Cognition 244 (2024) 105712

Contents lists available at ScienceDirect

Cognition

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

Three-year-olds' ability to plan for mutually exclusive future possibilities is limited primarily by their representations of possible plans, not possible events

Esra Nur Turan-Küçük[a,*], Melissa M. Kibbe[b]

[a] Department of Psychological and Brain Sciences, Boston University, Boston, MA, USA
[b] Department of Psychological and Brain Sciences, Center for Systems Neuroscience, Boston University, Boston, MA, USA

ARTICLE INFO

Keywords:
Future-oriented thinking
Planning
Possibility
Modal reasoning
Cognitive development

ABSTRACT

The ability to prepare for mutually exclusive possible events in the future is essential for everyday decision making. Previous studies have suggested that this ability develops between the ages of 3 and 5 years, and in young children is primarily limited by the ability to represent the set of possible outcomes of an event as "possible". We tested an alternative hypothesis that this ability may be limited by the ability to represent the set of possible actions that could be taken to prepare for those possible outcomes. We adapted the inverted y-shaped tube task of Redshaw and Suddendorf (2016), in which children are asked to catch a marble that is dropped into the top of the tube and can emerge from either the left or right branch of the tube. While 4-year-olds typically place their hands under both openings to catch the marble, preparing for both possible outcomes (optimal action), 3-year-olds often cover only one opening, preparing for only one possible outcome (suboptimal action). In three Experiments, we asked whether first showing children the set of possible actions that could be taken on the tube would enable them to recognize the optimal action that should be used to catch the marble (Experiments 1 and 3, total $n = 99$ US 3- and 4-year-olds) and enable them to use the optimal action themselves (Experiment 2, $n = 96$ US 3- and 4-year-olds). We found that 3- and 4-year-olds performed similarly when they were given the opportunity to observe the set of possible actions beforehand. These findings suggest that 3-year-olds' competence at representing mutually exclusive possibilities may be masked by their developing ability to represent and deploy plans to act on these possibilities.

1. Introduction

Making decisions about what actions to take during complex situations in the real world often involves making predictions about the set of possible future events and then generating the potential set of actions that could be taken in response to those events. For example, imagine you are planning a dinner party, and you are uncertain whether your vegan friend is going to be able to attend. Although only one future outcome can *actually* occur (your friend either will be there, or they won't be there), you may opt to cook some vegan and some non-vegan dishes in order to prepare for *both* possible outcomes. While this is a small and somewhat trivial example, humans deploy this process regularly to make a range of decisions, in which actions now are planned in response to the potential future time course of events. These decisions range from relatively low-stakes decisions (like whether to carry an umbrella when there is a possible chance of rain) to much higher-stakes decisions (like whether to build additional supports into a building in an earthquake zone).

The ability to prepare for a variety of possible, often mutually exclusive future events requires the coordination of multiple cognitive and action processes. It requires the ability to think about events that have not yet occurred (i.e., mental time travel; Atance, 2015; Atance, Louw, & Clayton, 2015; Cheng, Werning, & Suddendorf, 2016; Martin-Ordas, 2018; Suddendorf, 2017; Suddendorf & Corballis, 2008), the ability to identify points of potential *divergence* in the future timeline (that is, points in time at which the time course of events branches off into distinct but possible future outcomes; Bulley, Redshaw, & Suddendorf, 2020; Redshaw et al., 2019), and the ability to plan and execute the relevant actions to take to prepare for those future outcomes (Atance & O'Neill, 2005; Blankenship & Kibbe, 2019; Prabhakar & Ghetti, 2020;

* Corresponding author.
E-mail addresses: esranurk@bu.edu (E.N. Turan-Küçük), kibbe@bu.edu (M.M. Kibbe).

https://doi.org/10.1016/j.cognition.2023.105712
Received 23 May 2023; Received in revised form 19 December 2023; Accepted 21 December 2023
Available online 30 December 2023
0010-0277/

Prabhakar & Hudson, 2014). That is, to prepare for multiple, mutually exclusive possibilities, one needs to not only represent the space of possible *outcomes* of future events, but also the space of possible *actions* that could be taken on those outcomes.

Earlier work established that, at least by age 5 years, children can prepare for multiple possible but mutually exclusive outcomes of an event (Beck, Robinson, Carroll, & Apperly, 2006; Robinson, Rowley, Beck, Carroll, & Apperly, 2006), and recent research has examined the developmental emergence of this ability. Redshaw and Suddendorf (2016) adapted the task of Robinson et al. (2006) for younger children and non-human primates by reducing the verbal instructions required for the task. They showed 2–4-year-old children an inverted y-shaped tube into which an experimenter could drop a marble such that it was possible for the ball to emerge from *either* the left *or* right exit of the tube. Children were then given the opportunity to catch the marble when it emerged from one of the two exits on the bottom of the tube. Since children could not be certain which branch of the tube the ball would emerge from, in order to catch the ball, children needed to represent *the set of potential future locations* of the ball (left branch or right branch), and then *take the correct actions to prepare for both potential future outcomes* – placing a hand under both openings of the tube to cover both possibilities and guarantee the catch. They found that 4-year-old children spontaneously and consistently placed a hand under each exit of the inverted y-shaped tube, while 3-year-olds were more likely to spontaneously place their hands under only one opening (for replications, see Redshaw, Leamy, Pincus, & Suddendorf, 2018; Redshaw et al., 2019; Suddendorf, Crimston, & Redshaw, 2017). Later work suggested that, when younger children do spontaneously cover both exits, they do not necessarily do so because they recognize that it is the correct strategy (Leahy, 2024).

Children's behavior in tasks like the y-shaped tube task has led to the suggestion that children's ability to simultaneously represent multiple, mutually exclusive possible future outcomes of an event does not develop around until the age of 4 or older (e.g., Leahy & Carey, 2020; Redshaw & Suddendorf, 2016; Redshaw et al., 2018; Suddendorf et al., 2017; Harris, 2022; Gautam, Suddendorf, & Redshaw, 2021; see also Beck et al., 2006; Robinson et al., 2006). Converging evidence for this view comes from other variations on the Y-shaped tube task in which children are asked to prepare for future events that either are guaranteed or are merely possible. For example, Leahy (2023) showed 2.5–4-year-old children scenarios involving two "slides": an inverted y-shaped slide, and a straight slide that did not branch. Children were told that an experimenter would drop two marbles, one into each of the openings at the top of the slides, and children were asked to place a container under the opening of one of the slides to catch a marble. If children perform optimally, they should always place the container beneath the straight slide, since doing so would guarantee that they would catch the marble. While 3- and 4-year-old children did use this strategy on the majority of trials, both groups sometimes placed a container under one of the branching openings. Similar results were obtained in studies in which children did not have to prepare for a future action, but to instead reason about the present location of two objects (Leahy, Huemer, Steele, Alderete, & Carey, 2022; Mody & Carey, 2016). This pattern of results has led to the suggestion that, while some young children may represent mutually-exclusive possible future outcomes of an event, other young children may have only a minimal representation of possibility (c.f. Leahy et al., 2022): they focus on only one of the future outcomes, represent that outcome as a sure thing (even when it is only merely possible), and behave accordingly (by taking an action in response to that one possible outcome, like placing the container under one of the two branches of the y-shaped slide). Older children, who may have more conceptual competence to represent the modal relationship between possibilities, may be less likely to make errors (Leahy & Carey, 2020; Ozturk & Papafragou, 2015; Rafetseder, Cristi-Vargas, & Perner, 2010; Redshaw & Suddendorf, 2020). If this is the case, then 3-year-olds faced with Redshaw and Suddendorf (2016) Y-shaped tube task may struggle to prepare for multiple future possibilities because they lack the competence to represent multiple possibilities simultaneously, and are therefore unable to represent the full set of possible future *outcomes* of the ball dropping event.

However, recent evidence suggests that, under some circumstances, younger children may be able to differentiate between mutually exclusive possibilities in pared-down tasks that make fewer verbal and physical reasoning demands. For example, Goddu, Sullivan, and Walker (2021) found that toddlers showed evidence of being able to entertain mutually exclusive possible hypotheses for a causal relationship. Infants show some evidence of using mutual exclusivity to reason about the identity of an ambiguous object (Cesana-Arlotti, Kovács, & Téglás, 2020; Cesana-Arlotti, Varga, & Téglás, 2022) or the referent of a novel label (Pomiechowska, Bródy, Csibra, & Gliga, 2021). In a recent study, Alderete and Xu (2023) showed 3-year-old children two gumball machines, one which contained one or two gumballs of a single color and the other which contained two gumballs of different colors, and asked children to select which gumball machine would yield a specific color gumball. They found that 3-year-olds successfully chose the gumball machine that contained *only* the requested color gumball. While this study does not necessarily present unequivocal evidence that young children are *representing* possibility – since all the possible marbles were simultaneously visible, children could use a strategy of simply avoiding the undesired color – it suggests that very young children may at some level be able to distinguish possible things from sure things.

