DOI: 10.1111/cdev.14089

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

# Age differences in generalization, memory specificity, and their overnight fate in childhood

**Elisa S. Buchberger**[1] | **Ann-Kathrin Joechner**[1] | **Chi T. Ngo**[1] | **Ulman Lindenberger**[1,2] | **Markus Werkle-Bergner**[1]

[1]Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

[2]Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany

**Correspondence**
Elisa S. Buchberger, Ann-Kathrin Joechner and Markus Werkle-Bergner, Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.
Email: buchberger@mpib-berlin.mpg.de, joechner@mpib-berlin.mpg.de and werkle@mpib-berlin.mpg.de

**Funding information**
Deutsche Forschungsgemeinschaft, Grant/Award Number: NG 191/2-1; Jacobs Foundation, Grant/Award Number: 2016-1217-16 and 2021-1417-99

**Abstract**

Memory enables generalization to new situations, and memory specificity that preserves individual episodes. This study investigated generalization, memory specificity, and their overnight fate in 141 4- to 8-year-olds (computerized memory game; 71 females, tested 2020–2021 in Germany). The results replicated age effects in generalization and memory specificity, and a contingency of generalization on object conceptual properties and interobject semantic proximity. Age effects were stronger in generalization than in memory specificity, and generalization was more closely linked to the explicit regularity knowledge in older than in younger children. After an overnight delay, older children retained more generalized and specific memories and showed greater gains but only in generalization. These findings reveal distinct age differences in generalization and memory specificity across childhood.

Memories serve multiple purposes. For one, we rely on memory to form and apply knowledge to guide inferences in novel situations by extracting regularities across similar or related experiences (Bauer & San Souci, 2010; Zeithamova et al., 2012; Zeithamova & Preston, 2010). In addition, memory enables the reconstruction of specific events that make up our past by conserving the idiosyncrasies of individual experiences (Tulving, 2002). These complementary memory functions are crucial for deploying knowledge and preserving an autobiographical record over the course of our lives. But how do generalization and memory specificity abilities develop during childhood, given their interdependence and complementarity?

Traditional neurocomputational models suggest a division of labor between processes that support the memory functions of generalization and memory specificity. On the one hand, the extraction of commonalities across multiple experiences is thought to be a slow process that relies on the neocortex (McClelland et al., 1995). Such extraction of recurring patterns supports inferences in novel situations by applying the learned knowledge to new stimuli or contexts (Zeithamova et al., 2012). Memory specificity, on the other hand, is a multifaceted phenomenon that relies on several hippocampus-dependent computations. First, it requires forming an episode by binding the various aspects of an event into an integrated representation (Davachi, 2006; Ranganath, 2010; Tulving, 2002) allowing for reconstructing a specific episode based on partial cues (i.e., pattern completion; Hunsaker

**Abbreviations:** CA, cornu ammonis; GDPR, General Data Protection Regulation; GLMM, generalized linear mixed-effect model; GloVe, Global Vectors for Word Representations; LMM, linear mixed-effect model; mPFC, medial prefrontal cortex; MPIB, Max Planck Institute for Human Development; REML, restricted maximum likelihood variance estimation.

Elisa S. Buchberger and Ann-Kathrin Joechner contributed equally to this work.

& Kesner, 2013; Marr, 1971). Second, it depends on the ability to discern similar or overlapping episodes to guard against interference (i.e., pattern separation; Norman & O'Reilly, 2003; Yassa & Stark, 2011). However, while generalization and memory specificity seem to play a complementary role in human memory, they should not be understood as two ends of the same continuum. That is to say, that impaired memory specificity is not the same as successful generalization behavior, as these memory functions rely on distinct neural coding mechanisms (Hunsaker & Kesner, 2013). Interestingly though, recent neurocomputational models suggest that the hippocampal computations supporting memory specificity also contribute to the ability to rapidly generalize across recently encountered related episodes—either at encoding (Shohamy & Wagner, 2008; Zeithamova & Preston, 2010) or retrieval (Kumaran & McClelland, 2012). Indeed, empirical evidence from adults supports the idea that rapid generalization success is contingent on the memory for idiosyncratic episodes (Banino et al., 2016; Koster et al., 2018; Ngo et al., 2021; Tompary et al., 2020).

The notion that rapid generalization depends on the retrieval of individual episodes is intriguing given the uneven developmental patterns between generalization and memory specificity across childhood (Keresztes et al., 2018; Newcombe et al., 2007; Ramsaran et al., 2019). Generalization behaviors are already evident early in life, such that during the first years, children exhibit prodigious abilities in acquiring general world knowledge (Newcombe et al., 2007; Piaget et al., 1977). Infants as young as 8 months can already detect regularities in continuous language input after a brief exposure (Saffran et al., 1996). At 9 months, children show flexible application of sequence knowledge across different stimuli (Lukowski et al., 2009). Toddlers extend newly acquired knowledge onto novel stimuli from 12 to 21 months (Hayne et al., 1997) and successfully generalize event knowledge across different instantiations at 16 and 20 months (Bauer & Dow, 1994). In early to middle childhood, children improve in their ability to generalize across more complex demands, such as integrating novel facts (Bauer & San Souci, 2010) or generalizing across semantic categories (Ngo et al., 2021). Despite these early capacities, the ability to acquire schematic knowledge based on complex temporal regularities (Pudhiyidath et al., 2020) and linking memories based on overlapping elements (Schlichting et al., 2017) continues to improve from late childhood to adolescence and into adulthood. Interestingly, one potential contributor to this age difference lies in the notion that younger children manage to extract regularities in their surroundings, but fail to deploy this knowledge in a novel context (Pudhiyidath et al., 2020).

Unlike for statistical learning and generalization, there is little evidence of specific episodic memories in the first 2 years of life (Newcombe et al., 2014; Pillemer & White, 1989; but see Rovee-Collier, 1997). Over the course of early to middle childhood, children show remarkable improvements in binding and pattern completion capacities—that is, accurately retrieving the associated pairmate or context in the presence of a cue (e.g., Lloyd et al., 2009; Ngo et al., 2018; Riggins, 2014; Sluzenski et al., 2006). Pattern separation also strengthens over the same period: There is a strong tendency to confuse similar items with one another around age four, followed by remarkable age-related improvements in memory discrimination for similar objects (Canada et al., 2019; Ngo et al., 2018; Rollins & Cloude, 2018) and contexts (Lindsay et al., 1991; Ngo, Lin, et al., 2019) throughout early and middle childhood.

Given that generalization and memory specificity have been studied using vastly different paradigms and age windows, the relative associations with age and their interdependence across age are not well understood. Thus far, only one study showed that unlike in adults, young children's generalization success was less dependent on their memories of contextual information of individual episodes. Instead, children relied more on conceptual memories of individual items and the semantic structure that ties together the related episodes (Ngo et al., 2021). The robustness of these initial findings remains to be tested through replication efforts in an independent sample.

While rapid generalization and memory specificity both improve over childhood, less is known about whether children differentially consolidate generalized and specific memories over time. Both kinds of memories require processes that stabilize and integrate representations of the underlying experiences into existing memory networks, thereby shielding them from forgetting. Models of systems consolidation assume that repeated reactivation transforms initially labile, hippocampus-dependent mnemonic representations into more durable, integrated neocortical memories (Diekelmann & Born, 2010; Frankland & Bontempi, 2005; McClelland et al., 1995). Such consolidation processes not only support memory retention but may also facilitate the emergence of previously inaccessible memories—especially across-episode generalized knowledge (i.e., memory gain; Dumay, 2016; Fenn & Hambrick, 2013; Landmann et al., 2014; Lewis & Durrant, 2011). Systems consolidation can, in principle, take place over a period of wakefulness, but seems to be especially effective during sleep (Diekelmann & Born, 2010). While previous evidence points toward the beneficial effects of a sleep-filled delay on both generalization and memory specificity success when studied separately (e.g., Durrant et al., 2011; Ellenbogen et al., 2007; Friedrich et al., 2015; Hanert et al., 2017), only a handful of studies have investigated the time-dependent fate of memory for generalizations and specific details in tandem over a sleep-filled delay (Chatburn et al., 2021; Friedrich et al., 2020; Lau et al., 2011). These few studies propose quite opposing conclusions about the role of sleep for generalization and memory specificity, but notably in vastly different age groups. For instance, in adults, it has been suggested that specifically generalization processes might be facilitated by memory reactivation during sleep (Witkowski et al., 2021), while in infants the consolidation of specific episodic memories during sleep was found to rather disable their semantic processing (Friedrich et al., 2020). Further, inspired by traditional consolidation theories, recent studies in adults have fostered the idea that generalized knowledge comes at the cost of detailed memory (Richards et al., 2014; Tompary & Davachi, 2017; Witkowski et al., 2021). It remains unclear how age affects retention and gain of specific and generalized memories over a sleep-filled delay, and whether they occur in a trade-off fashion across different ages.

