Journal of Experimental Child Psychology 260 (2025) 106322

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

Young children's updating of mental representations of story characters and events based on verbal and pictorial information

Ruth Lee [a,*], Patrycia Jarosz [b], Patricia A. Ganea [b]

[a] York St John University, Lord Mayor's Walk, York YO31 7EX, UK
[b] University of Toronto, Department of Applied Psychology and Human Development, 252 Bloor Street West, Toronto, ON M5S 1V6, Canada

ARTICLE INFO

Keywords:
Updating
Situation model
Narrative
Language
Pictures
Inferencing

ABSTRACT

The ability to create mental models of story events is essential for narrative comprehension, yet little is known about the mechanisms that support children's ability to build and update an integrated mental representation of a story (a 'situation model') as it unfolds. The current study investigated very young children's ability to update their situation model of a simple story from verbal and pictorial information about a physical event, manipulating both the explicitness of verbal information and the informativeness (Study 1) and presence (Study 2) of pictorial information. Sixty-four 2-year-olds (35 girls) and 67 3-year-olds (36 girls) participated in Study 1, and 119 2-year-olds (69 girls) and 81 3-year-olds (43 girls) participated in Study 2. Two- and 3-year-olds updated their mental representation of the physical state of the story protagonist at a rate above chance, regardless of the informativeness of an accompanying picture (Study 1) and the explicitness of verbal information provided (Study 2). However, children's age in months significantly predicted 2-year-olds' performance across studies, and in the absence of a picture, 3-year-olds performed less robustly when receiving implicit than when receiving explicit verbal information. Findings suggest that 2- and 3-year-olds can integrate implicit information into their situation model of a story, even when the accompanying pictorial information is not maximally informative, but that implicit verbal information embedded in a narrative presents challenges for young children's updating when provided without pictorial support.

Introduction

Narratives engage children with the world beyond their direct experience, and so the sophistication with which children comprehend them is a question of importance. In order to comprehend a narrative, listeners must integrate various dimensions of the discourse, such as the spatial, temporal, and causal relations between characters and objects, with their prior knowledge about the world and with previously comprehended parts of the story. The resulting mental model represents the narrative in a unified global structure known as a situation model (Van Dijk & Kintsch, 1983; van den Broek & Lorch, 1993; Zwaan & Radvansky, 1998). As comprehenders encounter new information from the unfolding text, they must update the situation model. Some information may become outdated, for instance when a character changes location, or information may become peripheral to the narrative, as often happens when initial scene-setting gives way to the main plot.

Building a situation model demands rapid automatic judgments about which parts of the story should remain mentally activated

* Corresponding author.
E-mail addresses: r.lee@yorksj.ac.uk (R. Lee), patrycia.jarosz@utoronto.ca (P. Jarosz), patricia.ganea@utoronto.ca (P.A. Ganea).

https://doi.org/10.1016/j.jecp.2025.106322
Received 14 January 2025; Received in revised form 5 May 2025;
Available online 11 June 2025
0022-0965/

and available for integration with new information from the narrative as it unfolds (Zwaan & Radvansky, 1998). This includes not only explicit information, but inferences that have already been generated (Graesser et al., 1994; Van den Broek et al., 2005). While preschoolers find explicit information easier to process than implicit information in oral texts (Florit et al., 2011), the ability to generate inferences while processing a story predicts preschoolers' story comprehension both concurrently (Tompkins et al., 2012) and longitudinally (Kendeou et al., 2008) and does so independently of children's literal text comprehension, basic language skills, and vocabulary (Kendeou et al, 2008).

Many children's books convey important story information by providing affordances for inference. Some inferences are necessary in order to maintain comprehension. For example, when Winnie the Pooh gets stuck in Rabbit's hole, drawing an inference from Pooh's having eaten honey all morning that Pooh has become too big to fit through the hole is crucial to understanding the plot (that is, Pooh is now unable to move). Other inferences are elaborative and, while not immediately necessary for story comprehension, allow comprehenders to create a richer model of what a situation is like (for instance, hearing that snow fell heavily and inferring that the ground is white). Yet we do not know when children acquire the ability to successfully make connections between information in the narrative and their real-world knowledge in a way that enables them first to draw accurate inferences, and then to use these inferences to update their mental representation of the situation in the story.

By their second birthday, children can use explicit verbal information, delivered in person, to update representations about real events such as a change in an object's location or property. Children as young as 22 months can use language to understand changes in an object's properties (e.g., "Lucy got wet") even when they cannot see the object (Ganea et al., 2007). However, young children's ability to use verbal testimony to update their mental representations of absent entities varies depending on the task and their age: when verbal input describes a change in the location of an unseen object (e.g., "The puppy is moved from the bag to the box"), 30-month-olds can successfully find the object in its new location, whereas most 23-month-olds will continue to search in the object's original location (Ganea & Harris, 2010). In the location task, the requirement to activate a 'mental file' not just for an object, but also for its location, may explain children's additional difficulty (Özdemir & Ganea, 2020). Thus, while 2-year-old children demonstrate the ability to update information about an absent object, this ability appears fragile.

There is also evidence that children as young as 2 can update situation models when they are implied in make-believe. For example, if shown a toy monkey squirting pretend (i.e., invisible) ketchup onto a toy pig, children successfully infer that the pretend ketchup left the container and was transferred to the pig, and they then select a picture of the transformed pig (covered in ketchup) from a set including a clean pig and a transformed distractor animal (Harris et al., 1996). Thus, 2-year-old children enrich their mental representation of the pretend action by integrating it with their real-world knowledge, form a mental representation of imaginary events that are causally related, and infer the outcome.

An open question, unstudied to the best of our knowledge, is whether children of this age demonstrate the ability to update a mental representation with information they receive about the outcome of an action within the context of a story narrative. This process may present different challenges to those posed by either a real or a concurrent pretend transformation. First, it may be more difficult for children of this age to infer the outcome of an action without the social input of a pretence partner (Gelman, 2009) whose interactional style may support updating. Second, the drawing of an inference about an action within a story narrative requires the listener to suppress the activation of plot features that have become irrelevant. If updating is to keep pace with the unfolding story, then once a protagonist has rejected a particular course of action, objects associated with it should be activated to a lesser degree than objects associated with the chosen course of action (Glenberg et al., 1987). Thus, to succeed in updating a situation model, children must suppress the activation of an irrelevant action mentioned at the start of the story. As stories unfold, they afford simple elaborative inferences about outcomes that have not yet taken place, whereas when updating information about real (or indeed concurrent pretend) events, children are mentally representing a change that has already happened. In order to form elaborative inferences from a story narrative and to create expectations for the story action, children need to dynamically track, update, and coordinate representations of the situation based on described events, maintaining unified but continually shifting situation models (Zwaan & Radvansky, 1998). This requires keeping relevant information activated in memory so that it is accessible for retrieval during the inferencing process, across a number of dimensions of the situation in the story: time, space, causation, intentionality, and protagonist. If young children's greater difficulties with updating the location of real objects (versus their physical state) are, as Özdemir & Ganea (2020) suggest, partly explained by the need to maintain activation of multiple representations, then the tracking and coordination of multiple elements of a narrative may prove a challenge, particularly when stories provide implicit, rather than explicit information about events.

Surprisingly little recent work has addressed young children's narrative situation model competence, perhaps because research on discourse processing in young children has tended to focus on narrative production, rather than comprehension (as discussed by Paris & Paris, 2003), and theoretical perspectives on young children's narrative competence have tended to conceptualize it as an activity, rather than as a form of mental representation (Berman, 2009). There is some evidence that around their fourth birthday, children update the situation model of a story in terms of time and space by taking the mental perspective of a protagonist: they show slower processing times for descriptions of a visual scene when the protagonist is described as walking as opposed to driving (Fecica & O'Neill, 2010). Rall & Harris (2000) asked whether even younger children could take a protagonist's point of view when provided with a familiar fairytale as the narrative. However, to our knowledge, the ability of 3-year-olds to construct a situation model from an unfamiliar story has not been addressed; nor has situation model building in children younger than this been studied, although children of this age regularly hear unfamiliar stories in a number of genres, including narrative (Rudin, 2024).

