activation pattern proceeds from left to right, as shown by the arrows.

object contained a remote control that, when shaken, caused the marble dispenser to dispense a marble. The objects were solid and did not rattle; it was the adult’s action with the object that apparently triggered the machine. Once the marble was dispensed, the infant was allowed to retrieve it and place it in a marble run with opaque end points. In this manner, the marble was out of sight before the start of the next event.

The procedure during the stimulus presentation period was strictly controlled by a predetermined protocol. The participant was not given any prior experience with the marbles or with the objects to rule out trial-and-error learning. No causal linguistic descriptions were used to avoid linguistic bootstrapping. The initiator used attention-getting language, such as “Let’s watch,” but did not provide a causal narrative of events, such as “The block makes it go” or “I’m going to make it work.” Other pedagogical cues during the use of the effective and ineffective objects were also equated; the initiator looked to the infant an equal number of times when handling both the effective and ineffective objects—once before and once after shaking.

Controls and counterbalancing. The sequence of 12 observed events, 6 with each object, constituted the activation pattern for one test trial. There were two possible activation patterns (shown as the rows in Table 1). The activation patterns were designed to preclude infants’ use of common low-order heuristics to solve the task. For half of the infants the initiator began by shaking the effective object (Activation Pattern 1), and for the other half she began by shaking the ineffective object (Activation Pattern 2). Thus, choosing to interact with the object that the initiator used first (or last) would result in responding at chance. The experiment was also counterbalanced with respect to the sex of the participants and which object was the effective object. On the second test trial, the activation pattern was switched (e.g., from Activation Pattern 1 to 2), as were the test objects and the side on which the effective object was placed.

Response period. Following the stimulus presentation, the initiator presented the infant with the two objects side by side on the platform (Fig. 1A) separated by a distance of 50 cm. If infants are able to use observational causal learning to learn about physical events in the absence of any visible contact, they should selectively choose to interact with the effective object and to shake it. It is noteworthy for theory that the perception of statistical regularities alone is not sufficient to solve the task because the dependent measures require infants to initiate manual acts (e.g., touching, shaking). It is also noteworthy that to succeed on the task, infants could not simply duplicate the initiator's motor behavior because the initiator touched and shook both objects in the same way and did so an equal number of times.

Scoring and dependent measures

Responses were scored from the video records of the study in a random order by a coder who was kept blind to the test conditions. The coder viewed video segments of the response periods only. These segments contained no clues about which object was the effective versus ineffective object, and thus there was no artifactual way of influencing the scoring. One quarter (25%) of the videos were rescored by a second coder who was kept blind in the same way. There were two dependent measures: *touch* and *shake*. There were no intra- or inter-coder disagreements on either measure (Cohen's kappa = 1.0 for both).

Touch score. Touch was operationalized as the infant's finger or palm coming into contact with an object. During the response period of any given trial, the infant was presented with two objects. These objects were presented side by side equidistant from the infant and within the infant's reach. Thus, for each infant, a touch score could be calculated: The numerator was the number of trials in which the infant first touched the effective object, and the denominator was the number of trials in which the infant first touched the effective object plus the number in which he or she first touched the ineffective object. Thus, each participant received a touch score between 0 and 1, representing the proportion of trials in which the initial contact was made with the effective object rather than the ineffective one.

Shake score. Shake was operationalized as a rapid manual movement that produced a change in the object's trajectory. The coder reviewed the video segments of the response period in real time and sometimes in slow motion to score this behavior. The shake score was calculated similarly to the touch score. For each infant, the numerator was the number of trials in which the infant first shook the effective object, and the denominator was the number of trials in which the infant first shook the effective object plus the number in which he or she first shook the ineffective object. There were some infants who did not shake an object on either trial. These infants could not receive a shake score and were not included in the subsequent shake score analyses (the resulting sample sizes for all analyses are provided in Results and Discussion).