To better characterize the co-development of generalization and memory specificity from early to middle childhood, we set out to replicate two previous findings from Ngo et al. (2021). First, with a much larger sample size, we tested whether generalization and memory specificity varied with age from early to middle childhood. Second, we aimed to replicate the finding that children draw on conceptual memories of objects from past episodes and the semantic proximity between them and not on the rich contextual details surrounding the individual episode. We extended these findings with four specific questions. First, are improvements with age more pronounced in generalization than in memory specificity? Second, how is the ability to extract regularities linked to the ability to deploy this knowledge in novel situations? Third, do retention and gain of generalized and specific memories after an overnight delay vary with age? And last, on an exploratory level, is there a generalization-specificity trade-off after a sleep-filled delay in children? To address these questions, we adapted the behavioral paradigm from Ngo et al. (2021), which co-assesses generalization and multiple aspects of memory specificity immediately after learning and again in the morning after one night of sleep.

# METHOD

## Participants

A total of 146 German-speaking children with no signs of non-normative development between the ages of four and eight participated in the study ($M$ = 77.72 months, $SD$ = 19.81 months). This sample size was determined based on a previous study by Ngo et al. (2021) on 70 children. To increase statistical power, we here doubled the sample size. Given the global COVID-19 pandemic, children were either tested in-person ($n_{\text{in-person}}$ = 58) at the Max Planck Institute for Human Development, Berlin, Germany (MPIB), or virtually via the online meeting software GoToMeeting (©LogMeIn; $n_{\text{virtual}}$ = 88; in-person testing commenced in January 2020 and had to be put on hold in March 2020; online testing was resumed in November 2020 and completed in May 2021, for more descriptive information on the sample see Table S1). All behavioral performances were comparable between the two test formats (see Tables S2 and S3; Figure S1 for comparisons) and thus data from the two groups were collapsed in all subsequent analyses. Data from three children were excluded due to technical issues ($n$ = 1), experimenter error ($n$ = 1), or incomplete participation ($n$ = 1). Two additional children (49.64 and 50.30 months) showed chance-level performance across all subtasks (33%), suggesting a lack of attention or task procedural comprehension. These children were not included in the analyses. The final sample included 141 children (71 females, $M_{\text{age}}$ = 77.98 months, $SD_{\text{age}}$ = 19.61 months, $n_{\text{online}}$ = 86). Participants were recruited via the participant database of the MPIB and were screened through parental reports to ensure that they had no known chronic diseases, diagnosed sleeping, psychiatric, neurological, or learning disorders. Adhering to the principle of data minimization of the EU General Data Protection Regulation (GDPR), no demographic information other than age and sex was collected. Written informed consent was obtained from a legal guardian prior to the study. Verbal assent was obtained from the children at the beginning of each test session. The study was approved by the local ethics committee of the MPIB.

## Overall procedure

The study consisted of two behavioral sessions on consecutive days (see Figure 1). Session 1 took place in the evening, approximately 2–3 h before each child's habitual bedtime, and encompassed encoding and immediate test phases. In Session 2, children were tested in the morning of the following day, approximately 1.5–2.5 h after waking up (see Figure 1, Session 2; mean duration of delay = 14.88 h, $SD$ = 1.67 h). In the in-person test format, children were tested individually in a quiet room, where they sat in front of a computer screen. The experimenter controlled the tasks from a second screen, which was separated by a divider, and recorded children's responses on a keyboard. The online test format followed the same procedure with two deviations. First, experimenters shared their screens virtually so that children saw the same content as in the in-person setup on their own screens at home, along with the video of the experimenter. Second, brief technical checks preceded the experiment to ensure stable audio and video connections.

## Memory task

We administered an adapted version of the "collection game" paradigm from a previous study (Ngo

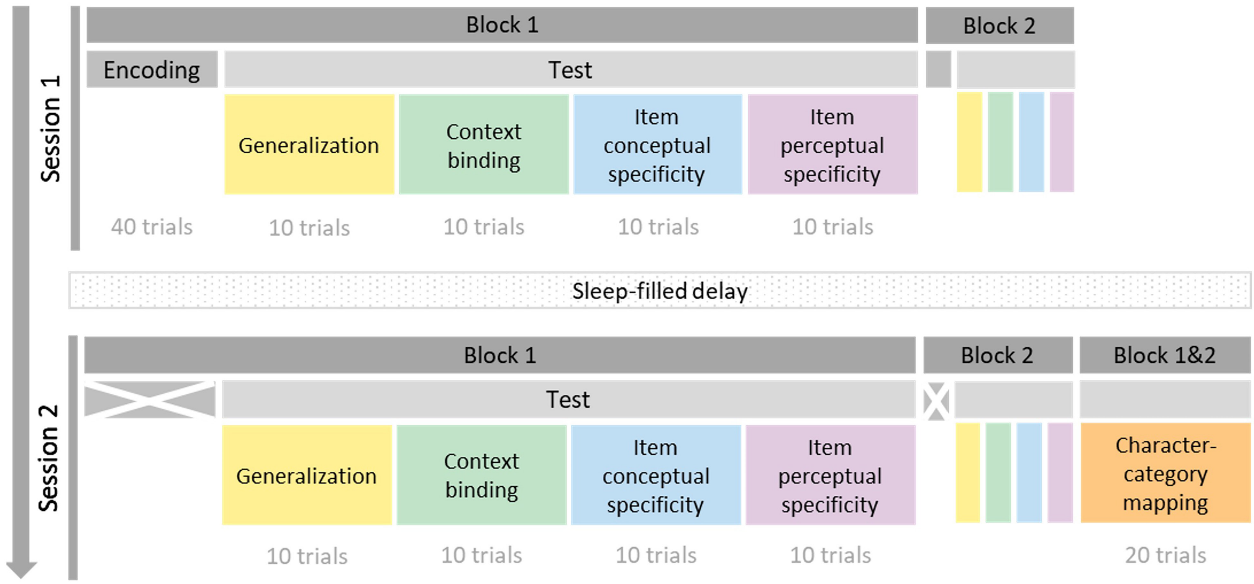


**FIGURE 1** Schematic depiction of the task procedure. Session 1 included an encoding phase and four subtasks in two blocks each (upper row). Session 2 followed the same overall structure as Session 1, with the exception that there was no encoding phase and that the character-category mapping was added at the end of the session (lower row). The two sessions were separated by one night of sleep.

et al., 2021). Similar to the previous study, the task was divided into two encoding-test blocks, which followed an identical procedure but with unique sets of stimuli (see Figure 1). Each block consisted of a character familiarization phase, followed by an encoding phase (Figure 2a) and four three-alternative forced-choice memory subtasks in a fixed order: generalization, context binding, item conceptual specificity, and item perceptual specificity (see Figure 2b). The two study sessions were separated by an interval of nocturnal sleep, which varied in length depending on the participant's habitual sleep schedule. Both sessions followed the same procedure, with three exceptions: (i) children were given a brief training phase to acquaint with the task procedure at the beginning of Session 1 but not in Session 2, (ii) the encoding phase was only presented in Session 1, and (iii) children were only administered a character-category mapping task at the end of Session 2. The average duration of Session 1 was approximately 0.83 h (min = 0.42 h, max = 1.5 h, *SD* = 0.19 h), and the duration of Session 2 was approximately 0.81 h (min = 0.42 h, max = 1.25 h, *SD* = 0.17 h).

## Materials

The behavioral paradigm was implemented using MATLAB version R2016b (MathWorks, Natick., 2016) and the Psychophysics toolbox version 3.0 (Brainard, 1997). The stimulus material for the “collection game” was adapted from the original paradigm in Ngo et al. (2021). The adaptation to the stimuli was necessary to ensure that (i) the materials were well suited for German children and (ii) there would be a sufficient number of stimuli for the delayed test session. The stimuli were cartoon images of 20 androgynous cartoon characters, 80 scenes, and 280 line-drawn common objects. In this study, 20 categories of semantically congruent objects and scenes were selected, with 18 out of 20 categories matching the categories in the original task, which had been identified based on their familiarity with young children (BatMon II, Price & Connolly, 2006) and validated in an additional sorting task (Ngo et al., 2021; e.g., furniture, kitchenware, musical instruments; see Figure S2 for mean accuracy per category, grouped by age). Each category consisted of 14 line drawings of objects (e.g., piano, guitar). There were four colored versions of each object image: one in black-and-white and three in distinct, naturalistic colors (manually painted using Adobe Photoshop). Each category contained four semantically congruent scenes. Each of the characters was randomly assigned to one of the categories and placed in their four respective scenes. Hence, the scenes form the context in which a specific character-object combination was encountered. The presentation of the 20 characters was divided into two blocks with 10 characters assigned to each block (40 encoding trials per block).

## Familiarization phase

First, children were told that they would meet many new friends in the “collection game.” At the beginning of each of the two blocks, they were introduced to 10 characters sequentially. Each character was presented in the center of the screen for 3 s and the name of the character was verbally introduced by the experimenter (e.g., “This is Luntik”).