One way to support young children's situation model competence, and thus their inferencing in a story context, may be to include informative pictures. Two-year-olds' visual attention to relevant storybook illustrations predicts their comprehension (Kaefer et al., 2017), and there are rapid developments in children's pictorial understanding during the first few years of life (for reviews, see

DeLoache, 2011a, 2011b; Jolley, 2008), suggesting that including pictures may facilitate younger children's ability to draw inferences and to use them to update their mental model of a story. Thus, when a story includes an accompanying picture that supports a target inference at a crucial causal juncture, we might expect it to support children's ability to update their mental model of the narrative. A less supportive picture (that is, one that does not directly support the target inference) may play no role in children's processing of the narrative, or may even divide children's attention and serve as a distraction. To date, we are not aware of any studies addressing the role of pictorial support in young preschoolers' situation model construction, nor of studies manipulating the extent of its informativeness. While there is evidence that pictorial representations support 7-year-olds' situation model representations of simple stories, as measured using a sentence recognition task (Seger et al., 2019), we do not yet know whether this is the case for much younger children whose competence in updating a mental model of real events is still emerging.

In the present study, we ask whether 2- and 3-year-olds can draw a simple inference from verbal and pictorial information in the context of a story, by integrating this information with their general knowledge of the world (e.g., water makes people wet). In Study 1, we presented children with verbal information that afforded an inference about an action in the story, and we manipulated the informativeness of accompanying pictorial information. In Study 2 (preregistered at https://osf.io/8jf39/), we manipulated the presence or absence of an informative picture. We also manipulated the explicitness of the verbal information. In both studies, the outcome measure was a forced choice between two alternative images of the story character and a distractor (as in Ganea et al., 2007; Özdemir & Ganea, 2020). All manipulations were administered between participants, and children received two trials in each study. Stories were presented as audiovisual narrations on a computer screen. While very young Western children's exposure to symbolic media is arguably becoming a fundamental part of their developmental context (Barr, 2019), young children's challenges in learning from digital media are well-documented (see Strouse & Samson, 2021 for a review), and it is therefore important to discover whether young children can update a simple situation model presented digitally.

Study 1

We first asked whether an informative picture would support young children's updating. Since the ability to update mental representations based on explicit information emerges and develops during the second year of life (Ganea & Saylor, 2013), we expected that 2-year-olds, but not 3-year-olds would update more successfully when the story was accompanied by an informative rather than uninformative picture (H1). Second, we expected an age effect, with 3-year-olds performing better than 2-year-olds overall (H2), given evidence for increases during this age period in symbolic reasoning ability and narrative comprehension (Apperly et al., 2004; DeLoache, 1997; Perfetti et al., 2005). Third, given prior evidence that 3-year-olds can simulate a character's movements and actions during story comprehension (Fecica & O'Neill, 2010), we also expected this age group to perform more consistently than 2-year-olds, succeeding at a rate above chance (H3).

Method

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study, and we follow JARS (Appelbaum et al., 2018). All data, analysis code, and research materials are available at osf.io/8jf39. Data were analyzed using SPSS versions 28 and 29 and R version 4.0.1 (R Core Team, 2020) utilising the packages tidyverse 1.3.0 (Wickham et al., 2019) and geepack 1.3-2 (Højsgaard et al., 2006). Data and materials are available at osf.io/8jf39.

Participants

Data were collected in 2016–17. We used prior studies on young children's updating as a guideline for determining the sample size (Galazka & Ganea, 2014; Ganea & Harris, 2010; Ganea et al., 2007; Özdemir & Ganea, 2020). Sixty-four 2-year-olds (*M*age = 31.09 months, *SD* = 2.51, range: 25.58 –35.97; 35 girls) and 67 3-year-olds (*M*age = 42.25 months, *SD* = 3.56, range: 36.26–47.84; 36 girls) participated. Inclusionary criteria were normal or corrected-to-normal vision, no history of diagnosis or treatment of cognitive, speech, language, hearing, or attentional issues, and hearing English spoken at home more than 75% of the time.

Children were randomly assigned to one of two conditions: a story accompanied by an informative picture (33 2-year-olds, 33 3-year-olds) and a story accompanied by an uninformative picture (31 2-year-olds, 34 3-year-olds). Data from 31 additional participants were excluded from analysis due to lack of attention (five 2-year-olds, one 3-year-old), failing to make a choice on either trial (18 2-year-olds, three 3-year-olds), parental interference (two 2-year-olds, one 3-year-old), and experimenter error (one 3-year-old). Exclusions were evenly distributed across conditions.

Children were recruited from a family volunteer participant pool at a public university, local day cares and a Science Centre in a large North American city. The majority of participants were White or mixed ethnicity. Due to data loss of demographic information, information about participants' specific ethnicity cannot be provided. Data collection took place in a developmental laboratory at the university, in quiet spaces at daycares, and in a designated room at the Science Centre.

Materials

A sequence of 18 pictures was created using online cartoon generation software (ToonDoo.com) and displayed using animated Power Point slides on a 24-inch monitor. The accompanying verbal input was recorded by a female native speaker of Canadian English.

Sound was presented through loudspeakers located behind the screen. A camera was placed on a nearby table or on a tripod at a 90-degree angle to the screen to record children's behaviour during the task.

Procedure

Children sat on a child-sized bench or chair at a table ~ 60 cm from the computer monitor, and the experimenter sat next to the child. The experimenter told the child 'Let's watch a story!' The experimenter advanced the PowerPoint slides manually, which allowed her to bring the child's attention back to the story if the child became distracted. When this occurred, the experimenter regained the child's attention by saying 'look' and gesturing towards the screen. Stimuli comprised two trials. Each trial comprised a brief description phase (40 s), during which a story about a protagonist was presented, and a test phase lasting between approximately 10 and 20 s, in which listeners were asked to identify the protagonist from several characters presented on the screen.

During the description phase, trials were administered within participants. All children saw two stories, both comprising a series of static pictures accompanied by narration, which concerned a child protagonist choosing between two activities. In the first story, a boy, Tom, chooses to swim in a pool rather than paint (henceforth, Pool Story), and in the second story, a girl, Sally, chooses to hug her muddy dog rather than eat ice-cream (henceforth, Dog Story) (Fig. 1). All stimuli are available at osf.io/8jf39.

The protagonist then begins to engage in their chosen activity (Tom jumps into the air above the pool; Sally reaches out her arms towards her dog). At this critical juncture, children heard 'Tom jumps into the water!' (Pool Story) or 'Sally hugs the muddy dog!' (Dog Story). In the informative picture condition, the accompanying picture included water (Pool Story: Tom mid-air above the pool) or mud (Dog Story: Sally stretching her arms towards the muddy dog). In the uninformative picture condition, the accompanying picture did not include the substance involved in the action (Tom in mid-jump against a background of the sky, with the pool below him located off-screen; Sally stretching her arms out towards an off-screen muddy dog).