Planning for mutually exclusive future possible events requires representing possible outcomes of the event *and* representing the potential actions that could be taken in response to those outcomes. The y-shaped tube task (Redshaw & Suddendorf, 2016; Robinson et al., 2006) and the slides task (Leahy, 2023) both require children to represent possibility (the possible future location of a marble/marbles) and to represent and select possible actions to prepare for those possibilities. If younger children have some competence with representing the space of mutually exclusive possibilities, perhaps their struggles in such tasks stem from difficulty representing the space of possible actions that could be taken on those future possible outcomes, or selecting the relevant action that they should perform from this possibility space, rather than representing the space of possible outcomes.

Indeed, 3-year-olds often struggle with tasks that require future-oriented thinking, even in circumstances where there is no mutual-exclusivity involved (see Atance, Ayson, & Martin-Ordas, 2023). For example, 3-year-olds struggle to generate narratives about what could happen in the future when asked open-ended questions (e.g., "Can you tell me something you are going to do tomorrow?"; Busby & Suddendorf, 2005). However, when the space of possibilities is constrained, 3-year-old children (and younger) show greater competence at thinking about the future, and even at making plans for future events (Atance & O'Neill, 2005; Blankenship & Kibbe, 2019, 2022, 2023; Prabhakar & Hudson, 2014). For example, Atance and O'Neill (2005) found that 3-year-olds were more successful at choosing the relevant items to take on a future trip, and explaining their reasoning behind their choices, when they were given fewer items to choose from (i.e., a set of four choices versus a set of eight choices). That is, when young children were shown a limited set of future options, they were able to recognize the correct options that could be needed in the future.

Given this previous research, we hypothesized that a primary limiting factor on younger children's ability to plan for multiple, mutually exclusive future possibilities may not be their ability to represent the space of mutually exclusive possibilities for the outcome of a future event, but instead their ability to represent the space of *possible future actions* that could be taken on those possible outcomes. To test this hypothesis, we conducted three experiments with 3-year-olds and 4-year-olds using variations on Redshaw and Suddendorf (2016) version of the y-shaped tube task.

In Experiment 1, we asked whether young children could recognize and select the optimal approach to prepare for mutually exclusive possible future events when they are not required to take any action themselves. We showed $n = 50$ 3- and 4-year-old US children videos of two adult actors catching a marble from an inverted y-shaped tube. One adult actor used the optimal strategy (covering both openings of the tube) while the other adult actor used a suboptimal strategy (covering only one opening of the tube). Crucially, *both* actors' strategies were successful: each actor caught the marble each time, and each caught the marble once from the left side of the y, and once from the right side of the y. Thus, while the actions that the actors took differed, the outcomes of those actions were identical. After children watched the videos, they were told that the actors would "play the marble game again" (i.e., in the future) and were asked to choose which of the two actors would catch the marble "for sure". Thus, the task did not require children to take an action to prepare for a future event themselves, but to instead select (from all available options) which actor should act – and hence which strategy should be taken – in the future. We reasoned that, if 3-year-olds' difficulty in the y-shaped tube task is largely driven by difficulty representing mutually exclusive possible *outcomes* of an event (as suggested by Leahy and colleagues), 3-year-olds should choose between the two actors at roughly equal rates, since both actors' strategies are consistent with a minimal representation of possibility. However, if 3-year-olds' difficulty in the y-shaped tube task is instead driven by difficulty generating representations of the space of possible *actions* that can be taken to prepare for these possibilities in the future, we predicted that younger children may be able to successfully *recognize* which of the strategies would result in a higher likelihood of catching the marble at a future time when they are shown the space of all possible actions, and would therefore select the actor who used the optimal strategy over the actor who used the suboptimal strategy. We predicted that 4-year-olds should consistently choose the optimal actor, since they have previously been shown to be able to consistently take the optimal action themselves in the inverted y-shaped tube task.

In Experiment 2, we asked whether showing children the space of possible actions that can be taken on the inverted y-shaped tube would impact their ability to select and deploy the optimal strategy themselves. Ninety-two 3- and 4-year-old US children were asked to catch a marble dropped into the inverted y-shaped tube. Half of the children ($n = 48$; Intervention condition) first watched the videos of the two adult actors from Experiment 1 demonstrating the suboptimal and optimal actions, while the other half ($n = 48$; Control condition) did not receive any intervention (a replication of Redshaw & Suddendorf, 2016). We reasoned that, if younger children's difficulty with the inverted y-shaped tube task is driven by difficulty with representing the space of possible actions that could be taken on the tube, then constraining the problem by showing children the space of all possible actions could help 3-year-olds use the optimal strategy themselves. We therefore predicted that 3- and 4-year-olds in the Intervention condition would perform similarly well on the task, while in the Control condition we should observe the developmental pattern observed in previous work by Redshaw and Suddendorf (2016), in which 4-year-olds outperform 3-year-olds.

Finally, in Experiment 3, we asked whether children would be able to recognize and select which strategy was optimal even when they did not observe the actors catching marbles. We told children that they were going to watch *how* the actors played the marble game, and then showed them cropped videos in which children only saw the actor place their hands under one or both exits of the tube (but did not observe a marble being dropped or caught). We then told children that the actors were ready to play the marble game, and asked children to choose which actor would "catch the marble for sure". We predicted that, if children's ability to recognize and select the optimal strategy was dependent on them simulating the path of the marble and observing how the actors' strategies lined up with their own simulations, children should select between the two actors at rates not different from chance. However, if children recognize that the act of covering both exits covers both mutually-exclusive possibilities, they should select the optimal actor. Also in Experiment 3, we asked whether children represent the optimal actor's actions as optimal *only* in a situation in which she is faced with mutually exclusive future possible outcomes. Specifically, we asked whether children who selected the optimal actor would go on to over-imitate her strategy even when that strategy was not relevant for another task. After children made their actor selection, we asked children to catch a marble that could be dropped into one of two openings of two *parallel tubes*. Because there is no ambiguity about where the marble will emerge, the optimal strategy is to place hands under only one opening. If children who chose the optimal actor over-imitate her strategy, they should place their hands under both exits of the parallel-tube apparatus. If they recognize that the optimal actor's strategy only makes sense in the context of uncertain, mutually exclusive possible future positions of the marble, they should not over-imitate that strategy on the parallel-tube apparatus, and instead place their hands under only the relevant exit.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Fifty children (24 3-year-olds, mean age: 42 months, 4 girls, 6 boys, 14 gender unreported; and 26 4-year-olds, mean age: 53 months, 7 girls, 6 boys, 13 gender unreported) were tested in the the Museum of Science, Boston. The sample size was determined to be sufficient to detect differences between 3- and 4-year-olds' choices using an estimate of effect size based on children's own approach to the tube task in previous work by Redshaw and Suddendorf (2016) (Fisher's exact test, two-tailed, $p1 = 0.38$, $p2 = 0.8$, $\alpha = 0.05$, $1\text{-}\beta = 0.8$, suggested $n = 48$). An additional two children participated but were excluded from analysis because they did not attend to the videos. The study was approved by the Institutional Review Boards of the Boston University Charles River Campus and the Museum of Science.

2.1.2. Apparatus and stimuli

Children at the museum watched a series of videos on a 13.3-in. laptop computer presented using the Qualtrics survey platform. Children watched a series of videos in which different actors interacted with an inverted y-shaped tube and a yellow plastic marble (~2 cm diameter) (see Fig. 1). The tube apparatus was constructed following the specifications provided in Redshaw and Suddendorf (2016). The tube was made of cylindrical PVC pipes (9 cm diameter) connected with 90° curved pipe fittings so that there was one opening at the top and a fork with two openings at the bottom. The entire apparatus measured approximately 50 cm high from the top opening to the center of the Y, and the bottom of the apparatus approximately 20 cm wide from the right to left opening. The experimenter could surreptitiously control which side of the fork the marble emerged from by manipulating a hidden cardboard switch that ran along the inside of the upper tube, closing off one side of the fork and forcing the marble to travel down the other. All videos used in this experiment can be found at https://osf.io/5k9ms/.

2.1.3. Procedure

Children were seated in a chair at a table across from the experimenter.

