Child Development, 2026, 1–13
https://doi.org/10.1093/chidev/aacaf019
Published: January 29, 2026
Empirical Article

OXFORD Society for Research in Child Development

Back to reality: Children's early temporal reasoning applies to real but not hypothetical events

Urvi Maheshwari* **and David Barner**

Department of Psychology, UC San Diego, La Jolla, CA, United States
*Corresponding author: Department of Psychology, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, United States. Email: umaheshwari@ucsd.edu

Abstract

Time words like "yesterday" and "tomorrow" are hard for children to learn, and for researchers to study, because their referents change from day to day. For example, "yesterday" means something different on Monday and on Wednesday. This study tested 3- and 4-year-old (n = 121; 52% female; no demographic data were collected) US and Canadian children's understanding of "yesterday" and "tomorrow" using three tasks that differed in their reliance on autobiographical and hypothetical events. Results across two experiments conducted between 2023 and 2025 indicated that 3-year-olds comprehend "yesterday" and "tomorrow" when they applied to autobiographical events. However, when asked about hypothetical timelines, even some 4-year-olds struggled to demonstrate knowledge, suggesting that children's early temporal reasoning may be limited to autobiographical events, and does not extend to hypothetical events.

Keywords yesterday, tomorrow, autobiographical memory, spatial timeline, hypothetical reasoning, episodic future thinking

Lay summary

Past research suggests that children learn "yesterday" refers to the past and "tomorrow" refers to the future gradually between 4 and 6 years of age. However, studies that test children's knowledge of temporal language often require them to reason about hypothetical events in time or map time to space. Here, we tested children's comprehension of "yesterday" and "tomorrow" using a hypothetical task and a spatial timeline task as in previous studies, and compared their performance on these tasks to a multi-day autobiographical task in which children had to reason about their own experiences. We found that even 3-year-olds understood the meanings of "yesterday" and "tomorrow" when asked about their own lives, but struggled to express this knowledge on other tasks.

A unique challenge to children learning temporal language is that, unlike words for objects, actions, colors, or numbers, the referents of time words are highly abstract (e.g., an hour), may not yet exist (e.g., tomorrow), or may only be a memory (e.g., yesterday). For example, although it is possible for a child to begin learning a word like "ball" by associating it with objects in their immediate environment, temporal expressions are more challenging because their referents are often fleeting, ethereal, and inaccessible to perception alone. However, as we argue in the present study, the abstract nature of time words—and the fact that the referents of temporal expressions often cover extended periods of time (e.g., last week vs next month)—may actually lead researchers to use proxy tests of temporal reasoning that rely on capacities that are even more sophisticated than reasoning about time, such as hypothetical reasoning. In the present study, we consider the hypothesis that children's temporal reasoning emerges earlier than previously imagined, but is initially limited to real events from their own lives.

While the problem of learning temporal language impacts many linguistic forms, a simple yet instructive example is the contrast between "yesterday" and "tomorrow". These words are interesting both because they are characteristically abstract—their reference shifts according to when they are spoken—but also because children appear to take years to master their meanings, despite beginning to use them early in development. In particular, previous studies find that children begin to produce these words by 2–3 years of age (Ames, 1946; Busby & Suddendorf, 2005; Busby-Grant & Suddendorf, 2011; Suddendorf, 2010) and comprehend them sometime between 3 and 5 years (Busby & Suddendorf, 2005; Busby-Grant & Suddendorf, 2011; Tillman et al., 2018; see also: Friedman, 1990, 1993; Nelson, 1998; Weist, 1989). However, children rarely use "yesterday" and "tomorrow" in an adult-like manner at these ages, with some studies estimating that adult-like comprehension emerges only around 7 to 8 years of age (Ames, 1946; Antinucci & Miller, 1976; Busby-Grant & Suddendorf, 2011; Eisenberg, 1985; Harner, 1981; Nelson, 1998; Szagun, 1978;

Handling Editor: Haley Vlach
Editor: Shauna Cooper
Received: June 18, 2025. **Revised:** August 22, 2025. **Accepted:** October 16, 2025

Tillman et al., 2017; Veneziano & Sinclair, 1995; Weist, 1989; Weist et al., 1991). For example, recent studies find that 3- to 5-year-old children understand neither the temporal remoteness information encoded in these words (that "yesterday" and "tomorrow" denote a period of time 1 day from the present; Tillman et al., 2017; Steele et al., 2025) nor the temporal-causal relations between them (e.g., that events that occur yesterday might cause the events of today and tomorrow to change; Zhang & Hudson, 2018).

A key challenge to the consensus view that time word knowledge emerges slowly is that many measures depend upon hypothetical reasoning capacities that are known to be challenging to young children. For example, in one study, Zhang and Hudson (2018) asked children to match sentences that denoted a past or future action (e.g., "I will open the present tomorrow" or "I opened the present yesterday") to corresponding pictures depicting the present state of an object (e.g., a gift box that has either been opened or remains unopened). The authors found that children failed at this task until they were 5 years old and that they continued to show substantial variability in performance even at this age. Crucially, however, this task requires children to reason about hypothetical points in time that they have not witnessed themselves and would not witness (e.g., a hypothetical opening of a gift in the past). This is important because previous studies demonstrate that children struggle to reason about both hypothetical past events, as in the case of counterfactuals (Leahy et al., 2014; Nyhout et al., 2023; Nyhout & Ganea, 2019; Rafetseder et al., 2013; Robinson & Beck, 2014; c.f., Harris et al., 1996; Weisberg & Gopnik, 2013) and hypothetical future events (Attance and O'Neill, 2001; Beck et al., 2006). Children's hypothetical reasoning abilities develop gradually in early and middle childhood. Given this, it is possible that children are able to reason about and label real temporal events from their own autobiographical timeline but struggle to reason about hypothetical events.

Interestingly, not all previous studies of deictic time words have required children to reason about hypothetical events. In one early study, Harner (1975) tested 2- to 4-year-old children's comprehension of "yesterday" and "tomorrow" by asking them to identify objects that were involved in events from the child's own life, rather than hypothetical events presented in a story. To do so, Harner tested children on two consecutive days. Children were told that on each day they would get to play with one set of toys. On the first day, children played with one set, and on the second day, they were presented with two new sets, and were allowed to play with one of them, leaving a third set that had not been played with. After this, they were asked to identify "yesterday's toys" (toys from day one) and "tomorrow's toys" (the set not previously used by the child). Harner found that while 2-year-olds showed little comprehension of deictic time words, 3-year-olds performed better (especially for "yesterday'), and 4-year-old children understood both terms equally well. Thus, using these methods, Harner found evidence of comprehension of the deictic status of "yesterday" and "tomorrow" up to 1 year earlier than in subsequent studies.

Harner's (1975) study raises the possibility that children have an early-emerging understanding of deictic time words, but that reasoning is restricted to real events from their own autobiographical timelines, and does not extend to hypothetical timelines. However, Harner did not directly compare her method to tasks that rely on hypothetical reasoning, and no subsequent study has done so. Thus, it is also possible that no such advantage for real events exists.[1] Given this, in the present study, we investigated the role of autobiographical experience and hypothetical reasoning in children's early understanding of "yesterday" and "tomorrow". To do so, we administered three tasks. In the first, children were tested with an adaptation of Harner's (1975) two-day task that tested children's application of "yesterday" and "tomorrow" to events from their own timeline (Experiments 1 and 2). In the second, we presented children with stories that introduced hypothetical past and future events that occur in either the life of another child (Experiment 1) or their own life (Experiment 2). Finally, in a third task, we tested children in a task that featured autobiographical events but also required them to map these events onto a spatial timeline—a commonly used metric of temporal reasoning (Experiments 1 and 2; Friedman & Kemp, 1988; Hudson & Mayhew, 2011; Tillman et al., 2017, 2018; Steele et al., 2025). Here, we asked if this additional step of analogically mapping time onto space would impact children's expression of time word knowledge.