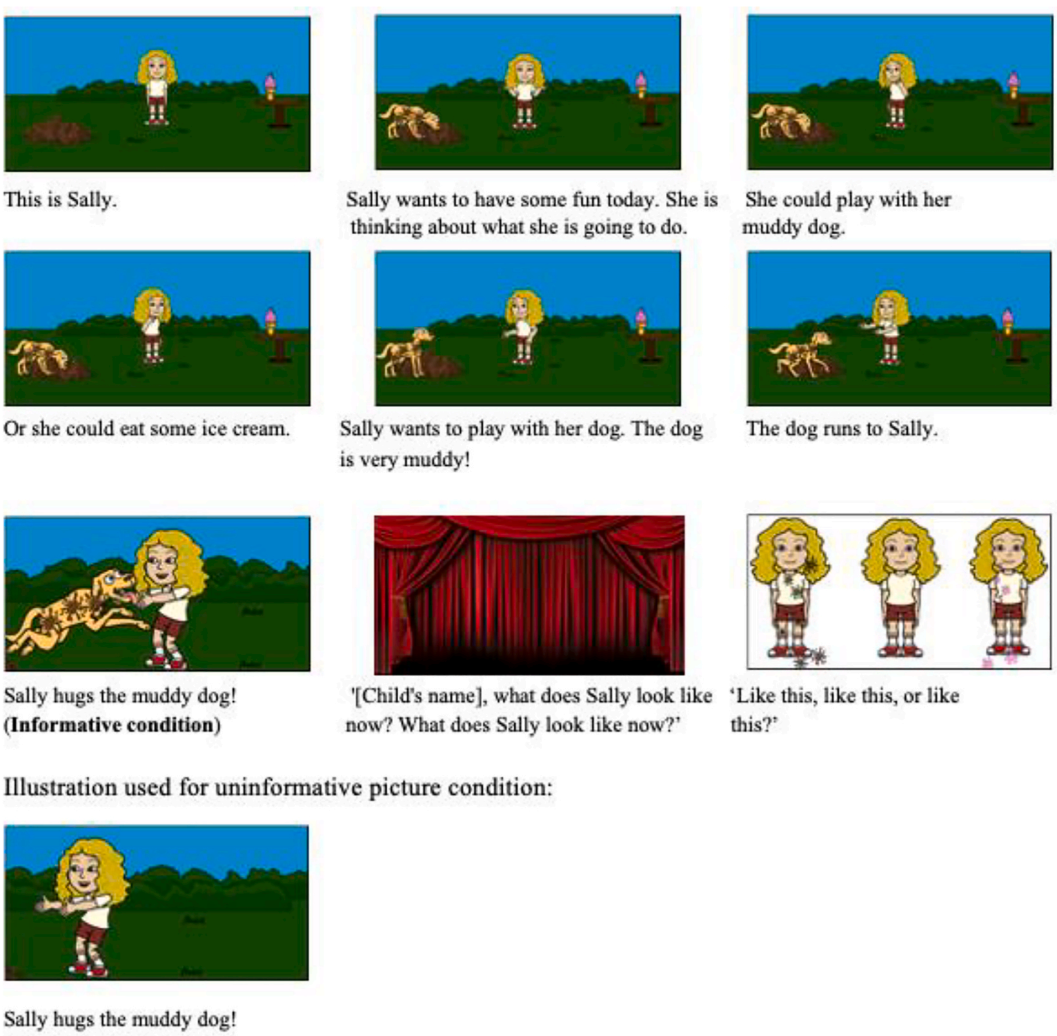

Fig. 1. Dog Story stimuli in informative and uninformative picture conditions, Study 1.

The subsequent test phase began with an image of a curtain. As soon as the curtain appeared, the experimenter made eye contact with the child and asked in an interested tone '[Child's name], what does (e.g.) Tom look like now?' Without pause the experimenter then advanced to the test slide, which presented three images of the protagonist: a Target image covered in dripping water (Pool Story) or with mud (Dog Story); a Distractor image covered in paint (Pool Story) or ice-cream (Dog Story); and an Untransformed image of a clean protagonist. Immediately following the appearance of the test slide, the experimenter swept her hand back and forth between the three versions of the protagonist displayed on the screen, while saying to the child 'What does (s)he look like now, like this, like this, or like this'? Thus, if children simply remembered that (e.g.) Tom had done something that would transform his physical appearance, they should choose randomly between the wet target and the painted distractor. If they had not updated their representation of Tom, they would choose the untransformed, dry Tom (an image matching that they had originally seen) or choose randomly between all three images. Finally, if the children had updated their representation of Tom on the basis of what they had heard and seen, they would choose the wet target Tom, ignoring both the untransformed, dry Tom and the painted distractor.

If children did not point, pointed to more than one image, or made multiple points that did not represent an intentional choice (such as touching every picture in quick succession), the experimenter repeated the question once. We counterbalanced the order of presentation of the stories (Pool and Dog) and the location of the target character on the screen (left or right, with the Untransformed character always displayed in the centre).

Data coding and scoring

Two research assistants coded children's points to the on-screen images at test (Target, Distractor, or Untransformed). If a child gave a verbal response that unambiguously identified one of the three images on the screen (e.g., 'Wet') but did not point, the verbal response was accepted. Only those points judged to be intentional were included in the analysis. Children who did not choose an image, or who indicated more than one image in a way that did not appear to indicate a choice (for instance, pointing to all three images in quick succession) were excluded from the analysis. Agreement between coders was high, $\kappa = .89$, $p < .001$. Disagreements were resolved by a third coder.

Results and Discussion

Before proceeding to the main analysis, we screened for potential effects of story (Pool Story or Dog Story) on children's performance. For each of the two stories, children were deemed successful if they pointed to the Target image and unsuccessful if they pointed to a non-target (Distractor or Untransformed image). Thus, children selecting the target image on a particular story were scored 1 for that story; failure to select the target image resulted in a score of 0 for that story. A Generalized Estimating Equation (GEE) analysis with a binomial distribution demonstrated a main effect of age, such that 2-year-olds had lower odds than did 3-year-olds of succeeding on any given trial (Wald x^2 (1) = 8.50, $p = .004$, $b = -.456$, 95 % CI [-1.45,.54], $SE = .51$, Exp (B) = .634, 95 % CI [.24, 1.71]; see Fig. 2). There were no other significant effects (all $ps > .129$). Thus, regardless of age or condition children did not perform differently across stories. Since there were two trials, this result also indicates that children's scores on the second trial were not significantly different than their scores on the first trial. Subsequent analyses therefore collapsed children's scores across Pool and Dog stories.

Next, we screened for potential effects on children's overall performance of the order in which children heard the two stories and of gender. Children selecting the target image on both trials resulted in a score of 1, and failure to select the target image on one or both trials resulted in a score of 0. A logistic regression analysis stratified by age group and condition was conducted on children's scores. There were no significant effects for either age group in either condition (all $ps > .099$).

To conduct the main analysis investigating the effect of condition (informative picture, uninformative picture), story (Pool Story, Dog Story), age in months, and their two-way interactions on children's choices, we submitted the data to a further Generalized

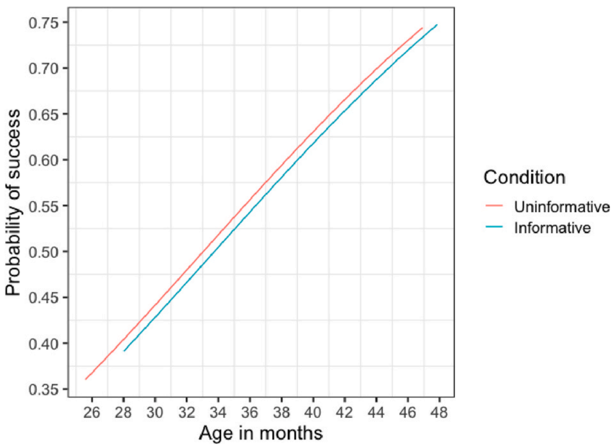

Fig. 2. Probability of success (0–1) collapsed across Pool and Dog stories, by condition (informative, uninformative).

Estimating Equation (GEE) analysis with binomial distributions, logit-log link functions and independent covariance structures. Testing location (laboratory or daycare versus Science Centre, which involved testing children in a somewhat noisier environment) and order of story presentation were included as covariates. This analysis accommodates the binary nature of the dependent variable and the presence of a within-subject factor (story) in the data. On each trial, children's responses were scored as 0 if they selected an incorrect response (Distractor or Untransformed) and 1 if they selected the correct response (Target). Table 1 shows the proportion of children who selected each response category.