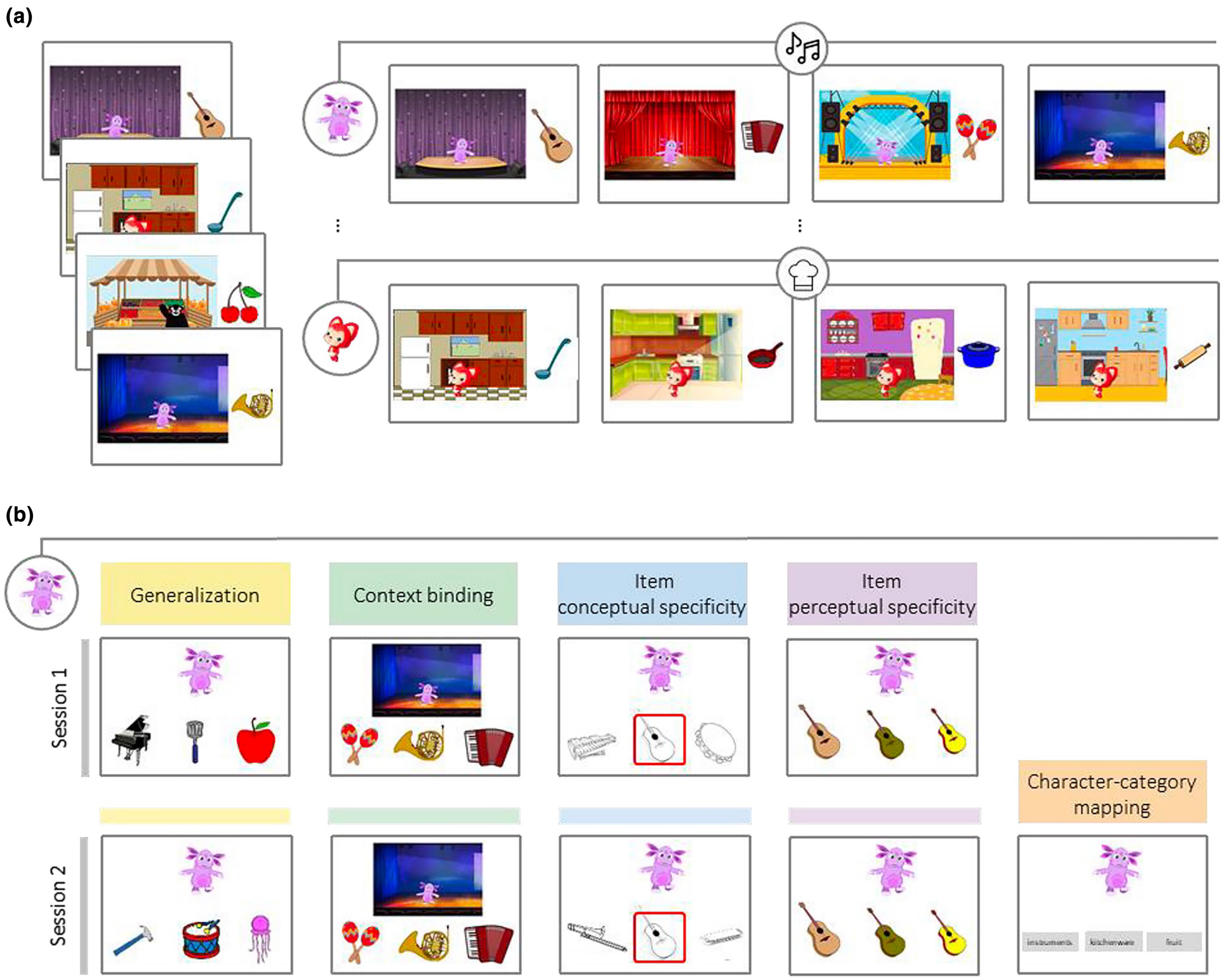


**FIGURE 2** (a) Schematic depiction of the encoding phase during Session 1. For visualization purposes, two sample characters are shown. Each character was presented in four different contexts along with four semantically related items throughout encoding (e.g., musical instruments in the top row and kitchen items in the bottom row). Encoding trials of the same character were separated by a minimum of two trials of other characters (as illustrated on the left). (b) Schematic depiction of the different subtasks, here shown for one example character. Immediately after encoding, children were tested on generalization, context binding, item conceptual specificity, and item perceptual specificity (top row) and again at the delayed test after one night of sleep (bottom row). The delayed test at Session 2 further included a character-category mapping test at the end.

## Encoding phase

Children were instructed that each character would go to different places and collect the things they wanted for their collection. They were told to pay attention to the places and objects paired with each character. In each block, the encoding phase entailed 40 trials, each consisting of a character in their context presented on the left side of the screen, along with an object presented on the right side of the screen (Figure 2a). A given character–context–object combination was presented for 5 s, with a 0.5 s intertrial interval. Throughout the encoding phase, each character appeared in four trials in an intermixed order, each time with a different context and object. The order of encoding trials within each block, as well as the block order, was randomized across participants, with the constraint that trials of a given character were separated by a minimum of two trials of other characters.

## Test phase

All subtasks were self-paced three-alternative forced-choice tasks. All test items only appeared once at each test session. The assignment of stimuli as encoding, target and lure objects was randomized across participants. On all subtasks, children responded verbally or pointed to the correct option on the screen.

### Generalization

There were 10 test trials in each block (one per character). Children were instructed to choose a new object to add to each character's collection. On each trial, children saw each character on the top of the screen, along with three unstudied objects, and were asked to choose the correct one. The three options consisted of a target, defined as an unstudied object from the semantic category assigned to the character, and two lure items, defined as unstudied objects from semantic categories assigned to other characters. All test objects for a given character differed between Sessions 1 and 2 for each child.

### Context binding

This subtask assessed children's memories of the specific item-context associations. There were 10 test trials in each block (one per character). On each trial, children were shown a character presented on one of the four contexts seen at encoding on top of the screen, along with three objects. The three options consisted of a target, defined as an object that had been paired with the character in the given context, and two lures, defined as objects that had been paired with the respective character in a different context. Children were asked to choose the correct object that the given character had collected in this context. All context-binding test trials were identical between Sessions 1 and 2 for each child.

### Item conceptual specificity and item perceptual specificity

The item conceptual specificity and item perceptual specificity tasks were linked, such that the item perceptual specificity trial immediately followed the item conceptual specificity trial for each character. The item conceptual specificity task assessed children's memories of the specific objects seen at encoding. There were 10 test trials in each block (one per character). On each trial, children were shown a character presented on the top of the screen, and three black-and-white drawings of objects beneath. The three options included a target, defined as the object that had appeared at the encoding phase, and two within-category lures, defined as unstudied objects, but both were semantically related to the target (e.g., unstudied music instruments). Children were asked to choose the object that had actually been collected by the given character and immediately received corrective feedback. If correct, the experimenter said "That's right!" and proceeded to the item perceptual specificity trial of that object. The item perceptual specificity task assessed children's memories of the specific perceptual attributes of learned objects. Here, they were asked to choose the object that looked exactly like the one that they had seen at encoding and were shown three options: a target, defined as an identical image to the one seen at encoding, and two lures, defined as object images in different colors. If children's response on the item conceptual specificity trial was incorrect, the target was highlighted in a red box and children were told "You actually saw this object instead." Then they proceeded to the item perceptual specificity trial as described above. The test trials were identical between Sessions 1 and 2 for the item perceptual task, while there were novel lure items for the item conceptual task during Session 2 for each child.

### Character-category mapping

To complement our measure of generalization abilities, we tested children's knowledge of the character-category associations at the end of Session 2. Note that this subtask could not have been administered at the end of Session 1 due to a potential contamination to the delayed generalization performance. There were 20 test trials (one per character). On each trial, children were shown a character on top of the screen and were asked which kind of things the given character had collected the day before. The three options included one target, defined as the semantic category that had been paired with this character, and two lure categories, defined as semantic categories that had been paired with other characters. The options were presented in written form (e.g., the words "instruments," "kitchenware," "fruit") and were read out one by one by the experimenter.

## Training phase

To ensure procedural understanding of the task, children completed a training phase prior to the first block at Session 1. For this, they were introduced to two example characters, which were not part of the main task. Children were presented with four interleaved training-encoding trials for each character, where they saw the character placed onto a context scene (e.g., a beach), paired with an object (e.g., a parasol), just like in the main experiment. After the mock-up encoding phase, the training phase proceeded exactly like the test phase in the main experiment, with the exception that experimenters provided corrective feedback on each trial. None of the stimuli from the training phase overlapped with those in the main task.

## Behavioral analyses

Memory performance was quantified separately for each subtask as accuracy, defined as the proportion of correct target selection relative to the total number of test trials. Immediate performance was calculated for Session 1. We assessed the effect of an overnight delay on an item level as *relative retention* and *relative gain* for each subtask separately (cf. Dumay, 2016; Joechner et al., 2021; Muehlroth et al., 2020). Memory retention was operationalized as the number of items correctly remembered in Sessions 1 and 2, relative to all items correctly remembered during Session 1. Memory gain was operationalized as the number of items correctly remembered in Session 2 that had not been correctly remembered in Session 1, relative to the number of incorrect trials in Session 1. Hence, memory gains reflect items that were inaccessible before but accessible after a delay (Dumay, 2016; cf. Habib & Nyberg, 2008). Mean accuracies for the immediate and delayed tests per age group can be found in Figure S3. However, all analyses on the overnight fate were performed on relative retention and gains, as this reflects a more nuanced analysis on an item level. Note that as corrective feedback was provided during the item conceptual specificity task in Session 1, this subtask was excluded for all retention and gain analyses.