The analyses were conducted using t tests; one-tailed tests were used in the experiments, unless otherwise noted, because across six prior related experiments (Meltzoff et al., 2012; Waismeyer et al., 2015) infants' performance was never significantly directed toward an ineffective object. Moreover, the design was also counterbalanced such that infants' choices based on object preferences (e.g., blue object), side preferences (e.g., right side), and first/last object touched by the experimenter would result in a score of .50. There is no reason to believe that, as a group, infants would systematically prefer the ineffective object. The experimental question is whether they would systematically choose to act on the effective object at greater than chance (.50) levels.

Results and discussion

The video record of the response period was not available for 2 of the 32 participants due to camera failure. These 2 participants were included in the touch score analysis using the live scoring done by the experimenter (shake was not scored live because it sometimes required slow-motion review of the video-recording). The mean number of trials in which infants touched either object was 1.9 ($SD = .25$).

The 24-month-olds' touch scores were significantly greater than .50 ($M = .83$, $SD = .24$), $t(31) = 7.69$, $p < .001$, $d = 2.76$. Fully 27 of 30 participants performed a shaking action, and their mean shake score (reflecting which specific object was shaken) was significantly greater than .50 ($M = .82$, $SD = .31$), $t(26) = 5.20$, $p < .001$, $d = 2.04$. This mean score indicates that infants systematically chose to shake the object that had triggered the marble dispenser. All infants were highly attentive. Their gaze during the stimulus presentation period was directed at the displays 93% of the time ($M = 93\%$, $SD = 3.60$).

Previous studies using related causal learning procedures have demonstrated that infants make anticipatory looks to the causal outcome before it occurs (Bonawitz et al., 2010; Meltzoff et al., 2012), suggesting that an enhancement of the test objects alone is insufficient to account for the response; moreover, we scored *both* that infants touched the effective object *and* that they reproduced the shaking act (the solid object made no rattling sound, so infants were reproducing the specific act seen without sound cues). We believe that the findings support the inference that infants can learn causal relations from observing another person's causal interventions even when the causal relations do not depend on spatial contact or linguistic descriptions binding the causal event. In Experiment 2, we examined infants' observational causal learning of social events.

Experiment 2: Observational learning of social causality

In the second experiment, participants observed the same causal displays as in Experiment 1 with one critical exception: Instead of the desirable event being a machine dispensing a marble, a person (the *responder*) dispensed the marble. The cause-and-effect relation was between the initiator's shaking of the test objects (potential causes) and a change in the responder's behavior (the effect). It appeared as if the initiator and the responder were interacting with one another using a novel "communicative" gesture (see below).

There were at least two interesting results that could potentially emerge from this novel task. First, infants' observational causal learning might be constrained to learning about *physical* outcomes, and infants' performance might drop to chance level. Second, once people are involved as stimuli, a set of factors beyond strict physical causal reasoning comes into play. Infants are able to make social inferences from observations. These inferences pertain to people's preferences, desires, emotions, and intentions (e.g., Kushnir, Xu, & Wellman, 2010; Ma & Xu, 2011; Meltzoff, 1995; Repacholi et al., 2016). For example, 18-month-olds are able to infer an adult's desires based on the adult's prior disgust responses to food items (Repacholi & Gopnik, 1997). If infants are also able to make socially relevant *cause-and-effect* inferences of the type tested here, they should perform above chance when the initiator's shaking of an object causes a person—instead of a machine—to dispense a marble. Perhaps their performance might even improve.

Method

Participants

The participants were 32 24-month-olds within ±14 days of their birthdays ($M = 23.95$ months, $SD = 6.35$ days). Equal numbers of boys and girls were tested. An additional 5 participants began testing but were excluded for refusal to touch the test objects ($n = 4$) or for ambiguous behavior at test ($n = 1$). All participants were recruited as in Experiment 1. The sample was primarily middle to upper-middle class with 81% White, 3% Asian, and 16% "other" (13% were of Hispanic ethnicity) according to parental report.