There was a main effect of age in months, such that younger children were less likely to succeed than older children on any given trial (Wald x^2 (1) = 7.54, p = .006, b = .11, 95 % CI [.03,.18], SE = .04, Exp (B) = 1.11., 95 % CI [1.03, 1.2]). There were no other significant effects (all ps > .136). Thus, regardless of age or condition, children did not perform differently across stories, and neither testing location nor the order of story presentation affected overall performance. All interactions were subsequently dropped from the model. The main effect of age in months remained, such that younger children had lower odds than did older children of succeeding on any given trial (Wald x^2 (1) = 12.46, p < .001, b = .08, 95 % CI [.04,.13], SE = .02, Exp (B) = 1.09, 95 % CI [1.04, 1.14]). There were no other significant effects (all ps > .148). Finally, all three control variables (story, testing location, and item order) were dropped from the model. There was no significant effect of condition (p = .84), and thus no support for H1. The main effect of age in months remained (Wald x^2 (1) = 11.81, p < .001, b = .08, 95 % CI [.03,.12], SE = .02, Exp (B) = 1.08, 95 % CI [1.04, 1.12]), confirming H2. This model was retained.

We then conducted tests against chance in order to examine children's robust performance, comparing the distribution of children's correct responses across two trials (0, 1, or 2 out of 2 trials correct; Table 2). Chance calculations across trials considered the fact that on each separate trial the probability of success was.33 (.67 for failure) given the 3-choice item test structure. Across 2 trials, the probability was .45 (.67 x .67) for scoring 0 on 2 consecutive trials, .44 (.33 x .67) = .22 for scoring 1 on 2 trials, and. 11 (33 x .33) for scoring 2 on 2 trials. Given that the GEE analyses revealed no effect of condition, we examined the performance of each age group collapsed across both conditions in which they were tested (informative and uninformative picture). Chi-square goodness of fit tests demonstrated that both 2-year-olds' (x^2 (2) = 10.01, p = .007, ω = 0.40) and 3-year-olds' (x^2 (2) = 95.99, p < .001, ω = 1.20) pattern of responses across trials was significantly better than that expected by chance. Thus, H3 was not supported.

Taken together, these results suggest that during their third year of life, children acquire the ability to infer a story outcome (i.e., wetness) when hearing implicit verbal information (i.e., Tom jumps in the water) by integrating verbal information with their existing knowledge. Contrary to our prediction (H1), not only 3-year-olds, but also 2-year-olds were able to draw this inference regardless of whether the picture was supportive of the transformation described in the story. However, as predicted (H2), children's success increased as a factor of age in months. The prediction that 3-year-olds, but not 2-year-olds would perform better than chance (H3) was not supported; however, it is important to note that the size of the effect was substantially larger in 3-year-olds than in 2-year-olds. Overall, the 2-year-olds had lower odds of succeeding on any given trial in Study 1.

To find out whether the source of 2-year-olds' difficulty came from the fact that they had to update based on an inference, in Study 2 we manipulated the explicitness of verbal information. Two-year-olds were tested in conditions where no inference was required (the verbal information was explicit) and this was done either in the presence or absence of a picture. Because 3-year-olds' performance was more robust in Study 1 on the basis of implicit information, in Study 2 we were particularly interested in whether this performance holds even when no picture is present.

Study 2

Study 2 manipulated both the availability of picture information and the explicitness of verbal information, allowing us to address three research questions and a number of pre-registered hypotheses (available with the data and materials for the experiment at osf.io/8jf39). Of particular interest was whether pictorial support is necessary for 3-year-olds to successfully update their mental representation of a narrative on the basis of an inference drawn from implicit information, and whether 2-year-olds who heard explicit information but saw no picture would be able to apply their demonstrable updating skills (Ganea et al., 2007) in a narrative context.

First, we asked whether updating is harder for young children on the basis of an inference than on the basis of explicit information. We predicted that explicitness of verbal information would influence 2- and 3-year-olds' performance, such that both groups would be

Table 1
Number and proportion of participants choosing each image (target, untransformed target, distractor) by age group and condition, Study 1.

Age group and trial	Condition	Target n (%)	Untransformed n (%)	Distractor n (%)
2-year-olds				
Pool Story	Uninformative	12 (38.7)	9 (29)	10 (32.3)
	Informative	18 (54.5)	9 (27.3)	6 (18.2)
Dog Story	Uninformative	17 (54.8)	5 (16.1)	9 (29)
	Informative	11 (33.3)	14 (42.4)	8 (24.2)
3-year-olds				
Pool Story	Uninformative	22 (64.7)	6 (17.6)	6 (17.6)
	Informative	22 (66.7)	5 (15.2)	6 (18.2)
Dog Story	Uninformative	22 (64.7)	9 (26.5)	3 (8.8)
	Informative	22 (66.7)	9 (27.3)	2 (6.1)

Table 2
Number and percentage of children giving correct responses across two trials (0, 1, or 2 out of 2 trials correct) as a function of age and condition, Study 1.

Condition	Informative picture *n* (%)			Uninformative picture *n* (%)		
Score	0/2	1/2	2/2	0/2	1/2	2/2
2-year-olds	9 (27)	18 (55)	6 (18)	9 (29)	15 (48)	7 (23)
3-year-olds	5 (15)	12 (36)	16 (49)	7 (21)	11 (32)	16 (47)

more likely to update when explicit verbal information was given than when implicit verbal information was provided (H1a). In the case of 3-year-olds, we predicted that receiving explicit verbal information, even in the absence of pictorial information, would support appropriate inferences at a rate significantly higher than expected by chance (H1b). We made no strong prediction regarding 3-year-olds' performance when receiving implicit verbal information with no picture.

Second, we predicted that when receiving explicit verbal information, the presence of an informative picture would influence 2-year-olds' performance, such that they would be more likely to update when a picture is provided than when it is not (H2a). We also expected that 2-year-olds seeing an informative picture while listening to explicit verbal information would demonstrate robust performance at a rate significantly higher than expected by chance, whereas 2-year-olds who received explicit verbal information and no picture would do so at a rate no different from chance (H2b); while previous research (Ganea et al., 2007) indicates that explicit verbal information can be sufficient to support updating, this work did not situate the information in a narrative context. Given that 2-year-olds gave a less robust performance than 3-year-olds when receiving implicit verbal information accompanied by a picture in Study 1, we made no strong prediction about 2-year-olds' performance in these circumstances.

Third, we investigated whether children's performance improves across their third and fourth years of life. We anticipated that age in months would predict appropriate inferencing in both 2-year-olds and 3-year-olds, such that older children were more likely to select the target referent, with no interaction between age in months and condition in either age group (H3).

Method

Participants

Data were collected in 2023. Power analysis using G*Power indicated that for tests against chance (chi-square goodness of fit), in order to detect a large effect (w = 0.5) at the conventional levels for alpha (.05) and power (.8), we required a minimum of 39 children per condition: that is, 117 2-year-olds and 78 3-year-olds. One hundred and nineteen 2-year-olds (*M*age = 30.05 months, SD = 3.64,