All statistical analyses were conducted in R 4.0.3 (R Core Team, 2020) using RStudio 1.1.383 (RStudio Team, 2020). For linear mixed-effect models (LMM), we used the lme function from the nlme package (Pinheiro et al., 2021) and for generalized LMMs (GLMM) the glmer function from the lme4 package (Bates et al., 2015). To specify the effects of subtask, we used treatment contrast coding, with the generalization task as the reference group. Age was treated as a continuous variable. All final LMMs were fit using restricted maximum likelihood variance estimation (REML) and the nlminb optimizer. GLMMs were fit by maximum likelihood laplace approximation and using the bobyqa optimizer. Since significant effects for subtasks in the LMMs reflect a difference between each of the specificity subtasks compared to the generalization subtask, respectively, effects for each subtask were further investigated through post-hoc simple linear regressions. *P*-values were corrected for multiple testing according to the Bonferroni method ($p_{adj}$) through multiplication by the number of individual tests (Bland & Altman, 1995). Hence, also $p_{adj}$ was compared to an alpha of 0.05.

The study protocol and the analyses were not preregistered but closely followed previous work. Hence, the present study cannot be described as strictly confirmatory, but it is also not fully exploratory, as it builds on (and supports) strong assumptions from prior work. In sum, we would consider the present study to be located somewhere in the middle of the continuum between confirmatory and exploratory.

# RESULTS

## Differential age effects on generalization versus memory specificity

First, we asked whether the performance on the four different subtasks differentially varied with participants' age. We computed LMMs to predict immediate memory performance by age and subtask, allowing for a random intercept per participant to account for the dependency within participants:

$$\text{Accuracy} \sim \text{subtask} \times \text{age} + (1|\,\text{ID})$$

A likelihood-ratio test indicated that a full model including all simple and interaction effects between age and subtasks provided a better fit than a model without the interaction between age and subtask ($\chi^2(1) = 74.63$, $p < .001$). Importantly, this suggests that the type of subtask significantly moderated the effect of age on performance. The results of this final model showed that the positive age effect on accuracy was more pronounced for generalization compared to all other subtasks (context binding × age: $\beta = -.07$, $SE = 0.02$, $t = -4.91$, $p < .001$; item conceptual specificity × age: $\beta = -.08$, $SE = 0.02$, $t = -5.09$, $p < .001$; item perceptual specificity × age: $\beta = -.13$, $SE = 0.02$, $t = -8.94$, $p < .001$; Figure 3). To examine the effect of age on accuracy in every subtask, we further conducted four post-hoc simple linear regression analyses, one for each subtask

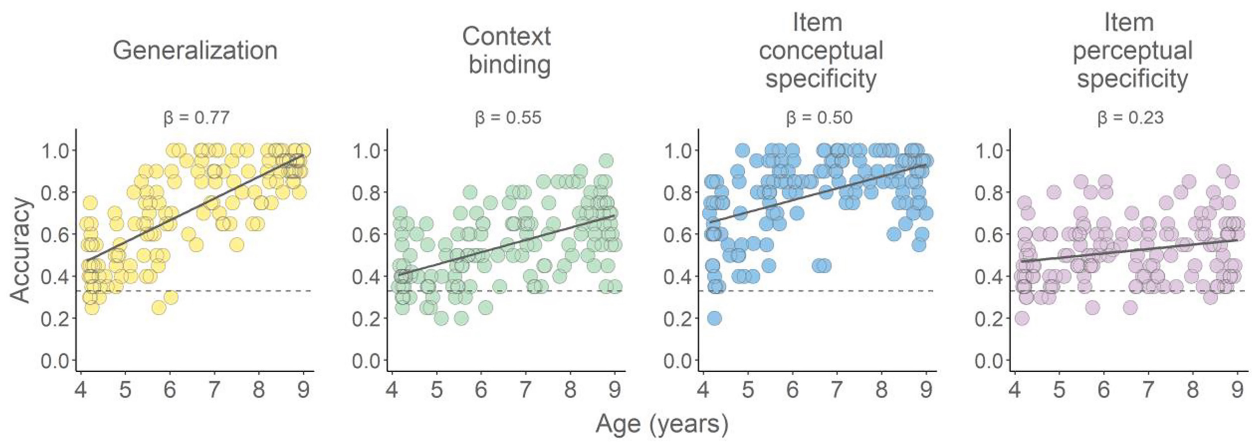


**FIGURE 3** Memory accuracy on the generalization, context binding, item conceptual specificity, and item perceptual specificity tasks for individual participants. Each dot represents a participant. Chance level of 0.33 is indicated by the dashed lines. The solid lines show the best-fitting least squares regression association between memory accuracy (plotted on the *y*-axis) and participants' age (plotted on the *x*-axis). The beta estimates indicate the standardized estimates from the post-hoc linear regressions (see Tables S4–S7).

separately. This analysis revealed a positive effect of age on each of the four subtasks (Tables S4–S7), suggesting that older children were better at generalization and all aspects of memory specificity. Taken together, we see age-related differences in all subtasks, with a particularly strong effect for generalization, over the observed age range from 4 to 8 years.

## Generalization is contingent on conceptual item specificity and within-category similarity

To better understand the interplay of rapid generalization and specificity in childhood, we then tested the interdependence of these memory functions. Specifically, we examined whether the likelihood of successful generalization performance for a given character was tethered to children's memory for the context-item associations as well as their memory for item conceptual specificity and item perceptual specificity for the objects that this character had collected.

In addition to memory specificity performance, we considered the role of semantic similarity for generalization success. Given that we used real-world objects, we asked whether within-category semantic similarity would promote generalization success. Semantic similarity was quantified by leveraging the GloVe algorithm (Global Vectors for Word Representation; Pennington et al., 2014), which derives measures of semantic similarity based on word co-occurrence statistics from a large corpus of text. To estimate the semantic closeness of within-category items, we computed the semantic similarity among the five items that each participant encountered for a given category: four shown at encoding and the target presented at the immediate generalization test (for the same approach, see Ngo et al., 2021).

We conducted a binomial GLMM with a logit link function to predict generalization success at immediate test for each of the 20 characters within each participant, with age, context binding, item conceptual specificity, item perceptual specificity, semantic similarity, age × context binding, age × item conceptual specificity, age × item perceptual specificity and age × semantic similarity as fixed effects, and category and participant as random intercepts (as each participant and each category contributed multiple data points in the analysis):

$$\begin{aligned}&\text{Generalization accuracy} \sim \text{age} + \text{context binding accuracy}\\&+ \text{item conceptual specificity accuracy}\\&+ \text{item perceptual specificity accuracy}\\&+ \text{sematic similarity} + (1|\text{ID}) + (1|\ \text{category})\\&+ \text{age} \times \text{context binding accuracy}\\&+ \text{age} \times \text{item conceptual specificity accuracy}\\&+ \text{age} \times \text{item perceptual specificity accuracy}\\&+ \text{age} \times \text{sematic similarity}\end{aligned}$$

We found that generalization success was significantly associated with age ($\beta = .84$, $z = 5.68$, $p < .001$), item conceptual specificity ($\beta = .48$, $z = 3.44$, $p < .001$), and semantic similarity ($\beta = .15$, $z = 2.08$, $p = .038$; see Figure 4). This suggests that older children showed higher levels of generalization success, in line with our results on generalization accuracy on a participant level (see section “Differential age effects on generalization versus memory specificity”). Further, memory for object identities was associated with the likelihood of correct generalization for the respective category. Last, greater semantic relatedness among items within a semantic category was linked to greater generalization success probability. None of the other predictors or interaction terms reached significance (all $-1.40 < z < 1.72$; all $p$s $> .086$), although we note a trend for the interaction effect of item perceptual specificity and age ($\beta = .19$, $z = 1.72$, $p = .086$).