Procedure and stimuli

Before testing, the family went to a waiting room and completed consent forms, and the child was introduced to the two female adults who would conduct the study. Prior to testing, the initiator and the responder each took a turn playing with the infant in the waiting room while the other experimenter discussed with the parent either consent documents or the guidelines for parent's behavior during the study. When interacting with each other, the two experimenters interacted naturally, neither emphasizing nor minimizing their social connection. After the child appeared to be comfortable, the family was escorted to the testing room.

The infant participated in the same procedure as in Experiment 1 except that the marble-dispensing machine was replaced by an adult responder (Fig. 1B). The objects, activation patterns, controls, and counterbalancing were identical to those used in Experiment 1. Again, no causal linguistic descriptions were used, pedagogical cues were controlled, and participants were not given any prior experience with the stimuli to rule out learning by trial and error.[1]

Scoring

Participants were scored in the same manner as in Experiment 1. They were scored in a random order by a coder who was kept blind to the test conditions, and 25% of the videos were rescored by a second coder who was also kept blind to the test conditions. Scoring agreement was high for the touch scores (Cohen's kappa = 1.0 for both intra- and inter-coder reliabilities) and shake scores (Cohen's kappa = 1.0 for intra-coder reliability and .87 for inter-coder reliability). As in Experiment 1, sample means closer to 1 for the dependent measures indicate that children chose to act systematically on the effective object more than on the ineffective object.

Results and discussion

Infants' mean touch score was significantly greater than .50 ($M = .64$, $SD = .41$), $t(31) = 1.96$, $p = .03$, $d = 0.57$. Of the 32 participants, 16 performed a shaking act on one of the two objects, and their mean shake score was significantly greater than .50 ($M = .71$, $SD = .44$), $t(16) = 1.95$, $p = .03$, $d = 0.98$. As in Experiment 1, the infants were highly attentive during the stimulus presentation period. Their gaze was directed at the displays 91% of the time ($M = 91\%$, $SD = 4.91$). The mean number of trials in which infants touched either object was 1.6 ($SD = 0.5$), which differed from Experiment 1, $t(62) = 3.78$, $p < .01$, $d = 0.96$, and could be due to the social nature of the task, requiring participants to interact with an adult experimenter to get the marble, or merely to the presence of two adult experimenters rather than a single experimenter (see footnote 1).

These results show that infants' observational causal learning is not limited to learning about causal relations involving physical outcomes. Infants are able to use their third-party observations of a social interaction between two people to make decisions about their own future acts in the same social situation. Observational causal learning may contribute both to infants' developing knowledge of physical–instrumental causality and to their blossoming social–communicative understanding.

One unexplored question, however, is whether observational causal learning of social events can incorporate *probabilistic* or imperfect evidence. In the real world, people do not always act reliably and deterministically (and neither do objects). In Experiment 3, we investigated this question by assessing infants' observational causal learning of both physical and social events from *probabilistic* evidence.

Experiment 3: Observational learning of probabilistic social and physical causality

Everyday social interactions often include an element of uncertainty. For example, when children asks their mother to bring them a toy, she may sometimes say yes and she may sometimes refuse the request; a well-formed linguistic request does not always (deterministically) cause a social effect. Physical causality is also not fully deterministic, especially from infants' point of view, because objects can be broken or simply not handled skillfully enough to work deterministically. In Experiments 1 and

---

[1] We chose this design to equate the number of significant objects in Experiments 1 and 2. There were five significant objects in Experiment 1 (the adult, two test objects, the marble dispenser, and the marble) and five significant objects in Experiment 2 (two adults, two test objects, and the marble). We could have equated the number of people present in both experiments by adding a second adult (stooge) in Experiment 1; however, that would have resulted in six significant objects, and it is unclear what specific activity this extra person should do. Thus, we swapped out the marble dispenser (physical event) for a person who dispenses marbles (social event). In the future, other manipulations could be tried.