Experiment 1

Experiment 1 tested children's comprehension of "yesterday" and "tomorrow" using three tasks that used a parallel structure. First, in a "Real Events" task modeled after Harner (1975), children played with two different sets of toys on consecutive days, and on the second day were asked to identify the toys they played with "yesterday" and those they would play with "tomorrow" among three sets. Second, children completed a novel "Hypothetical Events" task on the second day of testing. In this task, children were told a story about a character playing with different sets of toys every day and were asked to identify the toy associated with "yesterday" and one associated with "tomorrow". The primary difference between the events in the Real Events task and this Hypothetical Events task was that children could draw from memory to identify the toys for yesterday in the former, but had to construct a hypothetical timeline based on a story involving imaginary characters and events in the latter. Third, children completed a "Spatial Timeline Task" on the second of testing, in which they were asked to place the same toys from "yesterday" and for "tomorrow" on a left-to-right spatial timeline, as in Tillman et al. (2017). But unlike Tillman et al., this task tested children on actual, autobiographical events rather than hypothetical events.

Method

All methods and analyses were pre-registered, unless otherwise stated in-text. Data are available on the Open Science Framework (OSF; https://osf.io/wfk25/overview).

Participants

Participants were typically-developing, English-speaking children between the ages of 3;0 and 4;11 years. We recruited 30 3-year-olds (17 F) and 30 4-year-olds (19 F) from daycare centers and preschools in San Diego County between July 2023 and April

[1] Also relevant is that Harner (1975) is a 2 page paper that does not present descriptive statistics, and consequently it is difficult to know the absolute levels of performance of different age groups, and in which cases performance was better than chance, near adult-like, etc.

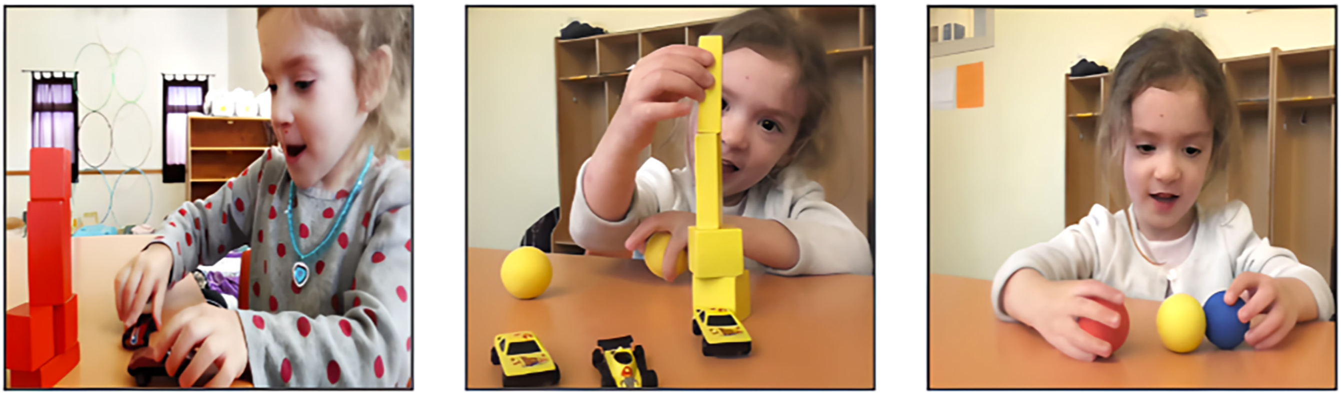

Figure 1 Example of a child playing with toys of a different color on two consecutive days: red on Day 1 (leftmost image), yellow on Day 2 (middle image), and at test on Day 2 (rightmost image), as described in the Real Events Task.

2024. Fourteen additional children were tested on Day 1 but were excluded for not returning for the second day of testing.

The sample size was based on Harner's (1975) original sample (30 participants per age group) and was adequate to achieve 80% power for detecting an effect of $f=0.25$ at $p=.05$ for a Mixed 2 Between (Age: 3-, 4-Years)×3 Within (Task: Real Events, Hypothetical Events, Spatial Timeline) study design, based on an *a priori* power analysis on G*Power version 3.1.9.7 (Faul et al., 2009). Data collection was stopped once we reached our target sample size (n = 30), post-exclusions.

Although individual demographics related to participants' racial/ethnic backgrounds and socioeconomic status (SES) were not collected, children aged 3–10 years in the San Diego County are predominantly Hispanic/Latino (43.3%) and white (37.0%), while other groups include African American/Black (4.8%), Native American (0.4%), Asian (8.9%), Hawaiian/Pacific Islander (0.4%), and Multiracial (5.1%), as per the KidsData Project (2021). The median annual household income in the region is $102,285 (US Census Bureau, 2024).[2]

Materials and procedure

Participants first completed the "Real Events" task over two consecutive days (described below), and on the second day of testing, completed a "Spatial Timeline" task, and a "Hypothetical Events" task.

Real events task

This task, adapted from Harner (1975), was administered over two consecutive days. Children played with one set of toys the first day (and another set the next day). On the second day, they were shown three sets of toys and asked to identify the toys associated with "yesterday" and those associated with "tomorrow". Each set varied in color (i.e., red, yellow, and blue), but contained the same toys: three cars, two balls, and five blocks. On the first day, the experimenter showed the child three bags, saying: "Look! Each of these has toys of a different color. We get to play with toys of a different color every day". The experimenter then chose one of the bags and took out its contents saying, "These are the toys for today!". Children played with toys of one color (e.g., red toys) on the first day. After 3–5 min, the child put the toys back in the bag. On the second day, the experimenter reminded the child of the game while showing them the three bags: "Remember, we get to play with toys of a different color every day! These are the toys for today." The child was then given toys of a different color (e.g., yellow toys) to play. The remaining procedure resembled the first day. After the child put the second day's toys back in the bag, the experimenter told the child that they would return the following day (i.e., Day 3): "When I come back, we get to play with toys of a different color!" The experimenter then took out one toy (e.g., a ball) from each of the three bags and asked the child to identify the toy from yesterday (i.e., the toy they played with the previous day) and the toy for tomorrow (i.e., the toy with which they had not yet played; see Figure 1).

The test questions were, "Show me the toy *from yesterday*/*for tomorrow*," consistent with Harner's (1975) paradigm. The test question was consistent across tasks. Questions were presented across six trials with each toy (i.e., ball, block, or car) presented twice. Children were asked to identify the toy associated with "yesterday" on half the trials and the toy associated with "tomorrow" on the other half. This meant that there was only one correct response (i.e., either "yesterday" or "tomorrow') on each trial. Children always saw red toys on Day 1 (i.e., yesterday's toys), yellow toys on Day 2 (i.e., today's toys), and blue toys were reserved for Day 3 and presented only at test (i.e., tomorrow's toys). The order in which children were asked about a lexical item (i.e., yesterday-first or tomorrow-first) was counterbalanced across participants, and the sets of toys were also presented in a pseudo-randomized order.

Spatial timeline task

In this task, adapted from Tillman et al. (2017), children were asked to place the same toys used in the Real Events task on a left-to-right spatial timeline that extended from infancy to adulthood. Children were familiarized with the timeline as follows, using the same language employed in Tillman et al.'s study: "Look, this is a timeline. It shows when different things happen. The line starts in the past [Experimenter (E) pointed to the left endpoint] and it goes to the future [E traced the line with her finger, ending on the right endpoint]. So, it goes from when you were a baby [E to pointed to left endpoint] all the way to when you're going to be a grown up [E gestured along line to right endpoint]. And here in the middle is right now [E pointed to vertical line at midpoint]. Each

2 We do not collect demographic data related to race/ethnicity and SES to encourage participation from local underrepresented groups who may not be comfortable declaring this information to outside organizations, but may otherwise be interested in participating in our studies.

time has its own place on the line. You're going to show me when different things happen by showing me where they go on the line. Look, when you were a baby goes here [E drew a vertical line on the left endpoint to demonstrate the procedure] and when you are going to be a grown up goes here [E drew a vertical line at right endpoint]. And right now goes here [E drew line at midpoint]. I'm going to give you a crayon, and your job will be to draw an up-and-down line to show me where each thing goes. Ready?" Once the child was ready to proceed, the experimenter introduced one of the toys used in the Real Events task and said, "Remember how we play with different toys every day? We are here [E drew a line in the middle of the timeline]. This [E placed the toy in front of the child] is the toy (e.g., ball) from yesterday/for tomorrow. Can you show me where it goes on the timeline?" If a child marked the toys from yesterday (i.e., the red toys) to the left of the experimenter's line, and the toys for tomorrow (i.e., the blue toys) to the right of the experimenter's line, the trial was marked as correct (1). If the child made a marking on the opposite side or scribbled elsewhere on the page (e.g., drew a horizontal line across the timeline, made a line over the experimenter's line marking "today'), the trial was marked as incorrect (0). See Figure 2 for examples of how children's timelines were coded. The toys were introduced in a pseudo randomized order, and a new timeline was used to mark each type of toy (i.e., block, ball, or car). The order in which toys were introduced (yesterday-first or tomorrow-first) was also pseudo randomized for each yesterday/tomorrow pair.