2.1.3.1. Familiarization. Children first watched a video of the six Familiarization trials (Fig. 1; Supplemental Video S1). In the first Familiarization trial, children saw an actor (henceforth referred to as the Dropper) standing behind the inverted y-shaped tube holding a yellow marble over the top opening of the tube. They then heard a recorded female voice-over saying, "I want to show you a game. This is a marble game. Someone will drop a marble in the tube, and another person gets to catch it. Okay? First, let's watch how the marble comes out of the

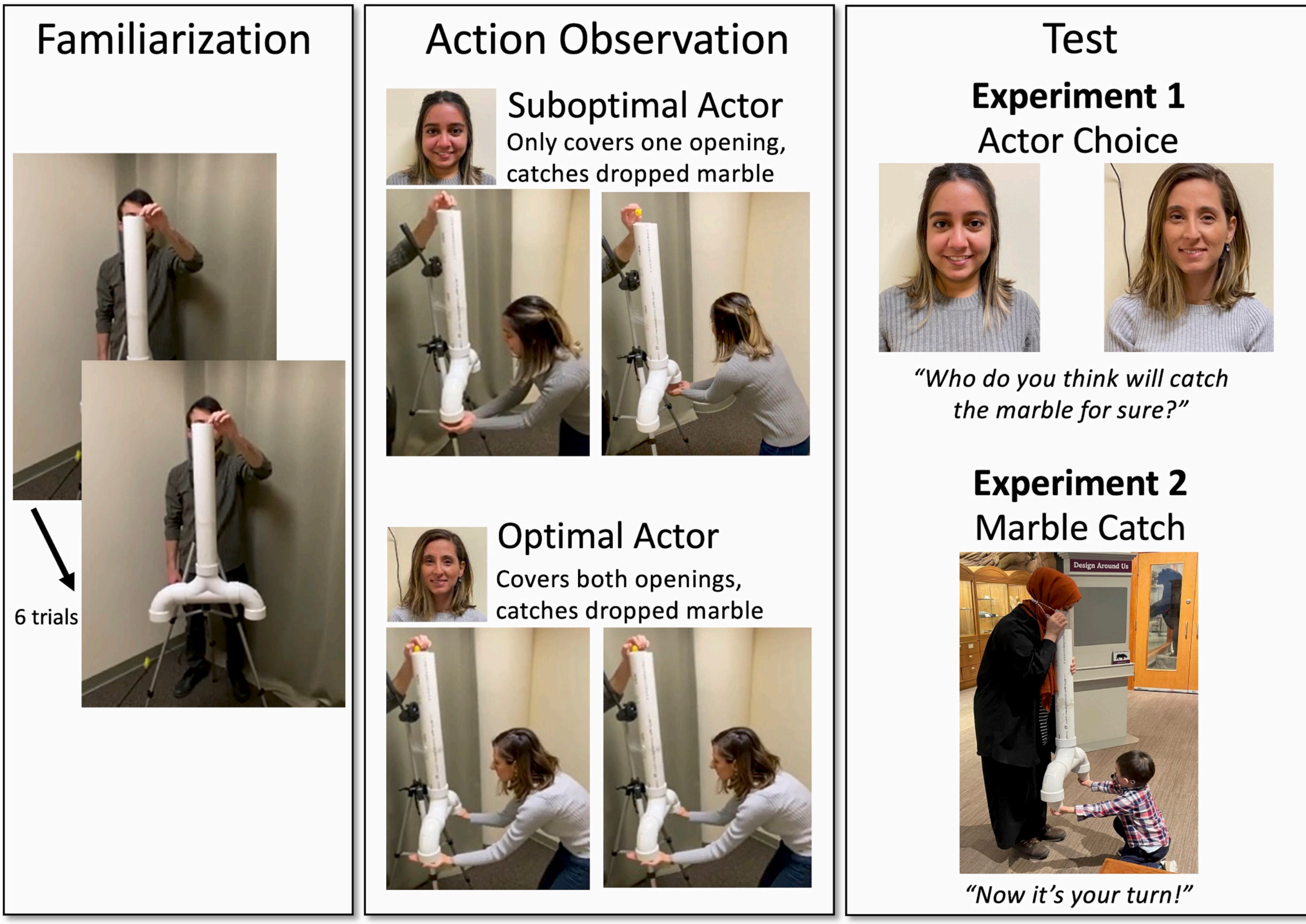

Fig. 1. Familiarization, Action Observation, and Test trials from Experiments 1 and 2. In Experiment 1, all children completed Action Observation trials. In Experiment 2, only children in the Intervention condition completed Action Observation trials.

tube. Watch, he will put the marble into the top, and then look what happens." The Dropper then released the marble, which subsequently dropped out of the right side of the tube and onto the floor. Children then saw a white screen displaying the words "Let's do that again" in black letters, and heard the voice-over say, "Let's do that again!".

Children then watched five more familiarization trials. On each of these trials, the Dropper held the yellow marble over the opening at the top of the tube while the voice-over said, "Watch, he will put the marble into the top, and then look what happens." The Dropper then released the marble, which emerged from one side of the tube and dropped to the floor. The side the marble emerged from was pseudorandomized across trials, following Redshaw and Suddendorf (2016) (right, left, left, right, left, right). The entire 6-trial sequence took 81 s to complete.

2.1.3.2. Action observation trials. Children first saw two still photos showing two different actors from the shoulders up. Captions above the photos showed the actors names, while a voice-over said, "This is Suzy [Suzy's photo wiggled back and forth to draw attention]. This is Jane [Jane's photo wiggled back and forth to draw attention]. They will play the marble game. But they do not know which side of the tube the marble will come out" (Supplemental Video S2).

Children then watched a series of videos depicting the two different actors using different strategies to catch a marble dropped into the tube by the Dropper (Fig. 1; Supplemental Videos S3 and S4). One actor (Suzy) used the "optimal" strategy of placing a hand under each of the openings of the tube (which guarantees the catch). The other actor (Jane) used the "suboptimal" strategy, placing both hands under only one opening of the tube (for a 50% chance of catching the marble on each trial). Children saw two examples of each actor using their respective strategies to catch the marble. Crucially, both actors caught the marble on each trial, regardless of the strategy they used, and both actors used both hands to deploy their strategies. This ensured that the *only* difference between the actors was the strategy they used to catch the marble, and not outcome of deploying the strategy or the effort required to use the strategy.

The order of presentation of the different actors' videos, and the side from which the marble emerged first, was counterbalanced across children (half of the children saw the Optimal Actor catch the marble two times, then saw the suboptimal Actor catch the marble two times; the other half of children saw the opposite order). For convenience, here we describe the condition in which the optimal actor is shown first (and the marble emerged from the left first), followed by the suboptimal actor.

In the first of two Optimal Actor trials (Supplemental Video S3, Fig. 1), children saw a screen depicting only Suzy's photo. The voice-over said, "See, this is Suzy. Suzy is going to try to catch the marble, but remember, she doesn't know where the marble is going to come out.[1] Let's see what Suzy does!" Children next saw the Dropper standing behind the tube holding the marble above the opening, while Suzy stood in front of the tube with her eyes fixed on the center of the bottom of the tube (i.e., the point at which the two branches of the tube diverge). Suzy then bent down and placed one hand under each tube opening (right hand under the right opening, left hand under the left opening). The Dropper then dropped the marble, which emerged from the *left* side of the tube. Suzy caught the marble and then lifted it up to show the Dropper. In the second Optimal Actor trial, children again saw a screen depicting only Suzy's photo. The voice-over said, "Suzy is going to play again! Remember, she doesn't know where the marble is going to come out. Let's see what Suzy does!" Children then watched Suzy place a hand under each opening, catching the marble from the *right* side of the tube, and showing it to the Dropper. In total, the entire Optimal Actor two-trial sequence took 45 s to complete.

Children then watched two Suboptimal Actor trials (Supplemental Video S4, Fig. 1). These trials proceeded similarly to the Optimal Actor trials, except that children were first shown Jane's photo, and were told that "Jane is going to try to catch the marble, but remember, she doesn't know where the marble is going to come out. Let's see what Jane does!" Children then watched Jane catch the marble after placing both hands (one hand on top of the other) under only one opening. In one trial, she covered the *right* opening, and in the other she covered the *left* opening (order counterbalanced across children). In both trials, she caught the marble and showed it to the Dropper. In total, the entire Suboptimal Actor two-trial sequence took 45 s to complete.

2.1.3.3. Test trial. In the single Actor Choice Test trial (Supplemental Video S5; Fig. 1), the experimenter showed children the photos of the Optimal Actor and Suboptimal Actor (whether the Optimal Actor was on the right or the left was counterbalanced across participants) and said, "We will play the marble game again, but this time *you* get to choose who plays the game. Remember, our goal is to catch the marble. Who will catch the marble for sure?" Children selected an actor by pointing to her photo or by saying the actor's name (43 children pointed, 7 children said the actor's name). The experimenter immediately entered children's responses by selecting a radio button under the chosen actor's photo. As an additional exploratory measure, the experimenter then asked children "Why did you choose [actor's name]?" and typed their responses into the form (we did not conduct formal analyses on these responses; descriptive statistics of children's responses to this question can be found in Supplemental Table S1).

2.2. Results

Children in both age groups chose the Optimal Actor at rates significantly above chance (chance = 50%; 3-year-olds: 20/24 children (83%) chose the Optimal Actor, binomial test $p = .002$); 4-year-olds: 23/26 children (88%) chose the optimal actor, binomial test $p < .001$; Fig. 2, left panel), and there was no difference in children's responses between the age groups (Fisher's exact test $p = .697$). Data for Experiment 1 can be found at https://osf.io/5k9ms/.