2, the effective and ineffective objects were deterministically causal (100% vs. 0% effective). Can infants succeed on observational causal learning tasks when there are imperfect contingencies?

Prior work has shown that infants can learn many types of relations from probabilistic displays, including both causal events and non-causal patterns (Denison & Xu, 2010; Gweon, Tenenbaum, & Schulz, 2010; Kuhl, 2004; Saffran, Aslin, & Newport, 1996; Téglás et al., 2011; Waismeyer et al., 2015). Although there are no experiments with infants using probabilistic information to test causal learning of social events (i.e., tests in which infants intervene to cause an outcome), there is informative work using *non*-causal tests. This work has shown that infants and toddlers can use probabilistic information to make predictions about people's attitudes such as their preferences and desires. What remains to be tested is whether infants' learning from probabilistic information extends to a social *causal* learning task. Experiment 3 was designed to explore this question.

We presented two groups of infants with the same causal displays as in Experiments 1 and 2 with one further modification. Instead of the potential causes being deterministically effective or ineffective, now each potential cause was probabilistically effective. At test, instead of choosing between a causal object and a non-causal object, infants would now choose between two causally effective objects with differing levels of effectiveness.

Method

Participants

The participants were 64 24-month-olds within ±14 days of their birthdays (M = 24.02 months, SD = 6.26 days). Equal numbers of boys and girls were tested. An additional 8 infants began testing but were excluded due to experimenter error (n = 2), refusal to touch the test objects (n = 2), ambiguous behavior at test (n = 2), or fussiness (n = 2). All participants were recruited as in Experiment 1. The sample was primarily middle to upper-middle class with 83% White, 2% Asian, and 16% "other" (5% were of Hispanic ethnicity) according to parental report.

Procedure and stimuli

Infants were randomly assigned to one of two independent groups: physical causal events or social causal events. The physical events group followed the procedures used in Experiment 1, and the social events group followed the procedures used in Experiment 2. The crucial change was that the two potential causes were probabilistically effective, rather than deterministically effective, during the stimulus presentation period (see below).

Stimulus presentation period. The *initiator* shook the two test objects as in Experiments 1 and 2. However, shaking one object caused a marble to be dispensed four of six times (.67 probability of being effective). This object was termed the *high-probability* object. Shaking the other object caused a marble to be dispensed only two of six times (.33 probability of being effective). This object was termed the *low-probability* object (Table 2). If infants learn probabilistic causal relations by observation of both physical and social events, they should systematically choose to use the high-probability object during the response period in both groups. As in Experiments 1 and 2, participants were not given any prior experience with the stimuli to rule out learning by trial and error, no causal linguistic descriptions were used, and pedagogical cues were rigorously controlled.

Controls and counterbalancing. A sequence of 12 observed events, 6 with each object, constituted the activation pattern for one test trial. There were four possible activation patterns (Table 2). As in Experiments 1 and 2, the activation patterns were designed to preclude the use of common low-order heuristics to solve the task. For half of the participants the high-probability object was used first, and for the other half the low-probability object was used first. Thus, choosing to interact with the object the experimenter used first would result in responding at chance. In addition, for any single participant (reading across a row in Table 2), the pattern for both objects began with the same object outcome, either effective or ineffective at producing a marble, and both ended with the opposite outcome. Thus, infants could not use either the first observed outcome or the last observed outcome to determine which object to interact with.

Both groups (physical events and social events) were counterbalanced with respect to the sex of participants, which object was the high-probability object, and the side on which the high-

Table 2
Activation patterns shown to 24-month-olds in Experiment 3.