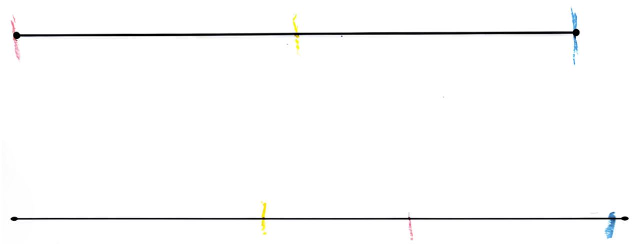

Figure 2 Top: Example of a child's timeline in which the line marking yesterday (i.e., for the red toy) is correctly placed to the left of the experimenter's (yellow) line (marked as 1), and the line for tomorrow (i.e., for the blue toy) is correctly placed to the right of the experimenter's line (marked as 1). Bottom: Example of a timeline on which the line marking yesterday (i.e., for the red toy) is incorrectly placed to the right of the experimenter's (yellow) line (marked as 0), but the line for tomorrow (i.e., for the blue toy) is correctly placed to the right of the experimenter's line (marked as 1).

Hypothetical Events task

After completing the spatial timeline task, children completed a Hypothetical Events task, wherein they were told a story about a character playing with a different toy every day. They were then asked to identify the toy associated with "yesterday" and one associated with "tomorrow" (Figure 3). The test question was the same as the Real Events task. A total of six stories were presented: three stories asked children to identify the toy associated with "yesterday" and three asked about the toy associated with "tomorrow". As in the Real Events task, this meant that there was only one correct response (i.e., either "yesterday" or "tomorrow') on each trial. The order in which children were asked about a lexical item (i.e., yesterday-first or tomorrow-first) was counterbalanced across participants, and questions about "yesterday" and "tomorrow" were alternated.

Results

Our primary analyses were based on a pre-registered plan available on OSF, reported below. Additional exploratory analyses are indicated below, or in supplementary materials. All model comparisons were performed using likelihood ratio tests, and best fitting models were selected on the basis of a significant chi-squared statistic and reduced AIC value. All pairwise comparisons were adjusted for multiple comparisons using a Bonferroni correction. All analyses were conducted in R studio 4.5.1., primarily using the lme4 (Bates et al., 2015), emmeans (Lenth, 2025), and tidyverse (Wickham et al., 2019) packages.

To test whether children's deictic time knowledge differed based on task, we constructed a base generalized linear mixed model with Age (3-/4-year-olds) and Task as fixed factors, and participant as a random effect, with Accuracy as the dependent variable. Age was a significant predictor of accuracy ($B = 0.93$, $p = 0.001$), and 4-year-olds demonstrated greater overall accuracy on each task than 3-year-olds. Children were significantly more accurate on the Real Events task relative to the Spatial Timeline ($B = -0.54$, $p = .01$) and Hypothetical Events task ($B = -0.76$, $p < .001$). In addition, children also performed better on the Spatial Timeline task than the Hypothetical Events task ($B = 0.46$, $p = .01$) overall. *Post hoc* pairwise comparisons found that 3-year-olds ($B = 0.54$, $p = .01$) but not 4-year-olds ($B = 0.52$, *n.s.*) were more accurate on the Real Events task compared to the Spatial Timeline task. Both 3-year-olds ($B = 0.76$, $p = .002$) and 4-year-olds ($B = 1.21$, $p < .001$) were more accurate on the Real Events task relative to the Hypothetical Events task. Finally, we found no differences in 3-year-olds ($B = 0.22$, *n.s.*) performance on the Spatial Timeline task compared to the Hypothetical Events task, though 4-year-olds were significantly better ($B = 0.69$, $p = .006$) on the Spatial Timeline task. These data not only suggest that children acquire temporal concepts before mapping these concepts to space but also that reasoning about hypothetical events may be harder than reasoning about autobiographical ones (Figure 4).

To determine whether there were differences in children's knowledge of the two words (yesterday/tomorrow) across tasks, we added an interaction term to the base model using the following formula: Accuracy ~ Age * Task * Item + (1|PID). However, Item was not a significant predictor of accuracy ($B = -0.02$, *n.s.*) and adding a Task × Item interaction term did not significantly improve model fit ($\chi^2(6) = 4.07$, *n.s*), suggesting that there were no significant differences in children's performance on "yesterday" relative to "tomorrow" trials (Figure S1).

To understand when children begin to comprehend the deictic status of these terms, we supplemented our pre-registered analyses by conducting t-tests examining children's performance to chance.[3] These tests found that 3-year-olds comprehend the deictic status of these time words at above chance levels when tested on the Real Events task ($t(179) = 5.3.$, $p < .001$) and the Spatial Timeline task ($t(173) = 2.088$, $p = .038$), but not the Hypothetical Events task ($t(179) = 0.71$, *n.s.*). By contrast, 4-year-olds performed above chance levels on all tasks [Real Events

[3] Note that we set chance at 33% (⅓) because there are three potential options from which to choose, consistent with Harner's (1975) task. However, it isn't always clear what incorrect response on the Spatial Timeline task is, and we chose to keep chance consistent across tasks.

Figure 3 Example of a story used in the hypothetical events task. Events were presented sequentially from (1) to (6).

($t(179) = 11.71$, $p < .001$), Spatial Timeline ($t(179) = 7.88$, $p < .001$), and Hypothetical Events ($t(178) = 3.57$, $p < .001$)]. See Table 1.

In a final set of exploratory analyses, we examined whether children's performance across tasks was correlated while controlling for age. We found no significant correlations between the Real Events task and the Hypothetical Events task ($r = 0.21$, *n.s.*), the Hypothetical Events task and the Spatial Timeline task ($r = -0.10$, *n.s.*), or the Real Events task and Spatial Timeline task ($r = 0.21$, *n.s.*), providing some evidence that different mechanisms beyond temporal reasoning may be involved in children's reasoning across tasks.

Discussion

In Experiment 1, we examined the developmental trajectory of children's acquisition of the deictic status of "yesterday" and "tomorrow" using tasks that followed the same structure, but relied differently on children's hypothetical reasoning skills and their ability to map time onto space. We found three main results. First, 3-year-olds performed above chance for these deictic time words on the Real Events task, but not on the Spatial Timeline or Hypothetical Events task. Thus, autobiographical events appeared easier for children than hypothetical events, though mapping autobiographical events to space posed an added challenge. Second, we found that children expressed their knowledge more accurately on the Spatial Timeline than the Hypothetical Events task, suggesting that mapping autobiographical experiences to space may be easier than reasoning about events that are entirely hypothetical in nature. Third, although children were above chance on all tasks by age 4, task-related differences in performance persisted, and some 4-year-olds struggled to identify the referents of deictic time words on the Hypothetical Events

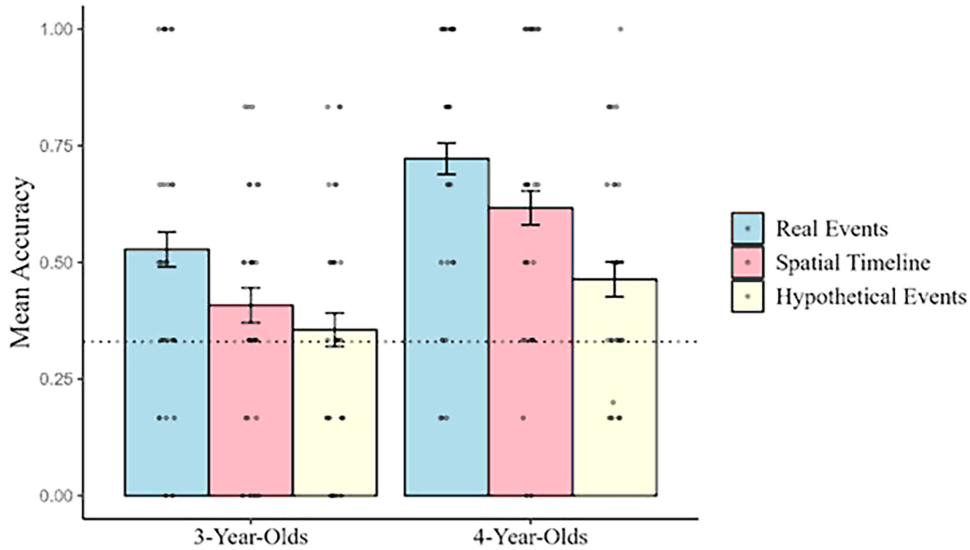

Figure 4 Figure showing overall differences between 3- and 4-year-olds accuracy across the real events, spatial timeline, and hypothetical events tasks. The dotted line represents chance. Error bars represent standard error of the mean. All data visualizations were created using the ggplot2 R package (Wickham, 2016).