## Generalization is contingent on character-category mapping

To further characterize age differences in generalization, we investigated the link between children's explicit knowledge of the character-category associations and generalization performance. Given that the generalization subtask requires knowledge extension based on the character-category regularities, it is possible that some children were able to extract the regularities from the multiple episodes (“Luntik always collects musical instruments”), but failed to apply their knowledge to a novel situation (“…and therefore would add a piano to his collection”). We tested (i) whether children's memory for the regularities of character-category co-occurrence was associated with their generalization success, and (ii) whether this association varied with age. Note that this analysis was restricted to performances from Session 2 because the character-category mapping task was only administered at the end of Session 2 (see Figure 1). Accuracy measures of the character-category mapping subtask can be found in Figure 5a (for delayed performance on all other subtasks see Figure S4). We conducted a separate binomial GLMM to predict generalization success during Session 2 on a trial-by-trial level by age, character-category mapping, and age × character-category mapping as fixed effects allowing participants and items belonging to different categories to vary in their intercept:

$$\begin{aligned}&\text{Delayed generalization accuracy} \sim \text{age}\\&+ \text{character}-\text{category mapping accuracy}\\&+ (1|\text{ID}) + (1|\ \text{category}) + \text{age}\\&\times \text{character}-\text{category mapping accuracy}\end{aligned}$$

We found that generalization success was positively associated with age ($\beta = .57$, $z = 3.27$, $p = .001$) and

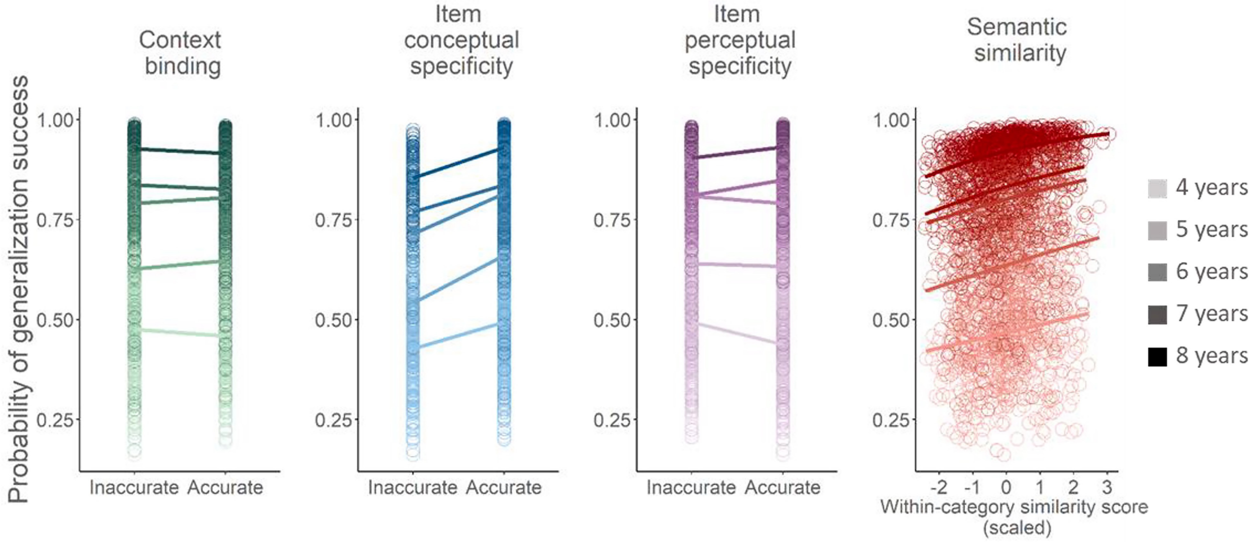


**FIGURE 4** Age patterns in the contingency of generalization on context binding, item conceptual specificity, item perceptual specificity, and within-category semantic similarity. Each plot shows the distribution of the predicted probability of generalization success ($y$-axis). For context binding, item conceptual specificity and item perceptual specificity inaccurate vs. accurate trials are plotted on the $x$-axis. For semantic similarity, the scaled within-category similarity score is plotted on the $x$-axis. For visualization purposes, the data are grouped by years of age, with color intensity representing age groups. Each dot represents one trial.

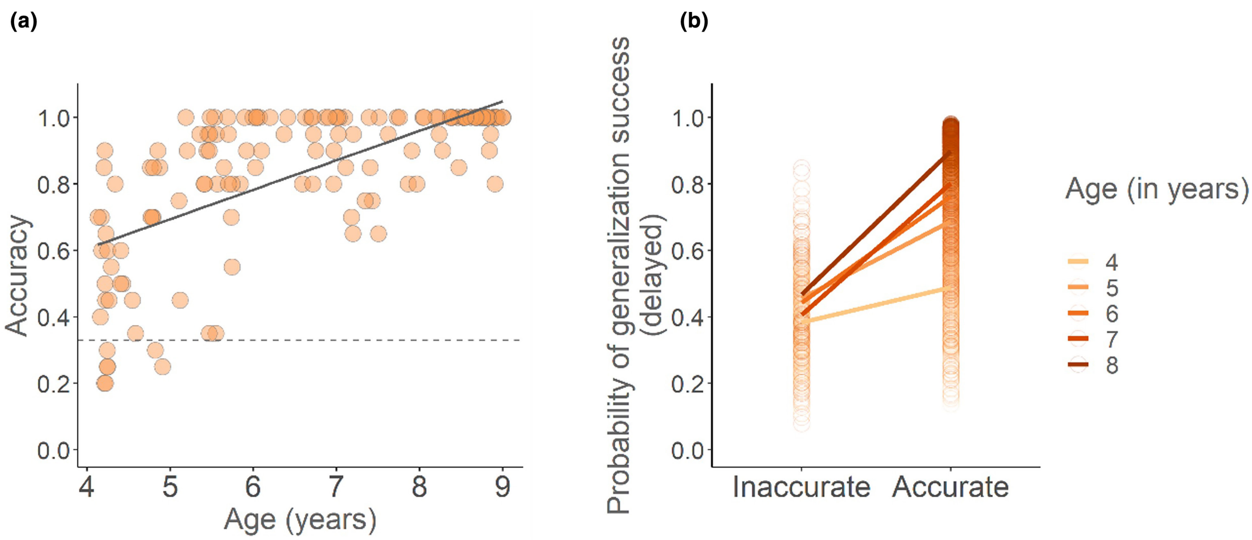


**FIGURE 5** (a) Accuracy on the character-category mapping task. The solid line shows the best-fitting least squares regression association between memory accuracy ($y$-axis) and participants' age ($x$-axis). Each dot represents a participant. Chance level of 0.33 is indicated by the dashed line. (b) Distributions of the estimated probability of generalization success on the delayed test ($y$-axis) by the character-category mapping accuracy (accurate vs. inaccurate trials, $x$-axis). For visualization purposes, the data are grouped by years of age, with color intensity representing age groups.

character-category mapping ($\beta$=.64, $z$=3.58, $p$<.001; Figure 5b). Interestingly, the age effects on generalization success interacted with character-category mapping accuracy ($\beta$=.42, $z$=2.51, $p$=.012). That is, the strength of the link between generalization success and character-category mapping differed depending on age: Generalization success was more closely associated with the explicit knowledge of character-category mapping in older children compared to younger children.

## Age effects on delayed memory retention differ between generalization and specificity

To characterize how stable children's generalized and specific memories would be over the course of a sleep-filled delay, we investigated relative memory retention from Session 1 to Session 2 (Figure 6a). As a reminder, relative memory retention was calculated as the proportion of items retained after a night of sleep among the