Activation pattern	Object 1 →	Object 2 →	Object 1 →	Object 2 →	Object 1: # of effective actions	Object 2: # of effective actions
1	1–1–0	1–0–0	1–1–0	1–0–0	4 of 6	2 of 6
2	1–0–0	1–1–0	1–0–0	1–1–0	2 of 6	4 of 6
3	0–1–1	0–0–1	0–1–1	0–0–1	4 of 6	2 of 6
4	0–0–1	0–1–1	0–0–1	0–1–1	2 of 6	4 of 6

Note. "1" indicates an effective event, and "0" indicates an ineffective event. There are four different activation patterns (rows). The time course for each activation pattern proceeds from left to right, as shown by the arrows.

probability object was located. On the second test trial, the test objects and the side on which the effective object was placed were switched and the activation pattern was changed such that if an infant had first observed a pattern starting with an effective event (e.g., Activation Pattern 1), on the second test trial the infant would observe an activation pattern starting with an ineffective event (e.g., Activation Pattern 4).

Scoring

Participants were scored in the same manner as in Experiments 1 and 2. Responses were scored from video in a random order by a coder who was kept blind to the test conditions, and 25% of the videos were rescored by a second coder who was also kept blind to the test conditions. Scoring agreement was high for the touch scores (Cohen's kappa = 1.0 for both intra- and inter-coder reliabilities) and shake scores (Cohen's kappa = 1.0 for intra-coder reliability and .94 for inter-coder reliability).

Results and discussion

The video record of the response period was not available for one of the participants due to camera failure. Responses for this participant were included in the touch score analysis using the live scoring done by the experimenter. Touch scores were significantly greater than expected by chance, .50 ($M = .62$, $SD = .36$), $t(63) = 2.57$, $p = .006$, $d = 0.65$. Of the 63 participants, 44 performed a shaking action on one of the two objects, and their mean shake score was significantly greater than .50 ($M = .67$, $SD = .39$), $t(43) = 2.91$, $p = .003$, $d = 0.89$. As in Experiments 1 and 2, participants were highly attentive. Infants' gaze during the stimulus presentation period was directed at the displays approximately 93% of the time ($M = 93\%$, $SD = 3.56$). The mean number of trials in which infants touched either object was 1.5 ($SD = 0.5$), which was different from Experiment 1, $t(94) = -4.31$, $p < .01$, $d = 0.89$, but not different from Experiment 2, $t(94) = -0.29$, $p = .77$, $d = 0.06$.

There were no significant differences between the physical and social events groups in either touch or shake measures. It is nonetheless of interest to provide results broken down at this finer grain. Considering just the physical events group alone, mean touch and shake scores both were significantly greater than .50 ($M = .64$, $SD = .34$, $t(31) = 2.33$, $p = .013$, $d = 0.84$ and $M = .67$, $SD = .37$, $t(25) = 2.37$, $p = .013$, $d = 0.95$, respectively. Considering just the social events group alone, the mean touch score was in the predicted direction but not significantly greater than .50 ($M = .59$, $SD = .39$), $t(31) = 1.36$, $p = .09$, $d = 0.49$, and the mean shake score was significantly greater than .50 ($M = .67$, $SD = .42$), $t(17) = 1.68$, $p = .05$, $d = 0.81$. In sum, these findings generally support the conclusion that infants' observational causal learning of physical and social causal relations is able to incorporate probabilistic evidence.

General discussion

The findings show that 24-months-olds can learn novel causal interventions from observing the outcomes of others' actions even without linguistic support and using spatially remote causes and effects, action at a distance. Across three experiments, participants learned which particular object to act on to produce a desirable effect, and they chose to duplicate the adult's specific causal act on that object. These findings advance what we know about the development of social and causal learning in several ways.

First, the results show that observational causal learning in infants goes beyond causal learning by firsthand manipulation of the objects and/or trial and error. In the designs used here, infants simply *observed* the adult's actions. They were not allowed to touch or handle the test objects before the response period, and first choices were scored for dependent measures. The results also go beyond Michottean billiard ball collision events because the designs involved action at a distance and relied on complex patterns of covariation across a sequence of demonstrations (see Tables 1 and 2 for activation patterns). The results depend on infants' statistical learning, perceptual predictions, and expectations, but they also require "something more" beyond that. To succeed on these tasks, infants needed to make a choice and *act;* they needed to choose which object to *shake* to bring about the desirable effect (distal marble dispensing).