Table 1 Children's mean (*standard deviation*) performance on the lexical items "yesterday" and "tomorrow" across tasks.

Age	3-year-olds		4-year-olds	
Task/item	Yesterday M (SD)	Tomorrow M (SD)	Yesterday M (SD)	Tomorrow M (SD)
Real events	0.52 *(0.50)*	0.53 *(0.50)*	0.73 *(0.44)*	0.71 *(0.45)*
Spatial timeline	0.36 *(0.48)*	0.44 *(0.50)*	0.62 *(0.48)*	0.61 *(0.49)*
Hypothetical events	0.41 *(0.49)*	0.30 *(0.46)*	0.46 *(0.50)*	0.46 *(0.50)*

task. These results suggest that children reason about real, autobiographical, events more easily than hypothetical events, and that mapping events to space may pose an additional challenge. However, this conclusion is tempered by the fact that the three tasks differed in a variety of ways that are not related to the distinction between autobiographical vs. hypothetical events, or to the problem of mapping events to space. Therefore, in Experiment 2 we sought to replicate the results of Experiment 1, while also addressing these differences.

Experiment 2

We aimed to replicate the findings of Experiment 1 while better accounting for low-level differences between tasks that may have led to differences in children's comprehension of "yesterday" and "tomorrow". We made five modifications in Experiment 2. First, in Experiment 1, the Hypothetical Events task differed from the autobiographical tasks not only in being hypothetical, but also in applying to a third person perspective, rather than to the child's own perspective. Therefore, in Experiment 2 the Hypothetical Events task asked children to imagine that they played with toys on different days, rather than imagining a character doing so. Second, in Experiment 1, toys in the Hypothetical task differed in kind across days rather than in color. Therefore, in Experiment 2 children were asked to imagine toys of different colors (rather than different toys) using the same images as in the other two tasks. Third, in Experiment 1, the Hypothetical Events task was always conducted on Day 2, raising the possibility that performance was poorer due to waning attention. Therefore, in Experiment 2, this task was administered on Day 1. Fourth, in Experiment 2 we placed images of a baby on the left-end of the spatial timeline and an adult on the right-end, to serve as anchors and minimize working memory demands for making left-to-right mappings between time and space. Finally, in Experiment 2, the experimenter explicitly labeled the toys from Day 2 as "today's' toys, so that their comprehension of "yesterday" and "tomorrow" was reduced to a clear two-alternative forced choice, with chance defined as 50% on all tasks (see footnote 2 for discussion of why assuming a chance rate of 33% in Experiment 1 may have overestimated children's knowledge).

In making these changes, we reasoned that if children can apply "yesterday" and "tomorrow" to real events more readily than to hypothetical events, then the results of Experiment 1 should replicate when low-level differences are removed. In particular, children should demonstrate earlier comprehension of "yesterday" relative to "tomorrow" on the Real Events task relative to the other tasks. However, if low-level task features explain differences between children's comprehension in these tasks, then they may perform similarly across all tasks.

Method

All methods and analyses were pre-registered, unless otherwise stated in-text. Data are available on OSF (https://osf.io/wfk25/?view_only=a528b92bd0194051bd3bddaec911b4f3).

Participants

Participants were typically-developing, native English-speaking children between the ages of 3;0 and 4;11 years. We recruited 30 3-year-olds (12 F) and 31 4-year-olds (15 F) from daycare centers, preschools, and parks in Comox Valley, British Columbia (n = 23) and San Diego County, California (n = 38) between March and June 2025. Thirteen additional children were tested on Day 1, but were excluded for not returning for the second day of testing. The sample size was kept the same as Experiment 1, and data collection was stopped once we reached our target sample size (n = 30 per age group) post-exclusions. Data from one additional 4-year-old was collected during a testing session and retained.

As in Experiment 1, we did not collect individual demographics related to participants' race/ethnicity and socio-economic status. The sample was drawn from the populations of San Diego County (reported in Experiment 1) and the Comox Valley, which is comprised predominantly by people of European origin (87.01%), with others of Indigenous & Métis (7.33%), East Asian (1.73%), South-East Asian (1.48%), South Asian (0.93%), African (0.6%), Latin American (0.37%), and Middle Eastern (0.17%) origin, as per the Canadian Census of Population. The median household income in the region is $40,000 (Statistics Canada, 2021).

Materials and procedure

Participants first completed the Hypothetical Events task on the first day of testing, before being introduced to the Real Events task. The Spatial Timeline task was completed last on Day 2.

Real events task

This task was administered in the same way as Experiment 1, with one modification to the test question to mark "today" to ensure consistency across tasks. The experimenter pointed to the Day 2 toy, saying: "This is the toy from today. Show me the toy *from yesterday/for tomorrow*".

Spatial timeline task

The procedure and coding scheme for the spatial timeline task were the same as in Experiment 1, with one exception: the ends of the timelines were marked with images of a baby on the left-end and an adult on the right end (Figure 5). This task was administered after the Real Events Task on Day 2.

Hypothetical events task

In this instantiation of the Hypothetical Events task, children were told to imagine that they had a lot of toys and that they would play with toys of a different color every day (Figure 6). They were then asked to identify the toy associated with "yesterday" and one associated with "tomorrow", once today's toy was identified by the experimenter. The test question was the same as the Real Events task. A total of six questions were presented: three questions asked children to identify the toy associated with "yesterday" and three asked about the toy associated with "tomorrow". As in the Real Events task, this meant that there was only one correct response (i.e., either "yesterday" or "tomorrow') on each trial. The order in which children were asked about a lexical item (i.e., yesterday-first or tomorrow-first) was counterbalanced across participants, and questions about "yesterday" and "tomorrow" were alternated.

Results

Our primary analyses were based on a pre-registered plan available on OSF and are reported below. Additional exploratory analyses are indicated below, or in supplementary materials. All model comparisons were performed using likelihood ratio tests, and best fitting

Figure 5 Example of a child's timeline in Experiment 2, in which the line marking yesterday (i.e., for the red toy) is correctly placed to the left of the experimenter's (yellow) line (marked as 1), and the line for tomorrow (i.e., for the blue toy) is correctly placed to the right of the experimenter's line (marked as 1).

models were selected on the basis of a significant chi-squared statistic and reduced AIC value. All pairwise comparisons were adjusted for multiple comparisons using a Bonferroni correction.