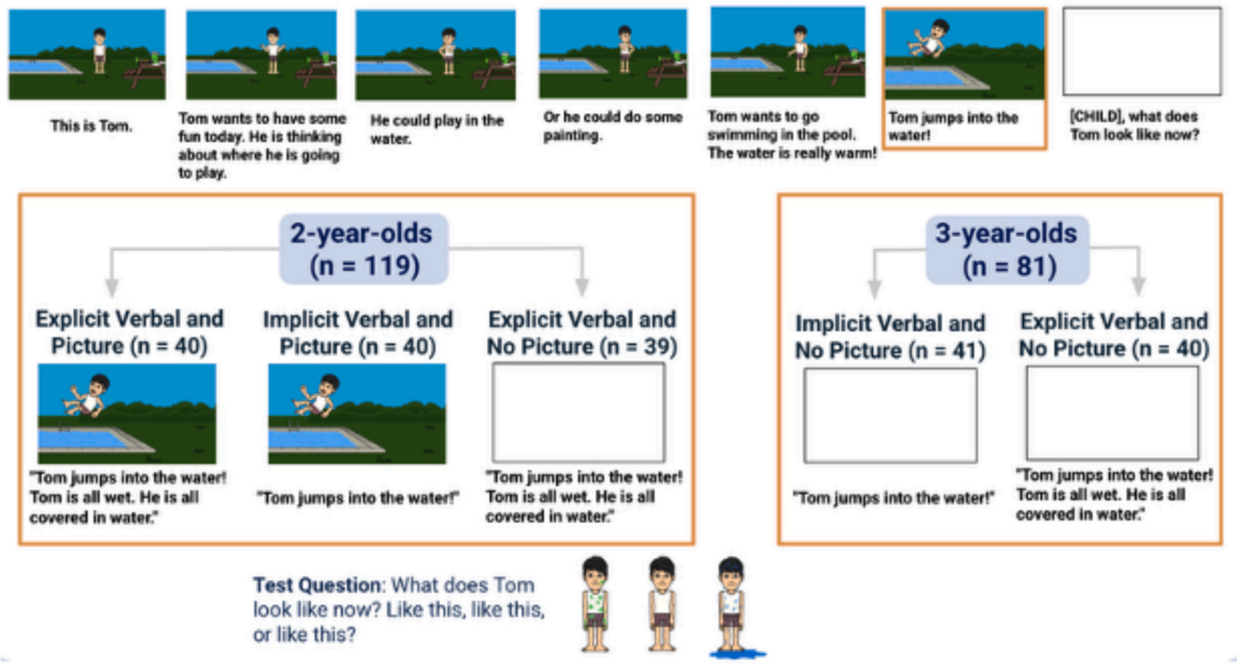

Fig. 3. Conditions to which 2- and 3-year-olds were randomly allocated and procedure, Study 2. The penultimate slide varied by condition, as illustrated by the boxes in orange outline. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

range: 24–36 months, 69 girls) and 81 3-year-olds (*M*age = 41.95 months, SD = 3.21, range: 36–47 months; 43 girls) participated. Parents reported their ethnicity as White (39 %), mixed (24.5 %), South Asian (11 %), Chinese (9.5 %), Black (6.5 %), Filipino (3.5 %), West Asian (2 %), Latin American (1 %), Arab (1 %), and Southeast Asian (.5%). An additional 1.5 % of families declined to disclose ethnicity information. Inclusionary criteria were identical to those in Study 1.

Children were randomly assigned to conditions as follows (Fig. 3). Two-year-olds were randomized between explicit verbal information with picture (40 2-year-olds), explicit verbal information with no picture (39 2-year-olds), and implicit verbal information with picture (40 2-year-olds). Three-year-olds were randomized between explicit verbal information with no picture (40 3-year-olds), and implicit verbal information with no picture (41 3-year-olds).

We counterbalanced the order of presentation of the stories (Pool and Mud) and the location of the target characters on the screen (left or right, with the Untransformed character always displayed in the center). An additional 29 children were tested but excluded from the analysis due to not meeting the language or age requirement ($n = 17$), failing to make a choice on either trial ($n = 11$) or being inattentive ($n = 1$). Exclusions were evenly distributed across conditions and age groups.

Children were recruited via a family volunteer participant pool at a public university in a large North American city and from local daycares. Data collection took place in a developmental laboratory at the university and in quiet spaces at daycares.

Materials

Stimuli were identical to those used in Study 1 save for the following changes. First, the visual representation of muddy and water-soaked protagonists at test was somewhat enhanced. Second, a blank screen was presented in the no-picture conditions. Finally, when presented at test, verbal information was additionally paraphrased (e.g., in the Pool Story, the protagonist was said not only to be ‘all wet’, but also ‘covered in water’: see Fig. 3). The procedure for presentation of pictures and audio was identical to that in Study 1. All stimuli are available at osf.io/8jf39.

Design

In picture conditions, a character was seen in the process of completing the verbally described action, including visual context (equivalent to the informative picture information provided in Study 1). In no-picture conditions, the screen was blank. In implicit verbal information conditions, verbal information described an action, but not its consequences (e.g., a character was said to jump into a pool). In explicit verbal information conditions, actions and their consequences were described (e.g., a character was said to jump into a pool, be swimming in the water, be “all wet”, and be “covered in water”).

Procedure

Stimuli again comprised two trials. The narrative, description, and test phases were structured as in Study 1. However, the verbal and visual information presented across conditions was patterned differently from that in the narrative and description phases of Study 1. As the protagonist begun to engage in their chosen activity (e.g., Tom jumps into the air above the pool), children in the explicit verbal with picture condition heard “Tom jumps into the water. Tom is all wet. He is all covered in water”, alongside a picture of Tom jumping into the air above the pool. Children in the explicit verbal with no picture condition heard identical verbal information and saw a blank screen. Children in the implicit verbal with picture condition heard only “Tom jumps into the water” and saw a picture depicting Tom jumping into the air above the pool. Children in the implicit verbal with no picture condition also heard only “Tom jumps into the water” and saw a blank screen.

Data coding and scoring

Data coding and scoring were performed in a manner identical to that in Study 1. There was nearly perfect agreement between the two coders, $k = .98$, $p < .001$.

Table 3
Number and proportion of participants choosing each image (target, untransformed target, distractor) by age group and condition, Study 2.

Age group and trial	Condition	Target *n* (%)	Untransformed *n* (%)	Distractor *n* (%)
2-year-olds				
Pool Story	Explicit verbal with picture	21 (52.5)	10 (25)	9 (22.5)
	Explicit verbal, no picture	19 (48.7)	8 (20.5)	12 (30.8)
	Implicit verbal with picture	24 (60)	6 (15)	10 (25)
Dog Story	Explicit verbal with picture	16 (40)	9 (22.5)	15 (37.5)
	Explicit verbal, no picture	24 (61.5)	2 (5.1)	13 (33.3)
	Implicit verbal with picture	22 (55)	10 (25)	8 (20)
3-year-olds				
Pool Story	Explicit verbal, no picture	36 (90)	0 (0)	4 (10)
	Implicit verbal, no picture	28 (63.3)	8 (19.5)	5 (12.2)
Dog Story	Explicit verbal, no picture	39 (97.5)	0 (0)	1 (2.5)
	Implicit verbal, no picture	29 (70.7)	8 (19.5)	4 (9.8)

Results and discussion

Table 3 shows the proportion of children who selected each response category.

Before proceeding to the main analysis, we screened for potential effects of story (Pool Story or Dog Story) on children's performance. As in Experiment 1, for each of the two stories, children were deemed successful if they pointed to the Target image and unsuccessful if they pointed to a non-target (Distractor or Untransformed image). Thus, children selecting the target image on a particular story were scored 1 for that story; failure to select the target image resulted in a score of 0 for that story. A Generalized Estimating Equation (GEE) analysis with a binomial distribution demonstrated a main effect of age, (Wald x^2 (1) = 42.16, $p < .001$, $b = .16$, 95 % CI [.11,.20], $SE = .02$, Exp $(B) = 1.17$, 95 % CI [1.12, 1.22]. There were no other significant effects (all $ps > .073$). Thus, regardless of age or condition children did not perform differently across stories. Subsequent analyses therefore collapsed children's scores across Pool and Dog stories.

Next, we screened for potential effects on children's overall performance of the order in which children heard the two stories and of gender. As in Experiment 1, children selecting the target image on both trials resulted in a score of 1, and failure to select the target image on one or both trials resulted in a score of 0. A logistic regression analysis stratified by age group and condition was conducted on children's scores. There were no significant effects for either age group in any condition (all $ps > .75$).