It is noteworthy for theory that the participants performed the appropriate causal intervention in the absence of a number of supporting cues to causality. In the current design, participants could not learn the appropriate intervention based on a linguistic causal narrative because the initiator provided no causal linguistic descriptions of the events. The infants also could not use differences in pedagogical cues to determine which object to choose because pedagogical cues such as eye contact and infant-directed speech were equated. Infants' ability to learn in such circumstances suggests that infants can learn about cause and effect from observing social interactions between other people.

There were some differences between infants' performance across the three experiments. Infants' mean touch score in Experiment 2 (social causality) was lower than that in Experiment 1 (physical causality), $t(62) = 2.24$, $p = .03$, $d = 0.70$, two-tailed. Similarly, the number of infants who shook either object (whether the correct or incorrect one) was lower for the social causal events than for the physical ones: Experiment 1 (27 of 30) versus Experiment 2 (16 of 32), $p < .001$, and Experiment 3 physical (26 of 31) versus Experiment 3 social (18 of 32), $p = .03$ (both using two-tailed Fisher's exact tests). In these ways, infants performed differently in the social versus physical causal conditions, and one is immediately drawn to exploring why this might be. The differences might be due to infants being more shy in the presence of two adults than in the presence of one adult or perhaps to infants not knowing what to do to "activate" the marble-dispensing adult. However, it is also possible that something more interesting may be occurring. We offer three speculative suggestions for future research.

One reason for a difference between physical and social causal events may be that infants use the presence of human intervention as a way of determining which causal hypotheses to pay attention to and learn about, and they can put this to better use when observing physical events. When observing covarying events between two people, the presence of human intervention is not a discriminative marker for highlighting which of the several covarying events may be important. Of course, for social events, infants may seek cues to the people's intentions, desires, or preferences (e.g., Meltzoff, 1995; Repacholi et al., 2016; Woodward, Sommerville, Gerson, Henderson, & Buresh, 2009). However, in the current experiments, the initiator's and responder's intentions, desires, and preferences were not indicated through differential emotions, language, and the like because these were controlled by design. Future work could explore whether the addition of such information might improve infants' performance in observational causal learning tasks involving two people (e.g., Repacholi et al., 2016; Sobel & Munro, 2009).

Another option is that infants' understanding of object functions may constrain causal hypotheses when observing physical events (e.g., Casler & Kelemen, 2005, 2007) but might not be applicable when observing social events. For example, in the current experiments, the marble dispenser's function appears to be that of generating marbles; once activated, the machine will presumably dispense another marble. In contrast, people are not limited to one function; if a person produces a marble in response to a signal at one time, the person might not produce a marble the next time. Even if the person displays a consistent pattern of behavior over time, this might not be enough for infants to form a firm prediction for what may happen on the *next* event—the one that they are now initiating. If infants believe that they cannot form a firm prediction for the next individual social event, they might not take action based on their observations. One way to explore this might be to compare infants' generalizations to novel situations. For physical causality, infants might systematically choose the same effective object even when presented with a novel physical outcome—as if the object were a universal remote or magic object. In contrast, this might not occur for social causality because infants might need evidence that the same act would cause a new person to change his or her behavior.

A third possibility is that infants' understanding of social causal mechanisms (e.g., social signals changing the behavior of the person) may be different from their understanding of physical causal mechanisms. When observing social events in which a person's behavior varies—for example, the responder produces a marble when a red object is shaken but not when a blue one is shaken—infants might infer that the human agent (responder) is unreliable. Young children have been shown to treat unreliable people as poor sources of information (e.g., Birch, Vauthier, & Bloom, 2008; Bridgers, Buchsbaum, Seiver, Griffiths, & Gopnik, 2016; Ganea, Koenig, & Millett, 2011; Jaswal & Neely, 2006; Kidd, Palmeri, & Aslin, 2013; Koenig & Harris, 2005). An interpretation that a person is unreliable might lead infants to assimilate less causal information from that person. Indeed, infants draw trait-like inferences about people (Repacholi et al., 2016), which could influence infants' predictions about what the responder is likely to do when the object is next shaken, taking into account the responder's past history.