As in Experiment 1, we first constructed a base generalized linear mixed model to examine whether children's comprehension of "yesterday" and "tomorrow" differed based on Age and Task using the following formula: Accuracy (1/0) ~ Age (3 years/4 years) * Task (Real Events/Spatial Timeline/Hypothetical Events) + (1|Site) + (1|PID). Once again, we found that children performed better on the Real Events task relative to the Spatial Timeline task ($B = -0.89$, $p < .001$) and Hypothetical Events task ($B = -1.13$, $p < .001$). Also, 4-year-olds performed better than 3-year-olds overall ($B = -0.65$, $p = .008$), suggesting that children's comprehension of deictic time words improved with age. *Post hoc* pairwise comparisons found that both 3-year-olds ($B = 0.89$, $p < .001$) and 4-year-olds ($B = 1.02$, $p < .001$) performed better on the Real Events task than the Spatial Timeline task. In addition, both 3-year-olds ($B = 1.13$, $p < .001$), and 4-year-olds ($B = 0.97$, $p < .001$) were more accurate on the Real Events task relative to the Hypothetical Events task, replicating the results of Experiment 1. Interestingly, unlike Experiment 1, there were no significant differences in 3-year-olds' ($B = 0.23$, *n.s.*) or 4-year-olds' ($B = -0.05$, *n.s.*) performance on the Spatial Timeline task relative to the Hypothetical Events task, again suggest that children might comprehend "yesterday" and "tomorrow" better for autobiographical events than hypothetical ones but that the additional problem of mapping events to space may impede expression of children's knowledge of time. Taken together, these results suggest that 3-year-olds can reason about the deictic status of time words when they apply to real, autobiographical events, but that they struggle to reason about hypothetical events, and to express autobiographical knowledge in a spatial timeline task (Figure 7).

As in Experiment 1, we examined differences in children's comprehension of "yesterday" and "tomorrow" by constructing a generalized linear mixed model with the following formula: Accuracy (1/0) ~ Age * Task * Item + (1|Site) + (1|PID), which significantly improved fit relative to the base model ($\chi^2(6) = 16.54$, $p = .01$). However, the fixed term for Item (yesterday/tomorrow) was not a significant predictor of accuracy ($B = -0.55$, *n.s.*). *Post hoc* pairwise comparisons revealed that children's comprehension of "yesterday" and "tomorrow" only differed on the Spatial Timeline task among 3-year-olds ($B = -1.16$, $p < .001$), but no other differences in comprehension were found for any task at any age (Figure S2).

Also as in Experiment 1, we conducted t-tests against chance to examine the age at which children comprehended deictic time words across tasks. We set chance to 50% in this Experiment (diverging from Experiment 1, where chance was 33%), because the experimenter always identified "today", leaving the items corresponding with "yesterday" and "tomorrow" as the only options to consider. This was done not only to make the potential

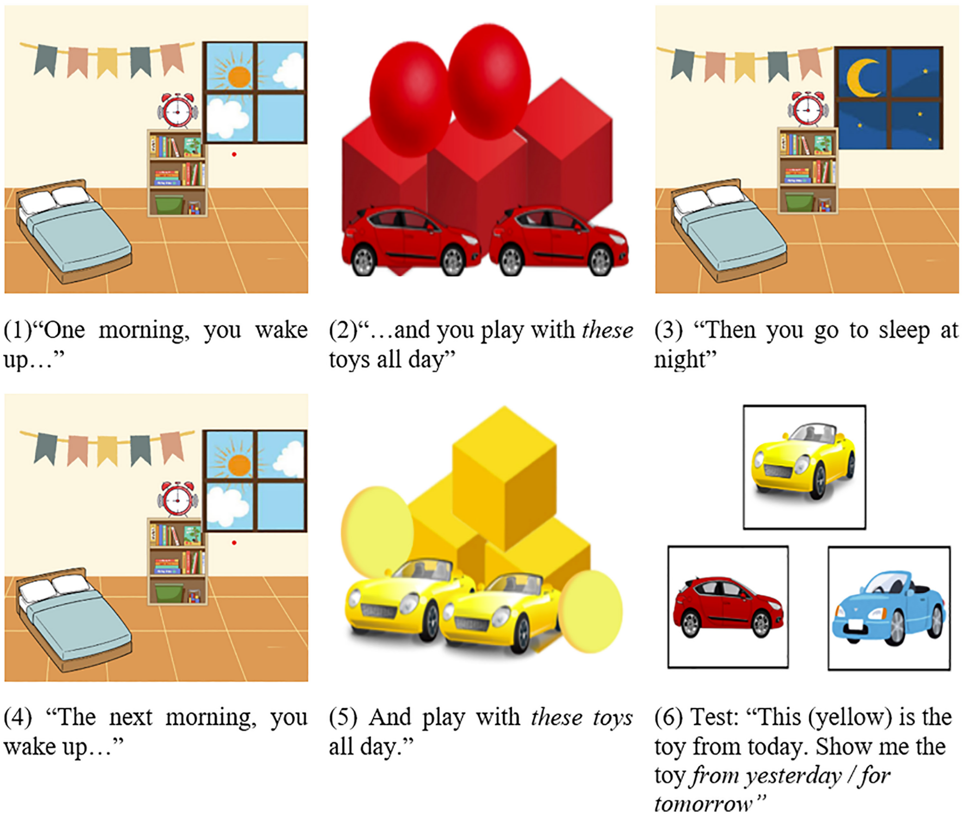

Figure 6 Vignette used in the hypothetical events task. Events were presented sequentially from (1) to (5), followed by the test trials (6).

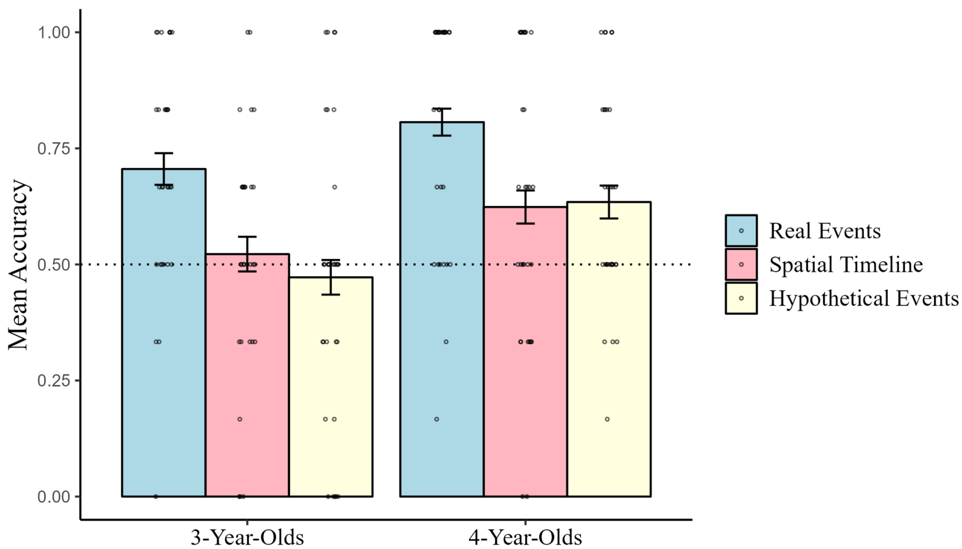

Figure 7 Bar graph showing overall differences between 3- and 4-year-olds accuracy across the real events, spatial timeline, and hypothetical events tasks. The dotted line represents chance (set to 50%). Error bars represent standard error of the mean.

memory load easier (since children did not need to recall which were the toys from "today') but also to ensure that chance was defined similarly across tasks. We found, once again, that 3-year-olds identified "yesterday" and "tomorrow" accurately on the Real Events task at above chance levels ($t(179) = 6.03$, $p < .001$), but not on the Spatial Timeline task ($t(179) = 0.59$, *n.s.*) or the Hypothetical Events task ($t(179) = -0.74$, *n.s.*). By age 4, children performed above chance on the Real Events task ($t(185) = 10.55$, $p < .001$), but also on the Spatial Timeline ($t(185) = 3.47$ $p < .001$) and Hypothetical Events ($t(185) = 3.79$ $p < .001$) tasks. We therefore replicated the pattern of findings in Experiment 1 (Table 2).

Finally, in exploratory analyses, we examined whether children who performed well on one task were also more likely to perform well on another. To do so, we examined whether children's performance across tasks was correlated while accounting for age. We found a small but significant correlation between the Spatial

Table 2 Children's mean (standard deviation) performance on the lexical items "yesterday" and "tomorrow" across tasks.

Age	3-year-olds		4-year-olds	
Task/item	Yesterday M (SD)	Tomorrow M (SD)	Yesterday M (SD)	Tomorrow M (SD)
Real events	0.75 (*0.43*)	0.65 (*0.47*)	0.82 (*0.37*)	0.78 (*0.41*)
Spatial timeline	0.40 (*0.49*)	0.64 (*0.48*)	0.60 (*0.49*)	0.64 (*0.48*)
Hypothetical events	0.47 (*0.50*)	0.46 (*0.50*)	0.65 (*0.47*)	0.61 (*0.48*)

Timeline and Real Events tasks (r = 0.28, p = .02). However, we found no significant correlation between the Real Events task and the Hypothetical Events task (r = 0.24, *n.s.*). Finally, the Spatial Timeline and Hypothetical Events tasks were also not related (r = 0.23, p = .07), suggesting once again that children may rely on different cognitive abilities beyond temporal reasoning when completing these tasks, which may underestimate their comprehension of temporal language.