To investigate the effect of condition (implicit, explicit verbal information), age in months, and their two-way interactions on trial-level data (choose target, scored 1; choose non-target, scored 0), data were submitted to Generalized Estimating Equation (GEE) analysis with binomial distributions, logit-log link functions and independent covariance structures. Story (Pool Story, Dog Story) and order of story presentation were included as covariates. Given that most experimental conditions were completed either by 2-year-olds or 3-year-olds but not both, analyses were stratified by age group.

Table 4 shows the proportion of 2-year-olds who selected each response category. For 2-year-olds there were no significant effects, though there was a marginal effect of age in months ($p = .052$; all other $ps > .125$). Interactions that included item order were removed from the model. There was a main effect of age in months (Wald x^2 (1) = 18.04, $p = .006$, $b = .178$, 95 % CI [.05,.30] $SE = .06$, Exp $(B) = 1.2$, 95 % CI [1.05,1.36]). There were no other significant effects (all $ps > .123$). Thus, regardless of age in months or condition, 2-year-olds did not perform differently across stories, nor did order of story presentation affect overall performance. All interactions were subsequently dropped from the model. The main effect of age in months remained (Wald $x^2(1) = 15.67$, $p < .001$, $b = .15$, 95 % CI [.08,.23] $SE = .04$, Exp $(B) = 1.16.$, 95 % CI [1.08, 1.25]). There were no other significant effects (all $ps > .179$). Finally, story and item order were dropped from the model. There was no significant effect of condition (all $ps > .261$). Given the lack of condition effects, we did not find support for our prediction that 2-year-olds would perform better when given explicit than when given implicit verbal information (H1a), nor that when presented with explicit verbal information, they would do better when shown an informative picture than when shown no picture (H2a). The main effect of age in months remained, such that older 2-year-olds had higher odds than did younger 2-year-olds of succeeding on any given trial, supporting our prediction of a significant effect of age in months (H3) for this age group (Wald x^2 (1) = 12.97, $p < .001$, $b = .14$, 95 % CI [.06,.21], $SE = .04$, Exp $(B) = 1.14$, 95 % CI [1.06, 1.23]). This model was retained and is illustrated, along with the 3-year-olds' model described below, in Fig. 4.

Table 5 shows the proportion of 3-year-olds who selected each response category. For three-year-olds, the model including all interactions between predictors and between predictors and covariates did not converge. Interactions that included item order were removed from the model. There was a main effect of story ($p = .016$) and an interaction between story and age in months (Wald x^2 (1) = 4.83, $p = .028$, $b = .27$, 95 % CI [.03,.51] $SE = .12$, Exp $(B) = 1.31$, 95 % CI [1.03, 1.66]), such that the older the child, the higher their odds of success on the pool story trial, but not the dog story trial. There were no other significant effects (all $ps > .200$). All interactions were subsequently dropped from the model. There was a main effect of condition (Wald x^2 (1) = 8.94, $p = .003$, $b = .1.80$, 95 % CI [.61, 2.97] $SE = .60$, Exp $(B) = 6.02$, 95 % CI [1.89, 19.5]), and no other significant effects (all $ps > .242$); hence, our prediction of a significant effect of age in months (H3) was not supported in this age group. Finally, both control variables (story, item order) were dropped from the model. The main effect of condition remained, such that 3-year-olds who heard explicit verbal information without a picture had higher odds of success than 3-year-olds who heard implicit verbal information without a picture (Wald x^2 (1) = 9.07, $p = .003$, $b = 1.76$, 95 % CI [.61, 2.9] $SE = .58$, Exp $(B) = 5.8$, 95 % CI [1.85, 18.22]), confirming for this age group our prediction that children would perform better when given explicit than when given implicit verbal information (H1a). This model was retained and, along with the 2-year-olds' model described above, is illustrated in Fig. 4.

Next, chi-square goodness of fit analyses were conducted to examine performance against chance across trials (0, 1, or 2 trials correct; Tables 4 and 5), stratified by age group and condition.

As in Study 1, the chance of succeeding on one trial was.22, and the chance of succeeding on two trials was.11. Given that the GEE analyses revealed no effect of condition for 2-year-olds, we examined their performance collapsed across all three conditions in which they were tested (explicit verbal information with picture; explicit verbal information with no picture; implicit verbal information with picture). Two-year-olds demonstrated success at a rate above chance, x^2 (2) = 46.10, $p < .001$, $\omega = 0.62$. Note that analysing 2-year-

Table 4

Number and percentage of 2-year-olds giving correct responses across two trials (0, 1, or 2 out of 2 trials correct) as a function of condition, Study 2.

Condition	Explicit verbal with picture *n* (%)			Explicit verbal, no picture *n* (%)			Implicit verbal with picture *n* (%)		
Score	0/2	1/2	2/2	0/2	1/2	2/2	0/2	1/2	2/2
	13 (32.5)	17 (42.5)	10 (25.0)	5 (12.8)	25 (64.1)	9 (23.1)	8 (20)	18 (45)	14 (35)

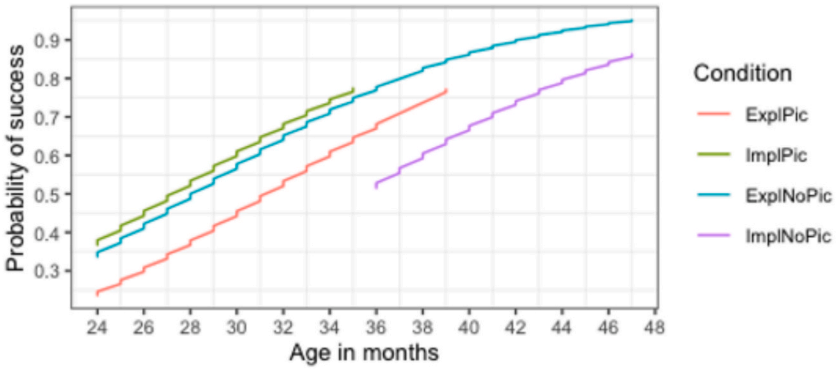

Fig. 4. Probability of success (0–1) collapsed across Pool and Dog stories, by condition (explicit verbal information with picture, implicit verbal information with picture, explicit verbal information with no picture, implicit verbal information with no picture).

Table 5
Number and percentage of 3-year-olds giving correct responses across two trials (0, 1, or 2 out of 2 trials correct) as a function of condition, Study 2.

Condition	Explicit verbal, no picture *n* (%)			Implicit verbal, no picture *n* (%)		
Score	0/2	1/2	2/2	0/2	1/2	2/2
	0 (0)	5 (12.5)	35 (87.5)	9 (22)	7 (17.1)	25 (61)

olds' responses separately by condition (our pre-registered analysis) produces analogous results: they performed above chance in the explicit verbal with picture condition, $\chi^2(2) = 8.71, p = .013$; in the implicit verbal with picture condition, χ^2 $(2) = 26.88, p < .001$; and in the explicit verbal with no picture condition, χ^2 $(2) = 17.14, p < .001$. Thus, our prediction that 2-year-olds hearing explicit verbal information would perform above chance when also shown a picture was confirmed, while our prediction that they would not do so if not shown a picture was disconfirmed (H2b).

Three-year-olds also succeeded at a rate above chance in both of the conditions in which they were tested: the explicit verbal information with no picture condition (x^2 $(2) = 242.64, p < .001, \omega = 2.46$), supporting H1b, and the implicit verbal information with no picture condition ($x^2(2) = 106.08, p < .001, \omega = 1.61$), about which we made no strong prediction.