2.3. Discussion

Previous work has shown that 3-year-olds who were asked to catch a marble from an inverted y-shaped tube often used a suboptimal or mixed strategy, placing their hands under only one opening of the tube (e.g., Redshaw & Suddendorf, 2016), which has been taken as part of the evidence that 3-year-olds struggle with modal representation of multiple, mutually exclusive possibilities (Leahy & Carey, 2020). In Experiment 1, we asked whether 3-year-olds could distinguish between suboptimal and optimal strategies and select which strategy should be used in the future to catch the marble "for sure". After watching videos of adult actors deploying different approaches to catch the marble from the inverted y-shaped tube, both 3- and 4-year-olds were able to choose which actor's strategy would likely result in future success. Importantly, children's choices could not have been driven by observing the success rate of these different strategies, since both actors always succeeded in catching the marble, regardless of the strategy they used and regardless of which branch of the inverted y the marble emerged from. Further, children's choices could not have been driven by the effort that was expended by the actors to catch the marble, since both actors used both hands to deploy their action (i.e., the suboptimal actor placed two hands under one opening, the optimal actor placed a hand under each opening). Instead, the results of Experiment 1 suggest that 3-year-olds, who typically struggle to take the correct action on their own, nevertheless recognized the best strategy to act on multiple, mutually exclusive future possible events when they were shown the set of possible actions that could be taken on the inverted y-shaped tube apparatus. The results of Experiment 1 suggest that 3-year-olds' ability to prepare for mutual exclusive possibilities of an event may be limited more by their ability to represent the set of possible actions that can be taken on an event.

In Experiment 2, we asked whether showing 3-year-olds the space of possible actions that can be taken on the inverted y-shaped tube also might help them to spontaneously *select and deploy* the optimal strategy themselves. We asked 3-year-olds, who typically use a suboptimal or mixed strategy, and 4-year-olds, who typically use the optimal strategy, to catch the object dropped into the inverted y-shaped tube. Half of the children observed the Action Observation trials from Experiment 1 before doing the inverted y-shaped tube task, while the other half of children did not observe any possible strategies for acting on the tube before they were asked to take an action themselves (replicating Redshaw & Suddendorf, 2016). We hypothesized that the 3-year-olds who saw the set of possible actions before they themselves did the tube task would perform more like 4-year-olds in the task.

3. Experiment 2

3.1. Method

3.1.1. Participants

Ninety-six 3- and 4-year-old children were tested at the Museum of Science, Boston. Sample size was determined as in Experiment 1, except we doubled the sample to examine effects of condition (Intervention vs. Control). Forty-eight children participated in the Intervention condition (24 3-year-olds, mean age: 43 months; 15 girls, 9 boys; and 24 4-year-olds, mean age: 54 months; 11 girls, 13 boys) and 48 children participated in the Control condition (24 3-year-olds, mean age: 43 months; 10 girls, 13 boys, and 1 non-binary; and 24 4-year-olds, mean age: 54 months; 9 girls, 15 boys). An additional 5 children participated but were excluded from analysis due to parental interference (3), declining to complete the study procedures (1), or equipment malfunction (1). We first collected the sample for the Intervention condition and then proceeded to collect the sample for the Control condition. The study was approved by the Institutional Review Boards of Boston University Charles River Campus and the Museum of Science, Boston.

3.1.2. Apparatus and stimuli

Children in both conditions were asked to catch marbles dropped into the inverted y-shaped tube depicted in the videos in Experiment 1 (see Experiment 1 *Apparatus* section for specifications). To accommodate child-sized hands, we made the bottom tubes' openings slightly

[1] By design, both actors caught the marble on each trial, regardless of the strategy they used. We therefore emphasized to children that the actors *did not know* which side of the inverted y-shaped tube the marble would emerge from, so that children would not attribute the actors' strategies to special knowledge about the outcome of the drop (i.e. so they would not think that the suboptimal actor covered only one opening of the tube because she had foreknowledge that the marble would emerge from that opening).

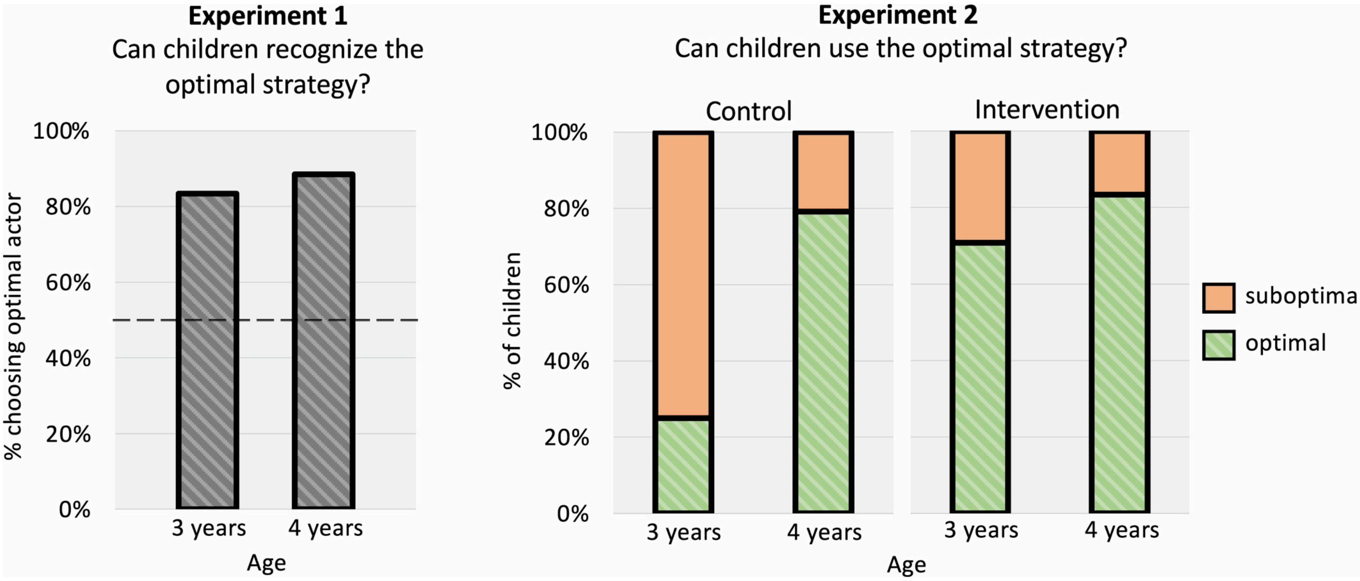

Fig. 2. The left panel shows the percentage of children choosing the optimal actor in response to the prompt "Who will catch the marble for sure?" in Experiment 1 (dashed line shows chance level, 50%). The right panel shows the percentage of children who used the optimal strategy (covering both openings of the tube) and the suboptimal strategy (covering only one opening) when asked to catch the marble themselves in Experiment 2. Children in the Intervention condition completed Action Observation trials before being asked to catch the marble. Children in the Control condition received no intervention.

smaller (5 cm) using white cardboard (see Fig. 1). We used the Familiarization and Action Observation videos from Experiment 1. The inverted y-shaped tube was placed out of children's view while they watched the videos.

3.1.3. *Procedure*

Children began by sitting with the experimenter at a table in the museum. Children in the Intervention condition watched both the Familiarization and Action Observation Trials videos from Experiment 1 (the order of the Action Observation trials was counterbalanced as in Experiment 1). Children in the Control condition watched only the Familiarization videos from Experiment 1. All children then proceeded immediately to the Test trials.

3.1.3.1. *Test trials.* The experimenter said to children: "Okay, now it is your turn to play the marble game. Let me grab the game." The experimenter took the tube out, held it up, and invited children to get up from the chair and stand in front of the tube. Children then completed two Marble Catch test trials. In the first Marble Catch test trial, the experimenter held the marble over the opening at the top of the tube and said, "Okay, are you ready? Can you catch the marble?" The experimenter waited until children placed their hand/s either under one or both openings (usually 3–5 s). Once children had positioned their hand(s), the experimenter dropped the marble into the top of the tube. If children caught the marble, the experimenter said, "Alright!" and asked for the marble back. If children did not catch the marble, the experimenter said, "Alright!" and retrieved the marble from the floor. In the second Marble Catch test trial, the experimenter said, "Okay, we will play again!" and then proceeded as in the first Marble Catch trial.

Children's catching strategy (placing their hands under both openings or only one opening, and which side they covered if they covered only one side) was recorded immediately after each trial by the experimenter.

For all children, in one Marble Catch test trial the marble emerged from the right side of the tube, while in the other test trial the marble emerged from the left side of the tube (side order was counterbalanced across children).

3.2. *Results*

Analyses were conducted on the first trial only in order to obtain a "pure" measure of children's catching strategies in the two conditions (i.e., before children have a chance to learn by trial and error with the apparatus; see Redshaw & Suddendorf, 2016). Results for the second trial can be found in the Supplement. We used pairwise Fisher's exact tests to investigate the differences between age groups in the Control and Intervention conditions. This decision was made after discussion with a statistics consultant; due to the low variability we observed in the data, other statistical methods such as Generalized Linear Mixed Models (GLMMs) or chi-square tests were deemed unsuitable for our analyses due to concerns related to model convergence and distribution assumptions. Data for Experiment 2 can be found at https://osf.io/5k9ms/.