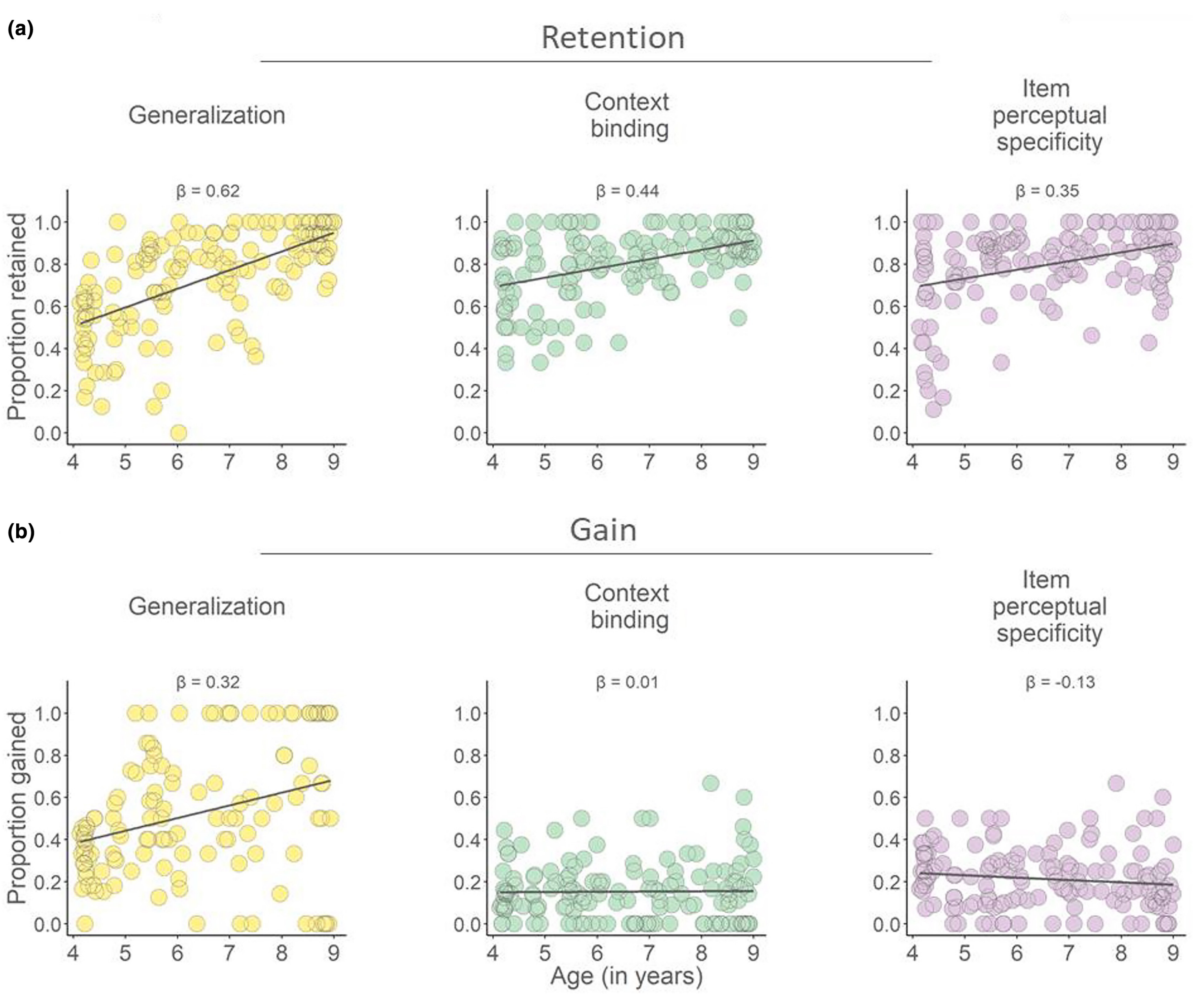


**FIGURE 6** (a) Proportion of items retained overnight for generalization, context binding, and item perceptual specificity in relation to participants' age. Ranging from 0 to 1, a perfect retention rate of 1 indicates that a child retained all of the correct items from the previous evening, whereas a retention of 0 indicates that they retained none. Each dot represents a participant. The solid lines show the best-fitting least squares regression between the proportion of retained items ($y$-axis) and participants' age ($x$-axis). The beta values indicate the standardized estimates from the post-hoc linear regressions (see Tables S9–S11). (b) Proportion of items gained overnight for generalization, context binding, and item perceptual specificity in relation to participants' age. Ranging from 0 to 1, a perfect gain of 1 indicates that a child succeeded on all of the incorrect trials from the previous night, whereas a gain of 0 indicates that those trials remained unsuccessful in Session 2. Each dot represents a participant. The solid lines show the best-fitting least squares regression between the proportion of gained items ($y$-axis) and participants' age ($x$-axis). The beta values indicate the standardized estimates from the post-hoc linear regressions (see Tables S15–S17).

successful trials in Session 1. Comparable to the previous analyses on immediate memory performance, we fit LMMs for relative memory retention, which included age and subtasks as fixed effects and allowed for random intercepts for each participant:

$$\text{Relative retention} \sim \text{subtask} \times \text{age} + (1 | \text{ID})$$

Interestingly, a likelihood-ratio test showed that the full model, including all interactions, had a significantly better fit than the corresponding model without interaction terms ($\chi^2(1)=20.20$, $p<.001$). This indicates that the age effect on retention differed across subtasks: the positive link between age and relative retention was stronger in generalization than in context binding (context binding × age: $\beta=-0.07$, $SE=0.02$, $t=-3.81$, $p<.001$) and item perceptual specificity (item perceptual specificity × age: $\beta=-.08$, $SE=0.02$, $t=-4.05$, $p<.001$). We examined the robustness of these findings by including overall memory performance (averaged immediate performance across all four subtasks) as a covariate (Table S8). This analysis showed comparable results to the main analysis, suggesting that the subtask × age interaction effect was as not only driven by differences in children's immediate performance.

Next, we also performed post-hoc simple linear regressions to test whether age significantly affected the retention in each subtask separately. Indeed, age had a positive association with relative memory retention in all three subtasks (generalization: $R^2=.38$, $F$(1,

139) = 86.25, $\beta = .62$, $t = 9.29$, $p_{adj} < .001$; context binding: $R^2 = .19$, $F(1, 139) = 32.90$, $\beta = .44$, $t = 5.74$, $p_{adj} < .001$; item perceptual specificity: $R^2 = .13$, $F(1, 139) = 19.92$, $\beta = .35$, $t = 4.46$, $p_{adj} < .001$, for full information, see Tables S9–S11). Together, these results suggest that older children retained relatively more memories than younger children across all subtasks and that this age effect was more pronounced for generalization than for memory specificity.

## Age effects on delayed memory gain differ between generalization and specificity

To further test whether there would be comparable age effects on the emergence of previously not correctly recalled generalized and specific memories over a sleep-filed delay, we conducted analogous analyses for relative memory gain (Figure 6b). As a reminder, relative memory gain was calculated as the proportion of correct responses after a night of sleep among the unsuccessful trials in Session 1. We fit a LMM for relative memory gain, which included age and subtasks as fixed effects and allowed for random intercepts for each participant:

$$\text{Relative gain} \sim \text{subtask} \times \text{age} + (1 | \text{ID})$$

A likelihood-ratio test showed that the full model, which included all interaction terms, had a significantly better fit than the corresponding model without interaction terms ($\chi^2(1) = 24.68$, $p < .001$). Again, the effect of age on relative memory gain differed by subtask, such that age had a more positive effect on relative memory gain for generalization compared to context binding (context binding × age: $\beta = -.10$, $SE = 0.03$, $t = -3.91$ $p < .001$, Figure 6b) and item perceptual specificity (item perceptual specificity × age: $\beta = -.12$, $SE = 0.03$, $t = -4.73$, $p < .001$, Figure 6b). Similar to the memory retention analyses, we tested the robustness of our findings when accounting for overall memory performance (Table S12). Given that performance during Session 1 restricted the number of items that could be gained in Session 2 and therefore led to extreme values in relative gains for high performers, we additionally repeated these analyses after excluding performances per subtask near ceiling (fewer than three items left to gain, $n = 54$; for age distribution and distribution across subtasks, see Table S13). This analysis ensures that the observed effects were not driven by close-to-perfect performance on specific subtasks in Session 1 (Table S14; Figure S5). Both additional analyses indicated comparable result patterns to the main analysis, suggesting the robustness of the effects.

Next, we tested the effect of age on relative memory gain for each subtask separately in post-hoc linear regressions. Interestingly, these analyses revealed a differentiated age pattern across the three subtasks. Age was only associated with relative gains in generalization ($R^2 = .10$, $F(1, 121) = 13.90$, $\beta = 0.32$, $t = 3.73$, $p_{adj} < .001$), but not in context binding ($R^2 < .001$, $F(1, 139) = 0.03$, $\beta = .01$, $t = 0.16$, $p_{adj} > .999$) or in item perceptual specificity ($R^2 = .02$, $F(1, 139) = 2.31$, $\beta = -.13$, $t = -1.52$, $p_{adj} = .390$, for full information see Tables S15–S17). This suggests that while older children showed higher relative overnight gains for generalized memories, they did not outgain their younger counterparts on memory specificity.

## No evidence for a trade-off between the ability to gain generalized knowledge and the ability to retain specificities of memories across an overnight delay

Finally, we explored the possibility that gains for generalization and memory retention for specificity trade-off with one another after an overnight delay. Due to the low number of trials per category that could both be gained for generalization and retained for specific memory aspects, this analysis could not be performed on a trial level. Instead, it was performed on a between-participant level and thus asks the question of whether children who showed a greater relative gain for generalization are also the ones with lower relative retention for the idiosyncrasies of individual episodes. We computed an exploratory multiple linear regression predicting gain in generalization with retention in context binding and in item perceptual specificity as predictors and age as covariates. Age was the only significant predictor ($\beta = .28$, $t = 2.86$, $p = .005$; see also section "Age effects on delayed memory gain differ between generalization and specificity"), whereas retention in context binding and in item perceptual specificity was not ($\beta = .03$, $t = 0.33$, $p = .746$ and $\beta = .10$, $t = 1.11$, $p = .268$, respectively; for full information see Table S18). We did not detect any evidence of a trade-off between the ability to gain generalized knowledge and the ability to retain specificities of individuated memories.

# DISCUSSION

The current study investigated age-related differences in two key memory functions—generalization and memory specificity—in a cross-sectional sample of children aged 4–8 years. Four main findings can be noted. First, both generalization and memory specificity showed robust age-related differences over this time period, with particularly strong effects in generalization. Second, generalization was tethered to children's memory for the conceptual identity of items and to the semantic similarity across objects from related events. Third, successful generalization was more closely associated with the explicit knowledge of the regularity in character-category co-occurrences in older than in younger children, and fourth, across a sleep-filled delay, older children retained both generalized and specific memories relatively more than younger children. Interestingly, we saw a stronger association between age and retention for generalization than memory specificity. However, age was only associated with delayed gains in generalization but not in any indicators of memory specificity. An exploratory analysis found no evidence for a trade-off between children's ability to retain memory specifics and the relative gain in generalizations.