To summarize, Study 2 results showed that contrary to expectation, neither the explicitness of verbal information (H1a) nor the presence of a supportive picture (H2a, H2b) significantly affected 2-year-olds' updating ability. They performed similarly across conditions, although as anticipated, older 2-year-olds performed better than younger 2-year-olds regardless of the explicitness of the verbal information and the presence or absence of a picture (H3). As predicted, explicitness of verbal information influenced 3-year-olds' updating ability, such that 3-year-olds who heard explicit verbal information without a picture had higher odds of success than 3-year-olds who heard implicit verbal information without a picture (H1a). Also as predicted, 3-year-olds given no picture performed above chance when they received explicit verbal information (supporting H1b) and, impressively, also when they received only implicit verbal information, about which we made no strong prediction. Their performance did not improve as a function of their age in months (failing to support H3 for this age group). Older 3-year-olds' improved odds of success on the pool story, but not on the dog story, may reflect more frequent encounters with mud than with pools; in previous work, the strength of a child's representation has been demonstrated to influence their ability to update (Galazka & Ganea, 2014). Future research could control for the role of parent-reported child experience.

Overall, these results indicate impressive and rapidly developing narrative situation model competence on the part of 2-year-olds when their updating is supported by an informative picture. They also demonstrate that even without pictorial support, 3-year-olds are able to update a narrative situation model using only implicit information, though explicit information improves their performance.

General discussion

We investigated for the first time whether young children aged 2 and 3 demonstrate an ability to update a situation model from an unfamiliar story narrative. We asked whether children's mental representation of the physical state of a story's protagonist (wet or muddy) used information drawn from two systematically manipulated sources: verbal statements that were provided either implicitly or explicitly in the story, and an informative picture. Two- and 3-year-olds updated their mental representation of the protagonist at a rate greater than chance, even when the verbal information provided was implicit rather than explicit. This success in using implicit information for updating in a story context was demonstrated regardless of the type of picture accompanying the verbal information: an informative picture explicitly depicting a physical substance involved in the event that was causally central to the outcome (water or mud: Studies 1 and 2), an uninformative picture that did not show this substance (Study 1), or – in the case of a condition presented only to 3-year-olds – no picture (Study 2). These findings provide evidence that children as young as 2 are beginning to integrate not only explicit, but also implicit information into their situation model of a story at a critical juncture when an inference is afforded, and that the accompanying pictorial information need not be maximally informative in order for this to take place; indeed, that 3-year-olds can manage this feat with no supporting pictorial information at all. The findings do not, of course, provide evidence that children of this age would draw the intended elaborative inference from our story when they are not prompted by a question such as that offered at test ("What does [protagonist] look like now?"). However, given the benefits conferred to children's comprehension when adults engage with children in scaffolded dialogic reading (Whitehurst et al., 1988; for a review, see Mol et al., 2008) and the benefits to children's learning when adults scaffold children's engagement with video (Strouse et al., 2013), the ability to use even 'cued' inferences to update a situation model is likely to be supportive of children's developing understanding of narrative.

Nevertheless, effects of both age and informativeness of verbal information were clearly apparent. Taken as a group, children's age in months significantly predicted their performance in Study 1 regardless of the explicitness of the verbal and visual information that they received, and this remained true of 2-year-olds in Study 2. This gradual improvement with age speaks against an account in which children succeed by using mere associations between linguistic references to the transformation of the protagonist (e.g., 'water' 'wet', 'swimming') and the image of a wet protagonist at test; on such an account, even the youngest 2-year-olds should find the task straightforward. Three-year-olds' performance in Study 2 stabilised, such that in this group there was no effect of age in months. However, in the absence of a picture, 3-year-olds' performance was significantly less robust when the verbal information they received was implicit than when hearing explicit verbal information. This suggests that implicit verbal information embedded in a narrative, when provided without pictorial support, presents challenges for young children's updating. It seems unlikely that this simply reflects a broader cognitive difficulty with updating a mental representation from information that is not explicit, since, as discussed, it has long been established that young children are able to update their mental representation of a situation that is largely imaginary, using information that is merely implied (Harris et al., 1996).

The current study used narrated pictorial stories were presented via digital media, meaning that children had to draw an inference without social input while avoiding activating representations of irrelevant actions (in the current study, Tom's opportunity to paint in the Pool Story, and Sally's opportunity to eat ice-cream in the Dog Story). These aspects of updating a situation model are not required during in-person verbal updating (e.g., Ganea et al., 2007), nor during pretend transformations in which the implied actions are physically acted out in front of the child by a social partner as they unfold (Harris et al., 1996). Across Studies 1 and 2, a substantial proportion of 2-year-olds appeared to have difficulty in suppressing representations of irrelevant actions, choosing an inaccurate representation of the protagonist's current state (e.g., Tom covered in paint). The nature of the distractor used in this study contrasts with the nature of the distractor used in both the work of Harris and colleagues (in which the distractor option was not mentioned) and with live interaction updating paradigms to date, in which children are required to suppress the activation of an irrelevant character, rather than an irrelevant action. For instance, Ganea et al. (2007) introduced children to Lucy, a toy frog, and to a nameless toy pig. Once out of sight of Lucy, children were told verbally that Lucy is 'all covered in water'; the distractor offered at test was a wet toy pig. a used a similar procedure in studying children's updating of imaginary states of affairs. Given the relative simplicity of distinguishing individual identities when compared with the complexity of human goals, which can change for a wide variety of reasons, the activation of actions that are irrelevant to the story may be more difficult for young children to suppress than the activation of an irrelevant character. Picture books for young children tend to establish narrative coherence through goal-directed action (Nicolopoulou, 2008). Thus, it is possible that in a narrative context, goal information that becomes irrelevant is particularly disruptive of young children's updating, their situation model building, and ultimately their story comprehension. Further work will be required in order to test this account. For instance, the nature of the information to be updated, and thus the distractor, could be manipulated while also investigating inhibitory control as a potential cognitive correlate of success.

Future work should also attempt to draw out the particular picture/text relations that support updating most effectively in a narrative context, both during the unfolding of the story and at the critical juncture when an inference is required. This work might usefully draw upon the category system for picture-text relationships developed by Martinez & Harmon (2012), which develops the notion of the informativeness of a picture within a story by examining not only how much critical information about character, plot, setting, and mood is conveyed by the picture, but also the degree of interdependence of picture and text in terms of the extent to which both are necessary in order to gain information considered critical to the story. For instance, the plot of a story in which a fox attempts to catch an oblivious chicken, and is continually foiled (Hutchins, 1968), is categorised as 'picture only' by dint of the fact that the fox and its thwarted threats to the survival of the chicken are only ever evident in the pictures. They are never referred to in the text, in which only the activities of the chicken are described. By contrast, a story in which the text informs readers that that a small boy has brought friends to see the king and queen, revealing *which* friends (animals) only in pictures (de Regniers, 1964), has a plot categorised by Martinez & Harmon (2012) as 'interdependent': that is, both picture and text are needed in order to understand it.

While the target inferences in our stories concerned physical changes that were salient to children, drew attention to the relations between characters and objects, and were within children's prior experience of the world (getting messy), future studies may reward attention to the developmental trajectory of situation model construction that requires updating of different types of transformations and events, including mental states. There is some evidence that even when provided with explicit information, 3-year-olds cannot yet take the mental perspective of a character who is thinking about a place (O'Neill & Shultis, 2007), nor attribute feelings to a story character that contradict their own privileged knowledge about upcoming story events (for instance, that the wolf, not grandma, is waiting for Red Riding Hood: Ronfard & Harris, 2013), suggesting that this question may reward future investigation in older 3- and 4-year-old children.

To gain insights into very young children's situation model updating is to understand how they come to know aspects of the world through narrative. Future research that seeks to clarify the nature and time course of the encoding and activation processes involved in preschoolers' narrative comprehension will move our understanding of children's early narrative comprehension beyond the *what* and towards the *how* of children's creation of meaning-based representations from narratives.