The results from the first Marble Catch test trial are shown in Fig. 2. In the Control condition, we found that 4-year-olds used the optimal strategy significantly more than 3-year-olds (3-year-olds: 6/24 children (25%) used the optimal strategy; 4-year-olds: 19/24 children (79%) used the optimal strategy; Fisher's exact test $p < .001$, two-tailed), consistent with previous work (e.g., Redshaw & Suddendorf, 2016). In the Intervention condition, we found that the majority of children in both age groups used the optimal strategy (3-year-olds: 17/24 children (71%) used the optimal strategy; 4-year-olds: 20/24 children (83%) used the optimal strategy) and there was no difference between the age groups in rates of optimal strategy use (Fisher's exact test, $p = .494$, two-tailed). Three-year-olds in the Intervention condition used the optimal strategy significantly more than their age-matched counterparts in the Control condition (Fisher's exact test, two-tailed, $p = .001$), while there was no difference in 4-year-olds' strategies between the two conditions (Fisher's exact test, two-tailed, $p = 1.0$). Children's performance in the second Marble Catch test trial was similar to their performance in the first Marble Catch test trial (Intervention: 96% of 3-year-olds and 96% of 4-year-olds used the optimal strategy; Control: 25% of 3-year-olds and 88% of 4-year-olds used the optimal strategy; see Supplement for details).

3.3. Discussion

In Experiment 2, we asked whether showing 3-year-olds the set of possible actions that can be taken on the inverted y-shaped tube would improve their ability to optimally prepare for multiple, mutually exclusive future outcomes of an event. We replicated the developmental effect of Redshaw and Suddendorf (2016) – with 4-year-olds outperforming 3-year-olds in using the optimal strategy – in our Control condition, in which 3- and 4-year-old children were simply asked to catch the marble from tube after observing a series of familiarization trials in which they saw the marble being dropped into the top of the tube and emerging from one side or the other. By contrast, in our Intervention condition, when children observed actors deploying the range of possible approaches to the tube task before they were asked to catch the marble themselves, both 3-year-olds and 4-year-olds used the optimal strategy at similar rates, and 3-year-olds in the Intervention condition used the optimal strategy significantly more than 3-year-olds in the Control condition.

Together, the results of Experiments 1 and 2 suggest that 3-year-olds can represent multiple, mutually exclusive possible future outcomes of an event, but this representational competence may be masked by their limited ability to represent the space of possible *actions* to take in response to those possibilities. However, there are some alternative explanations for younger children's performance in Experiments 1 and 2 that do not require granting children conceptual competence with mutual exclusive possibilities.[2]

First, 3-year-olds in Experiment 1 may have selected the optimal actor because her actions were more consistent with the child's own *minimal representation of possibility* than the suboptimal actor's actions. Because children watched the entire sequence of events unfold (Dropper holding marble, actor placing hands, dropper dropping marble, actor catching marble), children had the opportunity to *simulate* the path of the marble, and to then compare the outcome of their simulation to the outcome shown by the actor. If children were simulating the path of the marble and consequently representing the marble in one location (and not the other), then the optimal actor had a 100% chance of catching the marble from the exit that accorded with the child's own representation. By contrast, the suboptimal actor had only a 50% chance of catching the marble from the side that accorded with the child's representation of the marble's location on each trial (that is, a 25% chance across both trials). Children's subsequent selection of the optimal actor when asked "who will catch the marble for sure?" may therefore reflect the child's observation of the optimal actor's consistency with their own minimal representation of possibility, rather than the child's recognition that the optimal actor's strategy covered *both* possible future outcomes. If this is the case, then we would expect 3-year-olds' ability to choose the optimal actor when asked "who will catch the marble for sure" to be limited to conditions under which they had been given the opportunity to *simulate* the path of the marble and observe whether the actor's actions are consistent with their simulation.

Second, 3-year-olds in Experiment 2, who went on to use the optimal strategy themselves after observing the Action Observation trials, may have been doing so not because they recognized that the optimal strategy was the better one, but instead because they were *over-imitating* the optimal actor's actions. That is, 3-year-olds' behavior in Experiment 2 may show that they can use the action of the optimal actor, but not that they *understand why* the strategy is the better one. If this is the case, we would expect children to over-imitate the actor's actions even when it doesn't make sense to do so (e.g., when there is no ambiguity about where a marble will emerge).

We attempted to address both of these alternatives in Experiment 3. Children again observed a series of Action Observation trials. However, unlike in Experiments 1 and 2, we told children that they were going to watch *how* the actors play the marble game. We cropped the videos so that there was no Dropper in the video and therefore no marble dropped into the tube. Instead, children simply observed the suboptimal actor place her hands under one exit, and the optimal actor place her hands under both exits. Since no marble was dropped into the tube, there was nothing for children to simulate. And since neither actor caught a marble, there was no outcome for children to match to their own representation of the marble's location.

We had two dependent measures. First (similar to Experiment 1), we told children that the actors would play the marble game, and asked them who would catch the marble for sure. Second (similar to Experiment 2), we told children that they were going to get to play a marble game. We showed children an apparatus consisting of two, separate parallel tubes, such that there were two exits (like in the inverted y-shaped tube apparatus), but also two entrances. When the experimenter held a marble over one of the two entrances, there was no ambiguity about which exit the marble would emerge from. We measured children's hand positions when asked to catch a marble themselves. For the parallel-tubes apparatus, placing hands under both exits would be inefficient and non-obvious (although it would result in a catch), while placing hands under only the relevant exit would be the most efficient strategy.

4. Experiment 3

4.1. Method

4.1.1. Participants

Forty-nine children (24 3-year-olds, mean age: 40 months, 27 days, 13 girls, 11 boys; and 25 4-year-olds, mean age: 54 months 6 days, 15 girls, 8 boys, 2 gender unreported) were tested at the Museum of Science, Boston. Sample size was determined as in Experiment 1. The study was approved by the Institutional Review Boards of Boston University Charles River Campus and the Museum of Science, Boston.

4.1.2. Apparatus and stimuli

Children were presented with a series of videos displayed on a 13.3-in. laptop computer presented using the Qualtrics survey platform (full video stimuli are available at https://osf.io/5k9ms/). Children also interacted with a tube apparatus consisting of two parallel PVC pipes (each 60 cm in length and 9 cm in diameter). Two thin metal bars kept the tubes 20 cm apart and allowed the experimenter to grasp and hold the apparatus between the two tubes with one hand while dropping a marble into one of the openings with the other hand. The exits on the parallel-tubes apparatus were the same distance apart as the exits on the inverted y-shaped tube apparatus used in Experiment 2. Fig. 3 depicts the parallel tube apparatus.

4.1.3. Procedure

Children first observed the six Familiarization trials from Experiments 1 and 2. Children then observed a series of Action Observation trials. First (as in Experiments 1 and 2) children viewed images of Suzy and Jane and were told "This is Suzy and this is Jane. They will play the marble game. They want to try to catch the marble. But they do not know which side of the tube the marble will come out." Children were then told that they would watch how the actors play the marble game ("Let's watch *how* they play the game!"). Below we describe the condition in which the optimal actor was shown first, but whether the optimal or suboptimal actor was shown first was counterbalanced across children. To facilitate comparison between the experiments, deviations from the script of Experiments 1 and 2 are italicized. See Supplemental Video S6 for a video of the Action Observation trials of Experiment 3.

In the two Optimal Actor trials, children were shown a photo of the optimal actor and heard a voiceover say, "See, this is Suzy. *This is how Suzy plays the marble game*. Remember, she doesn't know where the marble is going to come out. *Let's see how Suzy plays the marble game*."

[2] We thank two anonymous reviewers for pointing out these alternatives.