## Development of generalization and memory specificity

The transition from early to middle childhood is characterized by fundamental changes in the ability to form episodic memories (Canada et al., 2021) and extract generalized knowledge (Ngo et al., 2021). Corroborating such previous findings, we found that multifaceted aspects of memory specificity improve with age from early to middle childhood. Spanning the period of preschool age to middle childhood, children's memories of object–context associations were strongly related to their age, consistent with prior research on similar age windows (e.g., Canada et al., 2020; Riggins, 2014). However, this result did not replicate that of Ngo et al. (2021), who surprisingly did not find an age effect on context binding in the very same task. The relatively larger sample size in the current study may have given us greater power to detect the age effects on context binding. Memory specificity for individual objects on both the conceptual and perceptual details also improved across this age window, in line with findings in Ngo et al. (2021). Compared to younger children, older children were better able to distinguish learned items from semantically related lures (Ngo et al., 2021). Memory discrimination for perceptual attributes of objects also positively scaled with age, consistent with previous studies on pattern separation development (e.g., Canada et al., 2019; Keresztes et al., 2017; Ngo, Newcombe, & Olson, 2019; Rollins & Cloude, 2018). Note that the majority of studies on pattern separation development employed variants of the Mnemonic Similarity Task (Kirwan & Stark, 2007; Stark et al., 2013). This task measures participants' memory discrimination for perceptually similar object exemplars, and thus, discrimination responses are mixed between conceptual and perceptual dimensions of learned items. Here, our findings demonstrate that, when assessed separately, younger children's memory specificity for both kinds of attributes is subpar to that in older children.

Complementing the abilities to remember the specificity of individual episodes, older children surpassed younger children in generalizing to novel situations. These findings again replicate a previous study that used the same task (Ngo et al., 2021) and others that used different paradigms to measure statistical learning (e.g., Schlichting et al., 2017), generalization (e.g., Pudhiyidath et al., 2020), or memory integration (e.g., Bauer & San Souci, 2010; Schlichting et al., 2017). Expanding previous studies, our study also examined the role of memory for the regularities in character-category co-occurrences in generalization success across age. Interestingly, we found that compared to younger children, older children showed a stronger association between memory for repeated co-occurrence and its respective generalization success. Put differently, if an older child remembers what tends to co-occur, s/he very likely would be able to accurately generalize based on newly acquired knowledge. Although the same was true for younger children, the coupling between the two expressions of generalizable knowledge was weaker. A previous study found that adults showed a stronger association between knowledge of statistical regularities and generalization than children and adolescents, who did not differ from each other (Pudhiyidath et al., 2020). With the current sample, which included younger children than the previous study (ages four to eight vs ages seven to nine), we detected age-related differences also across childhood in how the knowledge that is acquired on the category structure of related events is deployed to guide behavior in novel situations.

Adding to previous findings, our results further reveal that generalization and memory specificity differentially depend on participants' age. Generalization abilities were more strongly associated with participants' age than any of our three indicators of memory specificity. It has been noted before that the developmental profiles for generalization and memory specificity in early development are separable (e.g., Keresztes et al., 2018; Newcombe et al., 2007; Ramsaran et al., 2019). Indeed, the differential age effects on generalization versus specificity reported here resonate with the notion of differential development of these two complementary memory functions. However, given the cross-sectional design of the current work, we cannot draw conclusions about within-person changes in any of the constructs under investigation (Lindenberger et al., 2011; Voelkle et al., 2014). Rather, the current observations form an empirical basis for formulating hypotheses about patterns of correlated developmental changes within individuals.

The uneven age associations of the memory functions investigated in the present study have been hypothesized to be driven, at least in parts, by differences in maturational pace of the underlying neural substrates (Ghetti & Bunge, 2012; Keresztes et al., 2017). While generalization, as captured, for example, through associative inference abilities or integration of factual knowledge, has been linked to gray matter volume in the medial prefrontal cortex and in the hippocampal head (mPFC, Bauer et al., 2019; Schlichting et al., 2017), the improvement of memory specificity capacities has been thought to be related to maturational processes in specific hippocampal subfields, among those, for example, cornu ammonis (CA) structures 2–4 and the dentate gyrus, which was associated with measures of pattern separation (Canada et al., 2019), and CA1, previously linked to source memory (Riggins et al., 2018)—all structures that show a protracted development throughout childhood (e.g., Keresztes et al., 2017). However, understanding the precise relation between age-related changes in generalization and memory specificity—such as those that we observed in the current study—and within-person maturational processes of the underlying neural substrates remains an open question to be addressed in future research.

It is further worth noting that memory functions have been suggested to differ between mono- and bilingual children (Brito et al., 2014; Brito & Barr, 2012; Kormi-Nouri et al., 2003) and to be mediated by a variety of sociocultural factors (Botdorf et al., 2022; Fyffe et al., 2011; Nyberg et al., 1996). The current study did not assess any information on language development or socioeconomic status. Future research could thus shed light on whether memory functions, their overnight fate, as well as their development across childhood are differentially impacted by these factors.

## Generalization is contingent on conceptual item memory and semantic similarity

Theoretical models posit that rapid generalization is supported by retrieval processes of related but separate episodes (Kumaran & McClelland, 2012). Empirical findings in adults support this idea thus far (Banino et al., 2016; Ngo et al., 2021; Tompary et al., 2020; but see Ciranka et al., 2022). However, the asynchronous development of generalization and memory specificity suggests that children may differ from adults in their reliance on memory specificity when deriving generalizations (Ngo et al., 2021). Existing data pertinent to this question are scarce. Here, we replicated the initial finding that while children's generalization is linked to their memories of the conceptual identities of elements within a given episode, this does not apply to remembering the contextual circumstances or the perceptual details of that episode (Ngo et al., 2021). That is, generalization success was tethered to children's memory for the conceptual identity of the items, as well as the semantic proximity between them. However, we did not see a similar link to context memories or memories for the perceptual details of the objects. This corroborates the hypothesis that children might rely more on their memories for conceptual identities and the overall semantic structures that tie episodes together, rather than the reconstruction of full-fledged what–where–when memories on the fly when making inferences about novel situations (Ngo et al., 2021).

## Differential age effects on overnight retention and gain of generalization and memory specificity

Notably, in the current study, the retention and gain of generalized and specific memories after an overnight delay were differentially associated with participants' age. We found that relative retention was higher in older children than in younger children for generalized and specific memories. Importantly, the association between retention and age was more pronounced in generalization compared with memory specificity. Further, the relative degree to which children gained generalized knowledge positively scaled with age. The same did not hold true for memory for context or perceptual details of objects. These differential age effects on the time-dependent fate of generalized and specific memories from 4 to 8 years of age complement our findings on rapid generalization and memory specificity and hint toward distinct developmental trajectories of the consolidation of different kinds of memoranda. It is likely that the consolidation of generalization is specifically optimized within the studied age range, thereby promoting a prioritization of generalization over specificity in early development (Keresztes et al., 2018). Hence, by specifying how the time-dependent fate of generalized and specific memories may differ across development, our findings are consistent with the hypothesis that the development of different memory functions proceeds in a lead-lag manner during early and middle childhood. This hypothesis awaits to be tested directly by longitudinal studies.

While the current study design does not allow inferences about conditions underlying the differential age effects on retentions and gains of generalized and specific memories, our results provide an important foundation for future research that targets the consolidation mechanisms underlying these effects more specifically. In light of the wealth of studies suggesting the importance of sleep for memory across development (e.g., Backhaus et al., 2008; Kurdziel et al., 2013; Spencer & Riggins, 2022; Wilhelm et al., 2012), examining this link more stringently for the consolidation of generalization and memory specificity across development would be a fruitful line of research.

It is worth noting that the relative measures for retention and gain implemented in this study are dependent on evening performance in Session 1. This applies specifically to gains, with higher levels of immediate memories leaving less room for items to gain. Therefore, a given proportion of gain does not necessarily reflect the same absolute quantity of information gained. While it is important to keep this characteristic of the gain and retention measures in mind when interpreting the findings, the largest differences between relative versus absolute numbers of gained items come from high performers. However, the robustness analysis, which excluded trials close to ceiling performance (see Table S14), led to comparable results as our main analysis, thus supporting the assumption that the effects reported in this paper were not driven by a bias in the relative gain or retention.

## No evidence for a trade-off between the ability to gain generalized knowledge and the ability to retain specificities

In an exploratory analysis, we further tested the possibility that gaining generalized memories after a night of sleep may come at the cost of retaining specific memories, based on ideas from traditional consolidation theories (Diekelmann & Born, 2010; McClelland et al., 1995). Such theories suggest that the increasing expression of generalized knowledge requires specificities of their constituent memories to be less expressed (Richards & Frankland, 2017). However, in our results, we did not find evidence for a trade-off between the ability to gain generalized memories and retain specific memories on an interindividual level. While research on a trade-off between generalized and specific aspects of a memory is scarce until present, one comparable analysis in young adults has reported a link between better memory precision and increased generalization (Tompary et al., 2020). Together, these and the present findings therefore seem to contradict the idea of a strict trade-off between generalization and specificity. Do note, however, that the results presented here were on a participant level, not within trials, and thus do not reflect the balance between the two aspects of the same memory (but see Tompary & Davachi, 2017). Future studies should aim to understand how the retention of specific memories benefits or hinders the emergence of generalized knowledge that is extracted from such episodes.