Discussion

Experiment 2 found three main results, replicating Experiment 1. First, as in Experiment 1, 3-year-olds performed above chance for "yesterday" and "tomorrow" when tested in the Real Events task but not on the two other tasks. Second, we no longer found any differences in children's accuracy on the Spatial Timeline and Hypothetical Events tasks, suggesting that any advantage children have for reasoning about real, autobiographical events is reduced by the requirement to map these events to a spatial timeline, consistent with past reports that 3-year-olds struggle to map space to time (Tillman et al., 2017, 2018; Steele et al., 2025). Third, we again found that children showed some success on the Spatial Timeline and Hypothetical Events tasks by age 4, but that 4-year-olds remained better at reasoning about autobiographical events on the Real Events task. These results suggest that, initially, children's temporal reasoning is restricted to real, rather than hypothetical events, but also that this knowledge may be partially obscured by measures that require mapping time to space.

General discussion

The current study tested children's comprehension of "yesterday" and "tomorrow" on three tasks that relied differently on their autobiographical experience and hypothetical reasoning skills across two experiments. We found three main results. First, both 3– and 4-year-olds succeeded on the Real Events task, suggesting that children understand the tense information encoded by these deictic time words by age 3. This result is compatible with Harner (1975) and is earlier than reported by other investigations of children's deictic time word comprehension (e.g., Tillman et al., 2017, 2018; Steele et al., 2025; Zhang & Hudson, 2018). Second, we found that children performed better on tasks that required them to reason about autobiographical events (i.e., the Real Events task), than on a task that required them to reason about hypothetical events (i.e., the Hypothetical Events task). Third, both experiments found that children performed worse in an autobiographical task when it required mapping events onto a spatial timeline, consistent with previous studies which find that children struggle to map time onto space before age 4 (Tillman et al., 2017, 2018; Steele et al., 2025). Altogether, these results suggest that children find it easier to reason about time when it relates to their own autobiographical experiences, and that tasks which rely on reasoning about hypothetical events or mappings to space may underestimate children's representations of temporal language.

These results are important theoretically, because they suggest that early in life children's temporal thought favors reasoning about real, autobiographical, events over hypothetical events created in the imagination (for discussion, see McCormack & Hoerl, 2017; Nelson, 1998). In particular, they are consistent with previous reports that children reason more readily about real past events relative to counterfactual scenarios, and that children often struggle with hypothetical reasoning more generally, including reasoning about hypothetical future scenarios, or situations involving multiple mutually exclusive possible outcomes (Attance & O'Neill, 2001; Beck et al., 2006; Beck & Riggs, 2014; Leahy et al., 2014; Nyhout et al., 2023; Nyhout & Ganea, 2019; Rafetseder et al., 2013; Robinson & Beck, 2014; Turan-Küçük & Kibbe, 2024; 2025; Tillman & Walker, 2022; c.f., Harris et al., 1996; Weisberg & Gopnik, 2013). However, many questions remain about why reasoning about hypothetical timelines is more difficult for children. One potential factor relates to time itself: in the Real Events task, events from "yesterday" were actually in the past a day earlier, while events from "today" were in the present (or immediate past), and events for "tomorrow" had not yet occurred. By contrast, in the Hypothetical Events task, all three events occurred on the same day, in the present (immediate past). Given this, if children understand "yesterday" to refer to the day before "today" it may make little sense to apply it to events that actually occurred today. Another factor relates to the modal component of hypothetical reasoning, and the requirement that children imagine worlds that do not actually exist. Much like in the case of counterfactual reasoning, when children imagine a hypothetical timeline they are required to represent a past that does not actually exist, and to use this imaginary timeline to formulate predictions about the present and the future. Our data indicate that, although the future is always unknown—whether when reasoning about hypothetical or autobiographical timelines—children find it easier to predict and describe the future when it is situated on a timeline composed of actual, attested, events. What remains unsettled by our data is whether these difficulties reflect early limitations in processing capacity—e.g., due to the greater difficulty of holding in mind uncertain, hypothetical events—or instead reflect a conceptual change, wherein children slowly construct representational resources that allow higher-order meta-cognitive representations of time, alternative timelines, and recursive temporal relations (see for discussion: Gautam et al., 2019; McCormack & Hoerl, 2017; Redshaw, 2014, 2024).

Aside from suggesting that early temporal reasoning may favor real, autobiographical events, this work is also relevant to theoretical debates regarding the emergence of deictic expressions in particular. Although previous studies find that 3-year-olds can describe past and future events and use tense information in relation to the past and future (e.g., Harner, 1980; Hayne et al., 2011; Peterson, 2002; Peterson & Rideout, 1998; Quon & Atance, 2010; Weist & Zevenbergen, 2008), some have suggested that children this age may only represent the past and future status of the events but not where these events are located in time (e.g., Hoerl & McCormack, 2019; McCormack & Hanley, 2011; McCormack & Hoerl, 2017; Nelson, 1998). For example, Hoerl and McCormack (2019) cite children's failure to comprehend "yesterday" and "tomorrow" at age 3 as evidence that they fail to grasp "the systematic temporal relations that obtain between these events" (p. 11). Our study, however, suggests that by 3 years of age, children can use "yesterday" and "tomorrow" deictically to identify events before and after "today", and at least in this sense are able to reason about temporal relations between past, present, and future events denoted by these words.

Relatedly, our study suggests that although children are able to situate autobiographical events in time, they find it challenging to express this knowledge in a spatial format. Whereas children performed better than chance on the Real Events task at age 3, they did not reliably map these events to space on the Spatial Timeline task until age 4. This finding is compatible with results of previous studies, which report that children only map "yesterday" and "tomorrow" onto a spatial timeline sometime between 4 and 6 years of age. For example, in one study Tillman et al. (2017) asked children to map expressions like "breakfast yesterday" onto a timeline and found that while 3-year-olds performed at chance levels, 4-year-olds were above chance and only became adult-like around 7 to 8 years of age. In another study, Tillman et al. (2018) asked children to place "yesterday" and "tomorrow" in relation to "today" in a modified version of Tversky et al.'s (1991) sticker task, which also involved organizing events in space. Again, 4-year-old children were able to map deictic words onto a timeline when primed with a prior spatial mapping task, but unlike 5– and 6-year-olds, did not map time words to linear space spontaneously without prior priming. Finally, Steele et al. (2025) asked children to locate temporal terms like "yesterday" and "tomorrow", as well as the "day before yesterday" and "day after tomorrow" on a calendar-like template, wherein individual days were represented by individual squares of a calendar. Children once again had to place these terms in relation to "today", and only began to do so successfully at after age 4. In sum, these findings suggest that mapping time onto space may pose a challenge to young children, and likely emerges sometime after they label events in time with deictic expressions.

In addition to addressing the broad nature of children's reasoning about deictic time, our study also addressed the order in which children learn the deictic status of "yesterday" and "tomorrow". In our study, we found no significant difference in children's performance on "yesterday" relative to "tomorrow" in the three tasks. In particular, we found no significant differences in 3-year-olds' acquisition of these terms, which is somewhat at odds with Harner's (1975) finding that 3-year-olds understood "yesterday" better than "tomorrow". Although the literature on temporal language acquisition suggests that children generally comprehend language that refers to the past before they understand descriptions of future (e.g., Clark, 1973; Cromer, 1971; Harner, 1975, 1976; Herriot, 1969; Zhang & Hudson, 2018), this isn't always the case in studies of deictic time words. For instance, some studies find no differences in children's acquisition of these words (e.g., Tillman et al., 2017; Steele et al., 2025; c.f. Maheshwari & Barner, 2025), and some production studies have even found that children begin to produce "tomorrow" earlier than "yesterday" (Busby & Suddendorf, 2005; Busby-Grant & Suddendorf, 2011; Pawlak et al., 2006). Previous studies suggest that this asymmetry in children's acquisition may exist, either because children can rely on episodic memory for past events and do not have similar memory traces for future events (e.g., Friedman, 2000, 2002; Hinrichs, 1970; Zhang & Hudson, 2018), or alternatively, because children rely on syntactic cues (e.g., tense markings) to narrow the meanings of deictic time words (e.g., Gillette et al., 1999; Gleitman, 1990; Maheshwari & Barner, 2025; Tillman et al., 2017; Steele et al., 2025), and while English has a past tense marking, it does not have a similarly robust future tense marking (Shirai & Miyata, 2006; Weist, 1989). Teasing apart these factors in children's acquisition of "yesterday" relative to "tomorrow" should be the focus of future research.