CRediT authorship contribution statement

Ruth Lee: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Patrycia Jarosz:** Visualization, Project administration, Investigation, Conceptualization. **Patricia A. Ganea:** Writing – review & editing, Visualization, Software, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada under grant number RSD1270194 to P. Ganea, and by several awards by the Province of Ontario and University of Toronto to R. Lee. We thank the parents and children who participated in this research. We are grateful to the Ontario Science Centre and to several Toronto-area daycares for their invaluable support in finding participants. We thank members of the Language and Learning Lab for their help with creating the stimuli, testing and coding. We also thank Gabrielle Strouse and Angela Nyhout for feedback on earlier versions of this manuscript.

Data availability

Data and materials are available at osf.io/8jf39

References

Apperly, I. A., Williams, E., & Williams, J. (2004). Three-to four-year-olds' recognition that symbols have a stable meaning: Pictures are understood before written words. *Child Development, 75*(5), 1510–1522. https://doi.org/10.1111/j.1467-8624.2004.00754.x

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications Board task force report. *American Psychologist, 73*(1), 3.

Barr, R. (2019). Growing up in the digital age: Early learning and family media ecology. *Current Directions in Psychological Science, 28*(4), 341–346. https://doi.org/10.1177/0963721419838245

Berman, R.A. (2009). Trends in Research on Narrative Development. In Foster-Cohen, S. (Ed.), *Language Acquisition* (pp. 294-318). Palgrave Advances in Linguistics, Palgrave Macmillan. doi: 10.1057/9780230240780_13.

DeLoache, J. S. (1997). Rapid change in the symbolic functioning of very young children. *Science, 238*, 1556–1557. https://doi.org/10.1126/science.2446392

DeLoache, J. S. (2011a). Becoming symbol-minded. *Trends in Cognitive Sciences, 8*(2), 66–70. https://doi.org/10.1016/j.tics.2003.12.004

DeLoache, J. S. (2011b). Early development of the understanding and use of symbolic artefacts. *The Wiley-Blackwell Handbook of Childhood Cognitive Development, 312–336*. https://doi.org/10.1002/9780470996652.ch10

de Regniers, B. S. (1964). May I bring a friend? B. Montresor, illus. New York, NY: Aladdin.

Fecica, A. M., & O'Neill, D. K. (2010). A step at a time: Preliterate children's simulation of narrative movement during story comprehension. *Cognition, 116*(3), 368–381. https://doi.org/10.1016/j.cognition.2010.05.014

Florit, E., Roch, M., & Levorato, M. C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes, 48*(2), 119–138. https://doi.org/10.1080/0163853X.2010.494244

Galazka, M. A., & Ganea, P. A. (2014). The role of representational strength in verbal updating: Evidence from 19-and 24-month-olds. *Journal of Experimental Child Psychology, 121*, 156–168.

Ganea, P. A., & Harris, P. L. (2010). Not doing what you are told: Early perseverative errors in updating mental representations via language. *Child Development, 81*(2), 457–463. https://doi.org/10.1111/j.1467-8624.2009.01406.x

Ganea, P. A., & Saylor, M. M. (2013). Representational constraints on language development: Thinking and learning about absent things. *Child Development Perspectives, 7*(4), 227–231. https://doi.org/10.1111/cdep.12045

Ganea, P. A., Shutts, K., Spelke, E. S., & DeLoache, J. S. (2007). Thinking of things unseen: Infants' use of language to update mental representations. *Psychological Science, 18*(8), 734–739. https://doi.org/10.1111/j.1467-9280.2007.01968.x

Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*(1), 115–140. https://doi.org/10.1146/annurev.psych.59.103006.093659

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26*(1), 69–83. https://doi.org/10.1016/0749-596x(87)90063-5

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371. https://doi.org/10.1037/0033-295X.101.3.371 ·

Harris, P. L., Kavanaugh, R. D., & Dowson, L. (1996). The depiction of imaginary transformations: Early comprehension of a symbolic function. *Cognitive Development, 12*(1), 1–19. https://doi.org/10.1016/s0885-2014(97)90028-9

Højsgaard, S., Halekoh, U., & Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software, 15*, 1–11. https://doi.org/10.18637/jss.v015.i02

Hutchins, P. (1968). *Rosie's walk*. New York, NY: Macmillan.

Jolley, R. (2008). Children's understanding of the dual nature of pictures. In C. Lange-Küttner, & A. Vintner (Eds.), *Drawing and the non-verbal mind: A life-span perspective* (pp. 86–103). New York, NY, US: Cambridge University Press.

Kaefer, T., Pinkham, A. M., & Neuman, S. B. (2017). Seeing and knowing: Attention to illustrations during storybook reading and narrative comprehension in 2-year-olds. *Infant and Child Development, 26*(5), e2018.

Kendeou, P., Bohn-Gettler, C., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*(3), 259–272. https://doi.org/10.1111/j.1467-9817.2008.00370.x

Martinez, M., & Harmon, J. M. (2012). Picture/text relationships: An investigation of literary elements in picturebooks. *Literacy Research and Instruction, 51*(4), 323–343. https://doi.org/10.1080/19388071.2012.695856

Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. (2008). Added value of dialogic parent–child book readings: A meta-analysis. *Early Education and Development, 19*(1), 7–26. https://doi.org/10.1080/10409280701838603

Nicolopoulou, A. (2008). The elementary forms of narrative coherence in young children's storytelling. *Narrative Inquiry, 18*(2), 299–325. https://doi.org/10.1075/ni.18.1.07nic

O'Neill, D. K., & Shultis, R. M. (2007). The emergence of the ability to track a character's mental perspective in narrative. *Developmental Psychology, 43*(4), 1032–1037. https://doi.org/10.1037/0012-1649.43.4.1032

Özdemir, B., & Ganea, P. A. (2020). Variability in toddlers' ability to verbally update their mental representations of absent objects. *Journal of Experimental Child Psychology, 196*, Article 104843. https://doi.org/10.1016/j.jecp.2020

Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly, 38*(1), 36–76. https://doi.org/10.1598/rrq.38.1.3

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. *The science of reading: A handbook*, 227-247.

Rall, J., & Harris, P. L. (2000). In Cinderella's slippers? Story comprehension from the protagonist's point of view. *Developmental Psychology, 36*(2), 202–208. https://doi.org/10.1037//0012-1649.36.2.202

Ronfard, S., & Harris, P. L. (2013). When will little red riding hood become Scared? Children's attribution of mental states to a story character. *Developmental Psychology, 50*(1), 283–292. https://doi.org/10.1037/a0032970

Rudin, S. (2024). Concept, ritual, and narrative: What are toddler books? *Early Childhood Education Journal, 52*(7), 1349–1359. https://doi.org/10.1007/s10643-023-01480-0

Seger, B. T., Wannagat, W., & Nieding, G. (2019). How static and animated pictures contribute to multi-level mental representations of auditory text in seven-, nine-, and eleven-year-old children. *Journal of Cognition and Development, 1–19*. https://doi.org/10.1080/15248372.2019.1636799

Strouse, G. A., O'Doherty, K., & Troseth, G. L. (2013). Effective coviewing: Preschoolers' learning from video after a dialogic questioning intervention. *Developmental Psychology, 49*(12), 2368–2382. https://doi.org/10.1037/a0032463

Strouse, G. A., & Samson, J. E. (2021). Learning from video: A meta-analysis of the video deficit in children ages 0 to 6 years. *Child Development, 92*(1), e20–e38. https://doi.org/10.1111/cdev.13429

Tompkins, V., Guo, Y., & Justice, L. M. (2012). Inference generation, story comprehension, and language skills in the preschool years. *Reading and Writing, 26*(3), 403–429. https://doi.org/10.1007/s11145-012-9374-7

Van den Broek, P., Kendeou, P., Kremer, K., Lynch, J., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. G. Paris, & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 107–130). Mahwah, N.J.: Routledge.

Van den Broek, P., & Lorch, R. F. (1993). Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes, 16*(1–2), 75–98. https://doi.org/10.1080/01638539309544830

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24*(4), 552. https://doi.org/10.1037//0012-1649.28.6.1106

Wickham, et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162. https://doi.org/10.1037/0033-2909.123.2.162