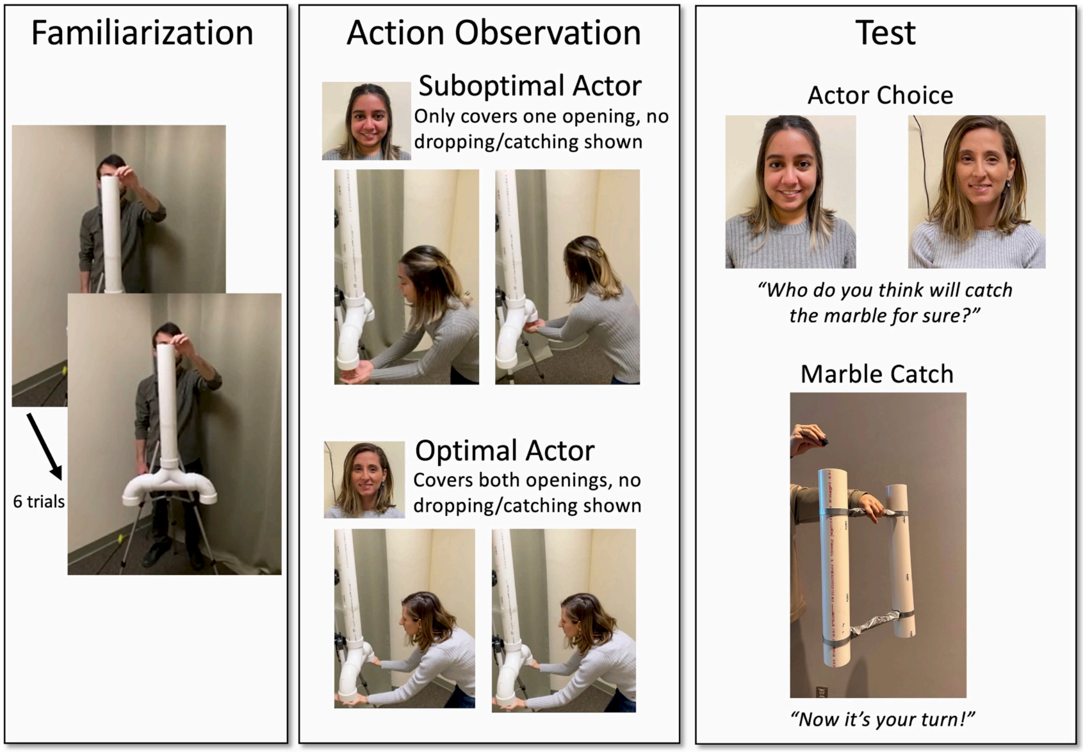

Fig. 3. Familiarization, Action Observation, and Test trials from Experiment 3. At Test, children in Experiment 3 completed an Actor Choice trial, immediately followed by two Marble Catch trials with the parallel tube apparatus.

Children then observed two videos of the optimal actor taking an action on the inverted y-shaped tube. In the first Optimal Actor trial, children observed a video in which the optimal actor stood facing the tube apparatus, and then placed one hand under each exit and paused for 5 s (Fig. 3). Children then observed the photo of the optimal actor, and heard the voiceover say, "*Let's watch Suzy again*. Remember, she doesn't know where the marble is going to come out. *Let's see how Suzy plays the marble game*." Children then again observed the Optimal Actor placing a hand under each exit. The two Suboptimal Actor trials proceeded similarly, except that children were told that they would watch "*how Jane plays the marble game*". They observed two Suboptimal Actor trials, one in which the suboptimal actor placed both hands under the right exit and paused, and one in which she placed both hands under the left exit and paused. Crucially, in all of the videos, children did not see the Dropper, did not observe a marble being dropped into the tube, and did not observe either actor catching a marble. The *only* thing depicted in the videos was the strategy by which the actors place their hands under the exit(s) of the tubes.

In the Actor Choice test trial, children then observed photos of the optimal and suboptimal actors. The experimenter said, "OK now Suzy and Jane *are ready to play the marble game*. But you get to choose who plays the game. Remember, the goal is to catch the marble. Who do you think will catch the marble for sure? Who do you want to pick?" (Fig. 3). As in Experiment 1, children indicated their selection by pointing or saying the actor's name, and the experimenter immediately entered children's responses by selecting a radio button under the chosen actor's photo (most children pointed, although we did not track individual children's response modality in Experiment 3). As in Experiment 1, we also included the exploratory question, "Why did you choose [actor's name]?"; descriptive statistics of children's responses to this question can be found in Supplemental Table S2.

All children then completed two Marble Catch trials (Fig. 3). The experimenter said, "Okay, *now I have a marble game for YOU to play*. Let me grab the game." Children then completed two Marble Catch test trials, which proceeded similarly to the Marble Catch test trials of Experiment 2, except the experimenter used the parallel-tube apparatus (Fig. 3). The experimenter held the marble over one of the two top openings of the tube (right or left, counterbalanced across trials and children), waited for children to place their hand(s) under the exit(s), and then dropped the marble. We recorded children's hand placement on each trial.

As in Experiment 2, children completed two Marble Catch trials, and we planned to analyze only the first Marble Catch trial in order to get a "purer" measure of children's strategies on the parallel-tube apparatus. We reasoned that, if children over-imitate the "both sides" strategy in a context in which it is not a relevant approach, they would be more apt to do so on the first Marble Catch trial, before they have experience with the parallel-tube apparatus. Children's performance on both trials can be found in the Supplement (Fig. S2).

4.2. Results

On the Actor Choice test trial, children in both age groups chose the optimal actor at rates significantly above chance (chance = 50%; 3-year-olds: 20/24 children (83%) chose the optimal actor, binomial test $p = .002$); 4-year-olds: 23/25 children (92%) chose the optimal actor, binomial test $p < .001$; Fig. 4, left panel) and there was no difference between 3- and 4-year-olds (Fisher's exact test $p = .42$). Children's actor choices in Experiment 3 (in which they had only observed the actors' strategies prior to the Actor Choice trial) were not significantly different from their choices in Experiment 1 (in which they had also observed the marble being dropped and caught) (Fisher's exact test $p = 1$).

On the first Marble Catch test trial, a majority of children in both age groups placed their hands under only the exit from which the marble could actually emerge (23/24 3-year-olds (96%); 20/25 4-year-olds (81%); Fig. 4, right panel), and there was no significant difference between 3- and 4-year-olds' strategies (Fisher's exact test $p = .19$).

We also compared the 3-year-olds who completed the Intervention condition of Experiment 2 to the 3-year-olds in Experiment 3. Recall that, in Experiment 2, 3-year-olds who observed the Action Observation

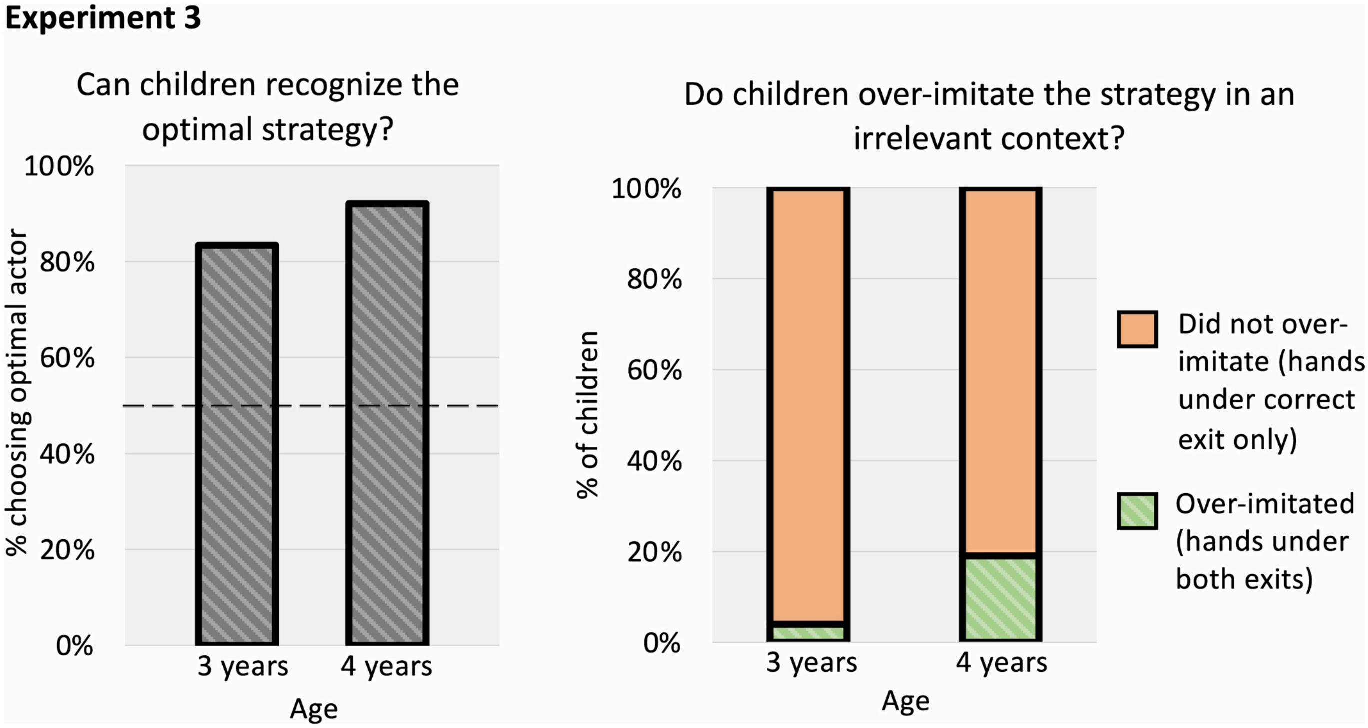

Fig. 4. Children's responses in the Actor Choice test trial (left panel) and the first Marble Catch test trial (right panel) of Experiment 3.

trials subsequently used the optimal strategy on the inverted y-shaped tube significantly more than children in the Control condition (who did not observe the Action Observation trials). If 3-year-olds' behavior in Experiment 2 was due to simply imitating the cover-both-exits strategy, we would expect children to use a similar approach in Experiment 3, even when it was not relevant to do so. Instead, 3-year-olds' approach in Experiment 2 was significantly different from their approach in Experiment 3, (Fisher's exact test $p < .001$).