Finally, before concluding, it is important to note that this work had a number of limitations that should be addressed in future work. First, in this study we tested only two time words, and limited our investigation to children's understanding of their deictic status—that yesterday is in the past and tomorrow is in the future. Notably, previous studies (Tillman et al., 2017; Steele et al., 2025; Zhang & Hudson, 2018) also measured other information that is encoded by these words (e.g., temporal remoteness, causality). A second limitation is that although we accounted for several lower-level features that might explain differences in Experiment 2, other confounds such as familiarity with or preference for particular colors in the paradigm, or memory of real toys relative to hypothetical ones might still account for some differences in children's comprehension.[4] A third limitation, alluded to above, is that while we found that children have difficulty with a task that involves hypothetical events, we didn't resolve why reasoning about such events is difficult. One possibility is that hypothetical events involve a form of modal reasoning that is still developing at age 3, and possibly requires the construction of meta-cognitive resources that are absent earlier in life (Amsel et al., 2005; Buchanan & Sobel, 2011; Gautam et al., 2019; Kuczaj & Daly, 1979; Leahy & Carey, 2020; Nyhout et al., 2023; Nyhout & Ganea, 2019; Weisberg & Gopnik, 2013). Another possibility is that hypothetical events pose a problem because they pose greater demands on other capacities, such as working memory and cognitive control (e.g., Arterberry & Albright, 2020; Buchsbaum et al., 2012; Beck et al., 2006; Beck & Riggs, 2014; Bettencourt, 2022; German & Nichols, 2003; Lohse et al., 2015; McCormack et al., 2018). Finally, it remains uncertain whether autobiographical events are easier because they occur on a real timeline (the past is in the past), because they are factual (rather than hypothetical), or because they are experienced more directly, resulting in richer representations. Future research should disentangle these different questions, to resolve both how hypothetical reasoning about deictic time emerges, and which factors make it challenging for children.

In summary, we find that 3-year-old children understand the deictic status of time words, and can reason about autobiographical events, but that they struggle to reason about hypothetical events, whether from their own perspective, or that of another person. We also find that even at 4 years of age children are more accurate when reasoning about autobiographical events than hypothetical events. Finally, although 3-year-olds can reason

[4] As one reviewer noted, since the toy for yesterday was always red and the toy for tomorrow was always blue, children might succeed on the task if they happened to associate the color red with "yesterday" and blue with "tomorrow," and if these associations may have increased with age.

about the deictic status of time words, they find it more difficult to express this knowledge when it requires mapping time onto space, compatible with previous reports that, at least initially, time is represented independent of space, and only gradually mapped onto it.

Supplementary material

Supplementary material is available at *Child Development* online.

Data availability

Data, code, materials, and pre-registrations for this research are available on the Open Science Framework (https://osf.io/wfk25/?view_only=a528b92bd0194051bd3bddaec911b4f3).

Author contributions

Urvi Maheshwari (Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Visualization, Writing—original draft, Writing—review and editing). David Barner (Conceptualization, Methodology, Project Administration, Resources, Supervision, Writing—original draft, Writing—review and editing)

Funding

This project was supported by an APA Dissertation Research Award to U.M.

Conflicts of interest

None declared.

Acknowledgments

We thank Marianna Thorne and Violet Scholar for their assistance in data collection.

References

Ames, L. B. (1946). The development of the sense of time in the young child. *The Pedagogical Seminary and Journal of Genetic Psychology*, *68*, 97–125. https://doi.org/10.1080/08856559.1946.10533358

Amsel, E., Trionfi, G., & Campbell, R. (2005). Reasoning about make-believe and hypothetical suppositions: Towards a theory of belief-contravening reasoning. *Cognitive Development*, *20*, 545–575. https://doi.org/10.1016/j.cogdev.2005.08.002

Antinucci, F., & Miller, R. (1976). How children talk about what happened. *Journal of Child Language*, *3*, 167–189. https://doi.org/10.1017/S0305000900001434

Arterberry, M. E., & Albright, E. J. (2020). Children's memory for temporal information: The roles of temporal language and executive function. *The Journal of Genetic Psychology*, *181*, 191–205. https://doi.org/10.1080/00221325.2020.1741503

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in cognitive sciences*, *5*, 533–539. https://doi.org/10.1016/S1364-6613(00)01804-0

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using {lme4}. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

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

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*, 413–426. https://doi.org/10.1111/j.1467-8624.2006.00879.x

Bettencourt, K. M. (2022). The relations between temporal memory abilities, time knowledge, executive function, and language in early and middle childhood [Doctoral Dissertation]. The University of North Carolina at Greensboro.

Buchanan, D. W., & Sobel, D. M. (2011). Mechanism-based causal reasoning in young children. *Child development*, *82*, 2053–2066. https://doi.org/10.1111/j.1467-8624.2011.01646.x

Busby, J., & Suddendorf, T. (2005). Recalling yesterday and predicting tomorrow. *Cognitive Development*, *20*, 362–372. https://doi.org/10.1016/j.cogdev.2005.05.002

Busby-Grant, J. B., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly*, *26*, 87–95. https://doi.org/10.1016/j.ecresq.2010.05.002

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

Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and acquisition of language* (pp. 27–63). Academic Press. https://doi.org/10.1016/B978-0-12-505850-6.50008-6

Cromer, R. F. (1971). The development of the ability to decenter in time. *British Journal of psychology*, *62*, 353–365. https://doi.org/10.1111/j.2044-8295.1971.tb02046.x

Eisenberg, A. R. (1985). Learning to describe past experiences in conversation∗. *Discourse Processes*, *8*, 177–204. https://doi.org/10.1080/01638538509544613

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*, 1149–1160.

Friedman, W. J. (1990). *About time: Inventing the fourth dimension*. The MIT press.

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, *113*, 44–66. https://doi.org/10.1037/0033-2909.113.1.44

Friedman, W. J. (2000). The development of children's knowledge of the times of future events. *Child development*, *71*, 913–932. https://doi.org/10.1111/1467-8624.00199

Friedman, W. J. (2002). Children's knowledge of the future distances of daily activities and annual events. *Journal of Cognition and Development*, *3*, 333–356. https://doi.org/10.1207/S15327647JCD0303_4

Friedman, W. J., & Kemp, S. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cognitive Development*, *13*, 335–367. https://doi.org/10.1016/S0885-2014(98)90015-6

Gautam, S., Suddendorf, T., Henry, J. D., & Redshaw, J. (2019). A taxonomy of mental time travel and counterfactual thought:

Insights from cognitive development. *Behavioural Brain Research*, *374*, 112108. https://doi.org/10.1016/j.bbr.2019.112108

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

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*, 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55. https://doi.org/10.1207/s15327817la0101_2

Harner, L. (1975). Yesterday and tomorrow: Development of early understanding of the terms. *Developmental Psychology*, *11*, 864–865. https://doi.org/10.1037/0012-1649.11.6.864

Harner, L. (1976). Children's understanding of linguistic reference to past and future. *Journal of psycholinguistic research*, *5*, 65–84. https://doi.org/10.1007/BF01067949

Harner, L. (1980). Comprehension of past and future reference revisited. *Journal of Experimental Child Psychology*, *29*, 170–182. https://doi.org/10.1016/0022-0965(80)90099-5

Harner, L. (1981). Children talk about the time and aspect of actions. *Child Development*, *52*, 498–506. https://doi.org/10.2307/1129167

Harris, P. L., German, T., & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, *61*, 233–259. https://doi.org/10.1016/S0010-0277(96)00715-9

Hayne, H., Gross, J., McNamee, S., Fitzgibbon, O., & Tustin, K. (2011). Episodic memory and episodic foresight in 3-and 5-year-old children. *Cognitive Development*, *26*, 343–355. https://doi.org/10.1016/j.cogdev.2011.09.006

Herriot, P. (1969). The comprehension of tense by young children. *Child Development*, *40*, 103–110. https://doi.org/10.1016/S0022-5371(69)80056-3

Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychological Review*, *77*, 223–233. https://doi.org/10.1037/h0029101

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

Hudson, J. A., & Mayhew, E. M. Y. (2011). Children's temporal judgments for autobiographical past and future events. *Cognitive Development*, *26*, 331–342. https://doi.org/10.1016/j.cogdev.2011.09.005

KidsData Project. (2021). A Project of the Population Reference Bureau. https://www.prb.org/projects/kidsdata/. Date accessed June 7, 2025.