Finally, an interesting question for the future will be to explore how other aspects of social–cognitive development interact with the phenomena reported here. For example, Meltzoff's (1988, 2007) "observation-only" design for imitation as well as Paulus's (2014) model of imitative learning seek to explain how and why infants take action to bring about effects based on observing others, and developmental neuroscience views on observational learning and imitation are also beginning to emerge (Marshall & Meltzoff, 2014). It is possible that with suitable simplifications of the current design, infants at younger ages than those tested here could succeed on our tasks.[2] It is also possible that there are developmental interactions with executive functions, memory, and statistical learning that could contribute to the emergence of observational causal learning.

Conclusions: Learning to make things happen

The findings across three experiments demonstrate that observational causal learning is a robust mechanism for learning how to make things happen even when the effects are at a distance and not spatially contiguous with the cause. Infants learn both when they see acts cause something to happen in an inanimate object and when they see that acts make a social agent do something. Furthermore, the results demonstrate that infants can learn these causal relations via third-party observation—simply by eavesdropping on social interactions between other people in the absence of specific dyadic teaching directed toward the infants themselves. Infants also learn not only when the cause–effect relations are deterministic but also when the patterns of evidence are probabilistic in nature.

Finally, we have discussed possible differences in young children's causal learning about physical and social events. Taken together, the results suggest that observational causal learning can contribute to the rapid social and physical causal learning during infancy and early childhood development. More speculatively, practice in reenacting the everyday causal procedures of others, which are often probabilistic in producing results, may support young children's developing skills for decision making and acting under uncertainty in a broader set of tasks and contexts.

Acknowledgments

This research was supported by a grant from the National Science Foundation (DRL-1251702). We thank the anonymous reviewers for several helpful suggestions.

---

[2] The current experiments involve causal learning across statistical patterns (Tables 1 and 2), which entails integrating patterns of successes and failures, and thus are somewhat more complex than those typically used in tests of infant imitation (for related studies of imitation integrating patterns of success and failure over time, see work in older children such as Want & Harris, 2001, and Williamson & Meltzoff, 2011).

References

Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science, 8*, 199–209.

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development, 72*, 416–430.

Baillargeon, R., Stavans, M., Wu, D., Gertner, Y., Setoh, P., Kittredge, A. K., & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative account. *Language Learning and Development, 8*, 4–46.

Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development, 19*, 159–170.

Barr, R., & Hayne, H. (2003). It's not what you know, it's who you know: Older siblings facilitate imitation during infancy. *International Journal of Early Years Education, 11*, 7–21.

Birch, S. A., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107*, 1018–1034.

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*, 104–117.

Bridgers, S., Buchsbaum, D., Seiver, E., Griffiths, T., & Gopnik, A. (2016). Children's causal inferences from conflicting testimony and observations. *Developmental Psychology, 52*, 9–18.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition, 120*, 331–340.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York: Academic Press.

Carpenter, M., & Call, J. (2002). The chemistry of social learning. *Developmental Science, 5*, 22–24.

Casler, K., & Kelemen, D. (2005). Young children's rapid learning about artifacts. *Developmental Science, 8*, 472–480.

Casler, K., & Kelemen, D. (2007). Reasoning about artifacts at 24 months: The developing teleo-functional stance. *Cognition, 103*, 120–130.

Denison, S., & Xu, F. (2010). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*, 798–803.

Elsner, B., & Aschersleben, G. (2003). Do I get what you get? Learning about the effects of self-performed and observed actions in infancy. *Consciousness and Cognition, 12*, 732–751.