4.3. Discussion

In Experiment 3, we found that when children were shown only the two strategies, but were not given the opportunity to simulate the paths of the dropped marbles nor observe the actors catching marbles, they were able to recognize the optimal strategy for preparing for both mutually exclusive future possibilities. Further, even though the majority of children chose the optimal actor when asked "who would catch the marble for sure", they did not subsequently imitate her actions in the parallel-tubes task when her strategy was not relevant.

Could children in Experiment 3 have been simulating the path of a marble in the Action Observation trials despite not having directly observing a marble being dropped into the tube nor observing either Actor interacting with a marble? Since children observed a series of Familiarization trials prior to viewing the Action Observation trials, they had some experience with observing that marbles dropped into the tube can emerge from either exit. Perhaps in the Action Observation trials, children simulated the trajectory of the marble based on their memory of the familiarization trials, matched the actor's actions with a possible outcome (e.g. simulated marble comes from the right exit, actor who places hands under both the exits would therefore catch it), and used these simulations as the basis for their choices. However, we think this unlikely for several reasons. First, it is not clear why children would do this given the structure of the Action Observation trials. Children were told that they would watch "how the actors play the marble game" and then saw videos of actors taking actions. The videos were shot from a different angle than the Familiarization trials, the top of the tube was obscured, and the actions of the actors were extremely brief. To the adult eye, there are no cues that would prompt one to simulate a marble (and we encourage readers to view Supplemental Video S6 to observe this for themselves). Second, the cognitive load involved in maintaining in memory the trajectories of the marbles of the Familiarization trials, and then applying simulations of those trajectories to the quite different Action Observation trials (entirely in the mind, in the absence of visual input), while also encoding and retrieving information from the Action Observation trials, would be considerable, and is not reasonably within the capacity of children in the 3–4-year-old age range. Instead, we argue that the results of Experiment 3 show that children's recognition of the effectiveness of the different actors' strategies does not depend on being able to simulate the path of the marble.

Together, the results of Experiments 1–3 suggest that young children have competence with representing multiple, mutually exclusive future possibilities, and seeing the set of possible actions that can be taken to prepare for those possibilities supports their ability to recognize the optimal action and to take the optimal action themselves.

5. General discussion

The ability to plan for multiple, mutually exclusive possibilities in the future is an essential part of everyday decision making (Bulley, Henry, & Suddendorf, 2016, Harris, 2022, Phillips, Morris, & Cushman, 2019, Redshaw & Ganea, 2022, Shtulman & Phillips, 2018). Previous research has suggested that children younger than 4–5 years often struggle to take the relevant actions to prepare for mutually exclusive future possibilities, often preparing for only one possible future outcome of an event (e.g., Leahy, 2023; Redshaw et al., 2018, 2019; Redshaw & Suddendorf, 2016; Suddendorf et al., 2017). This has led to the suggestion that a primary constraint on this ability in young children is the ability to simultaneously represent multiple, mutually exclusive possibilities (e.g., Leahy & Carey, 2020). However, preparing for mutually exclusive future possibilities requires not only representing the space of merely possible *outcomes* of future events, but also representing the space of possible *actions* that could be taken to prepare for those outcomes. In three experiments, we tested the hypothesis that a primary constraint on younger children's ability to act on mutually exclusive future possibilities may be their ability to represent the set of possible actions that can be taken to prepare for those outcomes and/or to select and deploy the optimal action from that set of possible actions. Our results suggest that young children's ability to prepare for mutually exclusive future possibilities may be limited more by their ability to generate representations of the set of possible actions that can be taken in response to possible outcomes, and less by their ability to represent the set of possible outcomes.

Our results are consistent with work that has shown that future-oriented thinking and planning abilities develop significantly between the ages of 3 and 5 years (see Atance et al., 2023, for review). Three-year-olds struggle with future planning in unconstrained tasks (e.g., Busby & Suddendorf, 2005) but show more success in more constrained tasks (e.g., Atance and O'Neill, 2005; Blankenship & Kibbe, 2019, 2022, 2023; Prabhakar & Ghetti, 2020). Our results extend this pattern to situations involving modal uncertainty and shed light on a potential source of developmental change in future-oriented planning. The fact that younger children recognize and quickly adopt optimal strategies, but do not typically spontaneously generate those strategies themselves, suggests that what develops in young children may not be the ability to *comprehend* future possibilities or the actions that should be taken to prepare for those possibilities, but the ability to *spontaneously represent the range of possible actions* to take on those possibilities.

We speculate that 3-year-olds faced with future scenarios that require planning may generate an initial representation of the set of possible actions that may be limited to something like one-to-one correspondence between actions and outcomes. For example, a 3-year-old in the inverted y-shaped tube task who first observes the series of familiarization trials (in which they watch the marble being dropped into the top of the tube and emerging from either the right or left side; Experiments 1 & 2; Redshaw & Suddendorf, 2016) may initially construct a limited representation of possible actions which could look something like: [the marble can emerge from the right branch, so a possible action is to cover the right side]; [the marble can emerge from the left branch, so a possible action is to cover the left side]. When then given the opportunity to catch the marble themselves, they select between these two, roughly equivalent representations of possible actions. However, when children were then able to observe a better strategy – one that did not directly correspond with one or the other possible outcome but accounted for both possible outcomes – children had an "Aha!" moment: they recognized that it was a better strategy (Experiments 1 and 3) and used that strategy themselves (Experiment 2). Our results suggest that children's initial limited representation of possible actions may not be a conceptual limitation, but may readily be expanded to include new strategies and new actions.

This hypothesis is consistent with the behavior of the subset of 3- and 4-year-olds in Leahy (2023) slides task who did not reliably take actions consistent with a guaranteed outcome. Children in this task first learn that they could place a container under the straight-path slide to catch a marble, and then they are shown an inverted y-shaped slide and learn that a marble can emerge either from the left or the right (but do not practice catching the marble from the inverted y-shaped slide). At test, children were shown two slides (e.g. a straight-path slide and an inverted y-shaped slide), the experimenter held one marble over each entrance (i.e. two marbles total), and children were asked to place the single container under an exit to catch one of the two marbles. Since the only action the child has learned to take (and indeed, the only action the child *can* take) is to place *one* container under *one* exit, children may initially construct a minimal representation of possible actions that includes [place the container under one opening to catch one marble]. When children are suddenly faced with two marbles in the test trials – one held over the entrance to the straight path, one held over the entrance to the branching path – this minimal representation of possible actions would drive children to place the container under the straight branch half the time, and place the container under one of the two other branches half the time (a similar explanation can be applied to the task of Leahy et al., 2022, a task with similar logic that does not require future-oriented planning). It is worth noting that 3- and 4-year-olds performed similarly in Leahy (2023) slides task, while 4-year-olds outperform 3-year-olds in the inverted y-shaped tube task of Experiment 2 (control condition) and Redshaw and Suddendorf (2016). Why would a large majority of 4-year-olds demonstrate conceptual competence with representing mutual-exclusive possibility when the inverted y-shaped tube is presented on its own, but fewer 4-year-olds show that competence (and look more like 3-year-olds) when the inverted y-shaped slide is contrasted with a straight slide? We speculate that this is because children's performance on these tasks is limited less by representational competence with possibility and more by planning competence. In the more complex two-marble, two-slide scenario, the demands of selecting the correct action in the face of two physical dropping events may have posed a greater challenge for both age groups.

Our results also are consistent with recent work with chimpanzees by Engelmann and colleagues (2023). Previous work had shown that chimpanzees typically approach the inverted y-shaped tube task using a suboptimal strategy, similar to young children (Redshaw & Suddendorf, 2016), which had been interpreted as a lack of the ability to represent multiple, mutually exclusive possibilities. However, Engelmann and colleagues found that, when chimpanzees were competing with a human for food rewards, they showed behaviors consistent with preparing for both possible outcomes of an object dropped into the top of the inverted y-shaped tube. The competition scenario, which is highly ecologically relevant for chimps, may have supported their ability to generate and select the optimal plan in an uncertain situation.

While we argue that the ability to represent possible actions is a primary limiting factor on children's ability to make future plans under uncertainty, we do not argue that it is the *only* limiting factor. A whole range of cognitive capacities are undergoing significant developmental change between the ages of 3 and 5 which may play a role in future-oriented thinking and planning, including executive functions and episodic memory (e.g., Scarf, Gross, Colombo, & Hayne, 2013; Zelazo et al., 2003), and which may separately or together support thinking and planning beyond generating representations of possible actions. We also do not argue that our results show that *all* 3-year-olds have an adult-like concept of possibility (although we believe our results do show that they are more competent at representing possibility than previously thought). It is possible that 3-year-olds' representations of possibility may be developing in a graded fashion, rather than an all-or-nothing fashion. For example, competence with recognizing mutual exclusivity (as in, e.g. Alderete & Xu, 2023; Cesana-Arlotti et al., 2020, Cesana-Arlotti et al., 2022) may develop earlier and potentially support conceptual competence with representing possibility in tasks that require children to make more sophisticated inferences (i.e. about things that have not happened yet). Future work would further examine the sources of limitation and developmental change in children's ability to think about and plan for future uncertain possibilities.