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

Leahy, B. P., & Carey, S. E. (2020). The acquisition of modal concepts. *Trends in Cognitive Sciences*, *24*, 65–78. https://doi.org/10.1016/j.tics.2019.11.004

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

Lenth, R. (2025). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.1. (Date accessed June 15, 2025).

Lohse, K., Kalitschke, T., Ruthmann, K., & Rakoczy, H. (2015). The development of reasoning about the temporal and causal relations among past, present, and future events. *Journal of Experimental Child Psychology*, *138*, 54–70. https://doi.org/10.1016/j.jecp.2015.04.008

Maheshwari, U., & Barner, D. (2025). Twice upon a time: Children use syntax to learn the meanings of yesterday and tomorrow. *Developmental Science*, *28*, e13600. https://doi.org/10.1111/desc.13600

McCormack, T., & Hanley, M. (2011). Children's reasoning about the temporal order of past and future events. *Cognitive Development*, *26*, 299–314. https://doi.org/10.1016/j.cogdev.2011.10.001

McCormack, T., & Hoerl, C. (2017). The development of temporal concepts: Learning to locate events in time. *Timing & Time Perception*, *5*, 297–327. https://doi.org/10.1163/22134468-00002094

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

Nelson, K. (1998). *Language in cognitive development: The emergence of the mediated mind*. Cambridge University Press.

Nyhout, A., & Ganea, P. A. (2019). The development of the counterfactual imagination. *Child Development Perspectives*, *13,* 254–259. https://doi.org/10.1111/cdep.12348

Nyhout, A., Sweatman, H., & Ganea, P. A. (2023). Children's hypothetical reasoning about complex and dynamic systems. *Child Development*, *94*, 1340–1355. https://doi.org/10.1111/cdev.13931

Pawlak, A., Oehlrich, J. S., & Weist, R. M. (2006). Reference time in child English and Polish. *First Language*, *26*, 281–297. https://doi.org/10.1177/0142723706059447

Peterson, C. (2002). Children's long-term memory for autobiographical events. *Developmental Review*, *22*, 370–402. https://doi.org/10.1016/S0273-2297(02)00007-2

Peterson, C., & Rideout, R. (1998). Memory for medical emergencies experienced by 1-and 2-year-olds. *Developmental psychology*, *34*, 1059. https://doi.org/10.1037/0012-1649.34.5.1059

Quon, E., & Atance, C. M. (2010). A comparison of preschoolers' memory, knowledge, and anticipation of events. *Journal of Cognition and Development*, *11*, 37–60. https://doi.org/10.1080/15248370903453576

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*, 389–404. https://doi.org/10.1016/j.jecp.2012.10.010

Redshaw, J. (2014). Does metarepresentation make human mental time travel unique? *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*, 519–531. https://doi.org/10.1002/wcs.1308

Redshaw, J. (2024). The recursive grammar of mental time travel. *Philosophical Transactions B*, *379*, 20230412. https://doi.org/10.1098/rstb.2023.0412

Robinson, E. J., & Beck, S. (2014). What is difficult about counterfactual reasoning? In R. Mitchell (Ed.), *Children's reasoning and the mind* (pp. 101–119). Psychology Press. https://doi.org/10.4324/9781315784670

Shirai, Y., & Miyata, S. (2006). Does past tense marking indicate the acquisition of the concept of temporal displacement in children's cognitive development? *First Language*, *26*, 45–66. https://doi.org/10.1177/0142723706060741

Statistics Canada. (2023). (Profile table). Census Profile. 2021 Census of Population. Statistics Canada Catalogue no. 98-316-X2021001. Ottawa. Released November 15, 2023. https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/index.cfm?Lang=E (Date accessed June 10, 2025).

Steele, K., Bánki, A., Markova, G., Hoehl, S., & Tillman, K. A. (2025). Today, Tomorrow, and Overmorrow: The Acquisition of Deictic Temporal Terms in English and German. *Open Mind*, *9*, 1826–1848.

Suddendorf, T. (2010). Linking yesterday and tomorrow: Preschoolers' ability to report temporally displaced events. *British Journal of Developmental Psychology*, *28*, 491–498. https://doi.org/10.1348/026151009X479169

Szagun, G. (1978). On the frequency of use of tenses in English and German Children's spontaneous speech. *Child Development*, *49*, 898–901. https://doi.org/10.2307/1128267

U.S. Census Bureau. (2024). *American Community Survey (ACS) and Puerto Rico Community Survey (PCRS), 5 Year Estimates: 2019-2023*. https://www.census.gov/quickfacts/fact/table/sandiegocountycalifornia

Tillman, K. A., Marghetis, T., Barner, D., & Srinivasan, M. (2017). Today is tomorrow's yesterday: Children's acquisition of deictic time words. *Cognitive Psychology*, *92*, 87–100. https://doi.org/10.1016/j.cogpsych.2016.10.003

Tillman, K. A., Tulagan, N., Fukuda, E., & Barner, D. (2018). The mental timeline is gradually constructed in childhood. *Developmental science*, *21*, e12679. https://doi.org/10.1111/desc.12679

Tillman, K. A., & Walker, C. M. (2022). You can't change the past: Children's recognition of the causal asymmetry between past and future events. *Child Development*, *93*, 1270–1283. https://doi.org/10.1111/cdev.13763

Turan-Küçük, E. N., & Kibbe, M. M. (2024). Three-year-olds' ability to plan for mutually exclusive future possibilities is limited primarily by their representations of possible plans, not possible events. *Cognition*, *244*, 105712. https://doi.org/10.1016/j.cognition.2023.105712

Turan-Küçük, E. N., & Kibbe, M. M. (2025). Children's reasoning about possible outcomes of events in the present and the future. *Developmental Psychology*. Advance online publication. https://doi.org/10.1037/dev0002028

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive psychology*, *23*, 515–557. https://doi.org/10.1016/0010-0285(91)90005-9

Veneziano, E., & Sinclair, H. (1995). Functional changes in early child language: The appearance of references to the past and of explanations. *Journal of Child Language*, *22*, 557–581. https://doi.org/10.1017/S0305000900009958

Weisberg, D. S., & Gopnik, A. (2013). Pretense, counterfactuals, and Bayesian causal models: Why what is not real really matters. *Cognitive science*, *37*, 1368–1381. https://doi.org/10.1111/cogs.12069

Weist, R. M. (1989). Chapter 2 time concepts in language and thought: Filling the piagetian void from two to five years. In I. Levin, & D. Zakay (Eds.), *Advances in psychology* (Vol. *59*, pp. 63–118). North-Holland. https://doi.org/10.1016/S0166-4115(08)61039-0

Weist, R. M., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of Child Language*, *18*, 67–92. https://doi.org/10.1017/S0305000900013301

Weist, R. M., & Zevenbergen, A. A. (2008). Autobiographical memory and past time reference. *Language Learning and Development*, *4*, 291–308. https://doi.org/10.1080/15475440802293490

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org (Date accessed June 15, 2025).

Wickham, H., Averick, M., Bryan, J., *et al.* (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*, 1686. https://doi.org/10.21105/joss.01686

Zhang, M., & Hudson, J. A. (2018). The development of temporal concepts: Linguistic factors and cognitive processes. *Frontiers in Psychology*, *9*, Article 2451. https://doi.org/10.3389/fpsyg.2018.02451