Ganea, P. A., Koenig, M. A., & Millett, K. G. (2011). Changing your mind about things unseen: Toddlers' sensitivity to prior reliability. *Journal of Experimental Child Psychology, 109*, 445–453.

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*, 3–32.

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

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 9066–9071.

Hauf, P. (2009). The interchange of self-performed actions in infants. In T. Striano & V. Reid (Eds.), *Social cognition: Development, neuroscience, and autism* (pp. 129–143). Oxford, UK: Wiley-Blackwell.

Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology, 37*, 668–683.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition, 8*, 164–181.

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best: Preschoolers use past reliability over age when learning new words. *Psychological Science, 17*, 757–758.

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition, 126*, 109–114.

Koenig, M. A., & Harris, P. L. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*, 1261–1277.

Kuhl, P. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843.

Kushnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43*, 186–196.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*, 1134–1140.

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

Ma, L., & Xu, F. (2011). Young children's use of statistical sampling evidence to infer the subjectivity of preferences. *Cognition, 120*, 403–411.

Marshall, P. J., & Meltzoff, A. N. (2014). Neural mirroring mechanisms and imitation in human infants. *Philosophical Transactions of the Royal Society: Biological Sciences, 369*, 20130620.

Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology, 24*, 470–476.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838–850.

Meltzoff, A. N. (2007). "Like me": A foundation for social cognition. *Developmental Science, 10*, 126–134.

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

Michotte, A. (1962). *Causalité, permanence, et réalité phénoménales: Études de psychologie expérimentale*. Louvain, Belgium: Publications Universitaires.

Paulus, M. (2014). How and why do infants imitate? An ideomotor approach to social and imitative learning in infancy (and beyond). *Psychonomic Bulletin & Review, 21*, 1139–1156.

Paulus, M., Hunnius, S., Vissers, M., & Bekkering, H. (2011). Bridging the gap between the other and me: The functional role of motor resonance and action effects in infants' imitation. *Developmental Science, 14*, 901–910.

Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). New York: Basic Books.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12–21.

Repacholi, B. M., & Meltzoff, A. N. (2007). Emotional eavesdropping: Infants selectively respond to indirect emotional signals. *Child Development, 78*, 503–521.

Repacholi, B. M., Meltzoff, A. N., Toub, T. S., & Ruba, A. L. (2016). Infants' generalizations about other people's emotions: Foundations for trait-like attributions. *Developmental Psychology, 52*, 364–378.

Rovee-Collier, C., & Barr, R. (2010). Infant learning and memory. In J. G. Bremner & T. D. Wachs (Eds.), *Wiley-Blackwell handbook of infant development* (Vol. 1, 2nd ed., pp. 271–294). Oxford, UK: Wiley-Blackwell.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Sobel, D. M., & Buchanan, D. W. (2009). Bridging the gap: Causality-at-a-distance in children's categorization and inferences about internal properties. *Cognitive Development, 24*, 274–283.

Sobel, D. M., & Munro, S. E. (2009). Domain generality and specificity in children's causal inference about ambiguous data. *Developmental Psychology, 45*, 511–524.

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*, 1054–1059.

Waismeyer, A., Meltzoff, A. N., & Gopnik, A. (2015). Causal learning from probabilistic events in 24-month-olds: An action measure. *Developmental Science, 18*, 175–182.

Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development, 72*, 431–443.

Williamson, R. A., & Meltzoff, A. N. (2011). Own and others' prior experiences influence children's imitation of causal acts. *Cognitive Development, 26*, 260–268.

Williamson, R. A., Meltzoff, A. N., & Markman, E. M. (2008). Prior experiences and perceived efficacy influence 3-year-olds' imitation. *Developmental Psychology, 44*, 275–285.

Woodward, A. L., Sommerville, J. A., Gerson, S., Henderson, A. M. E., & Buresh, J. (2009). The emergence of intention attribution in infancy. In B. H. Ross (Ed.). *The psychology of learning and motivation* (Vol. 51, pp. 187–222). San Diego: Elsevier Academic Press.