Our results have implications for how we think about constructing tasks to examine young children's conceptual competence. While a variety of tasks have been deployed to examine possibility concepts in recent years, all of these tasks make demands on other still-developing cognitive processes, including future-oriented thinking and planning, which undergoes protracted development throughout the preschool and early school years (see Phillips & Kratzer, 2022, for related discussion). We argue that careful manipulation of these demands can reveal a more nuanced picture of children's early competence with applying and using what they know about possibility to achieve behavioral goals, and can provide insights into how cognitive systems work together in early development. Our results also suggest that targeting action representation and action planning may be a fruitful avenue for intervention for researchers who are interested in improving younger children's ability to take actions under uncertainty.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105712.

Author note

Data and shareable materials are available on the Open Science Framework (https://osf.io/5k9ms/).

CRediT authorship contribution statement

Esra Nur Turan-Küçük: Conceptualization, Formal analysis, Funding acquisition, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Melissa M. Kibbe:** Conceptualization, Formal analysis, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Data are available at https://osf.io/5k9ms/

Acknowledgments

We are grateful to Eyüp Engin Küçük, Nadide Hazal Avcı, and Milagros Herrero for their assistance in creating the stimuli, and Becki Kipling, Rachel Fyler, and the rest of the Museum of Science Living Laboratories team for their support. This research was supported by a Boston University Clara Mayo Memorial Research Fellowship from the Department of Psychological and Brain Sciences, Boston University.

References

Alderete, S., & Xu, F. (2023). Three-year-old children's reasoning about possibilities. *Cognition, 237*, Article 105472.
Atance, C. M. (2015). Young children's thinking about the future. *Child Development Perspectives, 9*(3), 178–182.
Atance, C. M., Ayson, G., & Martin-Ordas, G. (2023). Moving beyond "spoon" tasks: When do children autocue their episodic future thought?. In *Wiley interdisciplinary reviews: Cognitive science* (p. e1646).
Atance, C. M., Louw, A., & Clayton, N. S. (2015). Thinking ahead about where something is needed: New insights about episodic foresight in preschoolers. *Journal of Experimental Child Psychology, 129*, 98–109.
Atance, C. M., & O'Neill, D. K. (2005). Preschoolers' talk about future situations. *First Language, 25*(1), 5–18.
Beck, S. R., Robinson, E. J., Carroll, D. J., & Apperly, I. A. (2006). Children's thinking about counterfactuals and future hypotheticals as possibilities. *Child Development, 77* (2), 413–426.
Blankenship, T. L., & Kibbe, M. M. (2019). Examining the limits of memory-guided planning in 3-and 4-year olds. *Cognitive Development, 52*, Article 100820.
Blankenship, T. L., & Kibbe, M. M. (2022). Two-year-olds use past memories to accomplish novel goals. *Journal of Experimental Child Psychology, 214*, Article 105286.
Blankenship, T. L., & Kibbe, M. M. (2023). "Plan chunking" expands 3-year-olds' ability to complete multiple-step plans. *Child Development,, 94*(5), 1330–1339.
Bulley, A., Henry, J., & Suddendorf, T. (2016). Prospection and the present moment: The role of episodic foresight in intertemporal choices between immediate and delayed rewards. *Review of General Psychology, 20*(1), 29–47.
Bulley, A., Redshaw, J., & Suddendorf, T. (2020). 26 the future-directed functions of the imagination: From prediction to Metaforesight. In *, 425. The Cambridge handbook of the imagination*.
Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development, 20*(3), 362–372.
Cesana-Arlotti, N., Kovács, Á. M., & Téglás, E. (2020). Infants recruit logic to learn about the social world. *Nature Communications, 11*(1), 5999.
Cesana-Arlotti, N., Varga, B., & Téglás, E. (2022). The pupillometry of the possible: An investigation of infants' representation of alternative possibilities. *Philosophical Transactions of the Royal Society B, 377*(1866), 20210343.
Cheng, S., Werning, M., & Suddendorf, T. (2016). Dissociating memory traces and scenario construction in mental time travel. *Neuroscience & Biobehavioral Reviews, 60*, 82–89.
Gautam, S., Suddendorf, T., & Redshaw, J. (2021). When can young children reason about an exclusive disjunction? A follow up to. *Cognition, 207*, Article 104507.
Goddu, M. K., Sullivan, J. N., & Walker, C. M. (2021). Toddlers learn and flexibly apply multiple possibilities. *Child Development, 92*(6), 2244–2251.
Harris, P. L. (2022). Young children share imagined possibilities: Evidence for an early-emerging human competence. *Philosophical Transactions of the Royal Society B, 377* (1866), 20220022.
Leahy, B. (2023). Don't you see the possibilities? Young preschoolers may lack possibility concepts. *Developmental Science, e13400*.
Leahy, B. (2024). Many preschoolers do not distinguish the possible from the impossible in a marble-catching task. *Journal of Experimental Child Psychology, 238*, Article 105794.
Leahy, B., Huemer, M., Steele, M., Alderete, S., & Carey, S. (2022). Minimal representations of possibility at age 3. *Proceedings of the National Academy of Sciences, 119*(52), Article e2207499119.
Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences, 24*(1), 65–78.
Martin-Ordas, G. (2018). "First, I will get the marbles." Children's foresight abilities in a modified spoon task. *Cognitive Development, 45*, 152–161.
Mody, S., & Carey, S. (2016). The emergence of reasoning by the disjunctive syllogism in early childhood. *Cognition, 154*, 40–48.
Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development, 11*(3), 191–214.
Phillips, J., Morris, A., & Cushman, F. (2019). How we know what not to think. *Trends in Cognitive Sciences, 23*(12), 1026–1040.
Phillips, J. S., & Kratzer, A. (2022). Decomposing modal thought. https://psyarxiv.com/g6tzc/.
Pomiechowska, B., Bródy, G., Csibra, G., & Gliga, T. (2021). Twelve-month-olds disambiguate new words using mutual-exclusivity inferences. *Cognition, 213*, Article 104691.
Prabhakar, J., & Ghetti, S. (2020). Connecting the dots between past and future: Constraints in episodic future thinking in early childhood. *Child Development, 91*(2), e315–e330.
Prabhakar, J., & Hudson, J. A. (2014). The development of future thinking: Young children's ability to construct event sequences to achieve future goals. *Journal of Experimental Child Psychology, 127*, 95–109.
Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: Developing a sense of "nearest possible world". *Child Development, 81*(1), 376–389.
Redshaw, J., & Ganea, P. A. (2022). Thinking about possibilities: Mechanisms, ontogeny, functions and phylogeny. *Philosophical Transactions of the Royal Society B, 377*(1866), 20210333.
Redshaw, J., Leamy, T., Pincus, P., & Suddendorf, T. (2018). Young children's capacity to imagine and prepare for certain and uncertain future outcomes. *PLoS One, 13*(9), Article e0202606.
Redshaw, J., & Suddendorf, T. (2016). Children's and apes' preparatory responses to two mutually exclusive possibilities. *Current Biology, 26*(13), 1758–1762.
Redshaw, J., & Suddendorf, T. (2020). Temporal junctures in the mind. *Trends in Cognitive Sciences, 24*(1), 52–64.
Redshaw, J., Suddendorf, T., Neldner, K., Wilks, M., Tomaselli, K., Mushin, I., & Nielsen, M. (2019). Young children from three diverse cultures spontaneously and consistently prepare for alternative future possibilities. *Child Development, 90*(1), 51–61.
Robinson, E. J., Rowley, M. G., Beck, S. R., Carroll, D. J., & Apperly, I. A. (2006). Children's sensitivity to their own relative ignorance: Handling of possibilities under epistemic and physical uncertainty. *Child Development, 77*(6), 1642–1655.
Scarf, D., Gross, J., Colombo, M., & Hayne, H. (2013). To have and to hold: Episodic memory in 3-and 4-year-old children. *Developmental Psychobiology, 55*(2), 125–132.
Shtulman, A., & Phillips, J. (2018). Differentiating "could" from "should": Developmental changes in modal cognition. *Journal of Experimental Child Psychology, 165*, 161–182.
Suddendorf, T. (2017). The emergence of episodic foresight and its consequences. *Child Development Perspectives, 11*(3), 191–195.
Suddendorf, T., & Corballis, M. C. (2008). Episodic memory and mental time travel. In *, 18. Handbook of behavioral neuroscience* (pp. 31–42).
Suddendorf, T., Crimston, J., & Redshaw, J. (2017). Preparatory responses to socially determined, mutually exclusive possibilities in chimpanzees and children. *Biology Letters, 13*(6), 20170170.
Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., et al. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, i–151.