# CONCLUSION

In conclusion, our study corroborates previous evidence that early and middle childhood is a pivotal period of memory development that is characterized by pronounced age-related improvements in rapid generalization, including its retention, and by the emergence of novel inferences after a sleep-filled delay. These findings lay the groundwork for future longitudinal studies investigating interindividual differences and commonalities of intraindividual changes in generalization and specificity, and their relation to maturational brain changes.

## ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

## FUNDING INFORMATION

This work was conducted within the project "Lifespan Rhythms of Memory and Cognition (RHYME)" at the Center for Lifespan Psychology of the Max Planck Institute for Human Development. A-KJ and ESB were fellows of the International Max Planck Research School on the Life Course (LIFE; http://www.imprs-life.mpg.de/en). MW-B received support from the Jacobs Foundation (Early Career Research Fellowship 2017–2019). CTN's work is supported by a grant from the German Research Foundation (DFG; NG 191/2-1) and an Early Career Research Fellowship by the Jacobs Foundation (2021-1417-99). We thank G. Faust for organizing data collection and all students of the RHYME project, specifically T. C. Meyer, M. K. Schmidt, C. Stein, W. Voigt, and I. Kerber, for their help in collecting the data. We are grateful to all members of the RHYME and LIME projects for valuable feedback. We further acknowledge support from the Max Planck Dahlem Campus of Cognition (MPDCC). Finally, we thank all our participants and their families for their time. The analyses presented here were not preregistered.

## CONFLICT OF INTEREST STATEMENT

None.

## DATA AVAILABILITY STATEMENT

All data and code to reproduce the present analyses and result figures are publicly available through the Open Science Framework (https://osf.io/afcxn/).

## ETHICS STATEMENT

The current study was approved by the ethics committee of the Max Planck Institute of Human Development (LIP 2019-11-COMIC and LIP 2020-12 COMIC).

## ORCID

*Elisa S. Buchberger* https://orcid.org/0000-0001-5377-9722
*Ann-Kathrin Joechner* https://orcid.org/0000-0003-4962-1089
*Chi T. Ngo* https://orcid.org/0000-0001-6962-7168
*Ulman Lindenberger* https://orcid.org/0000-0001-8428-6453
*Markus Werkle-Bergner* https://orcid.org/0000-0002-6399-9996

## REFERENCES

Backhaus, J., Hoeckesfeld, R., Born, J., Hohagen, F., & Junghanns, K. (2008). Immediate as well as delayed post learning sleep but not wakefulness enhances declarative memory consolidation in children. *Neurobiology of Learning and Memory*, *89*, 76–80. https://doi.org/10.1016/j.nlm.2007.08.010

Banino, A., Koster, R., Hassabis, D., & Kumaran, D. (2016). Retrieval-based model accounts for striking profile of episodic memory and generalization. *Scientific Reports*, *6*, 31330. https://doi.org/10.1038/srep31330

Bates, D., Maechler, M., & 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

Bauer, P. J., & Dow, G. A. (1994). Episodic memory in 16- and 20-month-old children: Specifics are generalized but not forgotten. *Developmental Psychology*, *30*, 403–417. https://doi.org/10.1037/0012-1649.30.3.403

Bauer, P. J., Dugan, J. A., Varga, N. L., & Riggins, T. (2019). Relations between neural structures and children's self-derivation of new knowledge through memory integration. *Developmental Cognitive Neuroscience*, *36*, 100611. https://doi.org/10.1016/j.dcn.2018.12.009

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, *107*, 452–465. https://doi.org/10.1016/j.jecp.2010.05.012

Bland, J. M., & Altman, D. G. (1995). Multiple significance tests: The bonferroni method. *BMJ*, *31*, 170. https://doi.org/10.1136/bmj.310.6973.170

Botdorf, M., Dunstan, J., Sorcher, L., Dougherty, L. R., & Riggins, T. (2022). Socioeconomic disadvantage and episodic memory ability in the ABCD sample: Contributions of hippocampal subregion and subfield volumes. *Developmental Cognitive Neuroscience*, *57*, 101138. https://doi.org/10.1016/j.dcn.2022.101138

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/10.1163/156856897X00357

Brito, N., & Barr, R. (2012). Influence of bilingualism on memory generalization during infancy. *Developmental Science*, *15*, 812–816. https://doi.org/10.1111/j.1467-7687.2012.1184.x

Brito, N. H., Grenell, A., & Barr, R. (2014). Specificity of the bilingual advantage for memory: Examining cued recall, generalization, and working memory in monolingual, bilingual, and trilingual toddlers. *Frontiers in Psychology*, *5*, Article 1369. https://doi.org/10.3389/fpsyg.2014.01369

Canada, K. L., Geng, F., & Riggins, T. (2020). Age- and performance-related differences in source memory retrieval during early childhood: Insights from event-related potentials. *Developmental Psychobiology*, *62*, 723–736. https://doi.org/10.1002/dev.21946

Canada, K. L., Hancock, G. R., & Riggins, T. (2021). Developmental changes in episodic memory across early- to mid-childhood: Insights from a latent longitudinal approach. *Memory*, *30*, 1–14. https://doi.org/10.1080/09658211.2021.2006233

Canada, K. L., Ngo, C. T., Newcombe, N. S., Geng, F., & Riggins, T. (2019). It's all in the details: Relations between young Children's developing pattern separation abilities and hippocampal subfield volumes. *Cerebral Cortex*, *29*, 3427–3433. https://doi.org/10.1093/cercor/bhy211

Chatburn, A., Lushington, K., & Kohler, M. J. (2021). Consolidation and generalisation across sleep depend on individual EEG factors and sleep spindle density. *Neurobiology of Learning and Memory*, *179*, 107384. https://doi.org/10.1016/j.nlm.2021.107384

Ciranka, S., Linde-Domingo, J., Padezhki, I., Wicharz, C., Wu, C. M., & Spitzer, B. (2022). Asymmetric reinforcement learning facilitates human inference of transitive relations. *Nature Human Behaviour*, *6*, 555–564. https://doi.org/10.1038/s41562-021-01263-w

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*, 693–700. https://doi.org/10.1016/j.conb.2006.10.012

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, *11*(2), 114–126. https://doi.org/10.1038/nrn2762

Dumay, N. (2016). Sleep not just protects memories against forgetting, it also makes them more accessible. *Cortex*, *74*, 289–296. https://doi.org/10.1016/j.cortex.2015.06.007

Durrant, S. J., Taylor, C., Cairney, S., & Lewis, P. A. (2011). Sleep-dependent consolidation of statistical learning. *Neuropsychologia*, *49*, 1322–1331. https://doi.org/10.1016/j.neuropsychologia.2011.02.015

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 7723–7728. https://doi.org/10.1073/pnas.0700094104

Fenn, K. M., & Hambrick, D. Z. (2013). What drives sleep-dependent memory consolidation: Greater gain or less loss? *Psychonomic Bulletin & Review*, *20*, 501–506. https://doi.org/10.3758/s13423-012-0366-z

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, *6*, 119–130. https://doi.org/10.1038/nrn1607

Friedrich, M., Mölle, M., Friederici, A. D., & Born, J. (2020). Sleep-dependent memory consolidation in infants protects new episodic memories from existing semantic memories. *Nature Communications*, *11*, 1298. https://doi.org/10.1038/s41467-020-14850-8

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, *6*, 6004. https://doi.org/10.1038/ncomms7004

Fyffe, D. C., Mukherjee, S., Barnes, L. L., Manly, J. J., Bennett, D. A., & Crane, P. K. (2011). Explaining differences in episodic memory performance among older African Americans and whites: The roles of factors related to cognitive reserve and test bias. *Journal of the International Neuropsychological Society*, *17*, 625–638. https://doi.org/10.1017/S1355617711000476

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*, 381–395. https://doi.org/10.1016/j.dcn.2012.05.002

Habib, R., & Nyberg, L. (2008). Neural correlates of availability and accessibility in memory. *Cerebral Cortex*, *18*, 1720–1726. https://doi.org/10.1093/cercor/bhm201

Hanert, A., Weber, F. D., Pedersen, A., Born, J., & Bartsch, T. (2017). Sleep in humans stabilizes pattern separation performance. *Journal of Neuroscience*, *37*, 12238–12246. https://doi.org/10.1523/JNEUROSCI.1189-17.2017

Hayne, H., MacDonald, S., & Barr, R. (1997). Developmental changes in the specificity of memory over the second year of life. *Infant Behavior and Development*, *20*, 233–245. https://doi.org/10.1016/S0163-6383(97)90025-4

Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience & Biobehavioral Reviews*, *37*, 36–58. https://doi.org/10.1016/j.neubiorev.2012.09.014

Joechner, A., Wehmeier, S., & Werkle-Bergner, M. (2021). Electrophysiological indicators of sleep-associated memory consolidation in 5- to 6-year-old children. *Psychophysiology*, *58*, e13829. https://doi.org/10.1111/psyp.13829

Keresztes, A., Bender, A. R., Bodammer, N. C., Lindenberger, U., Shing, Y. L., & Werkle-Bergner, M. (2017). Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 9212–9217. https://doi.org/10.1073/pnas.1710654114

Keresztes, A., Ngo, C. T., Lindenberger, U., Werkle-Bergner, M., & Newcombe, N. S. (2018). Hippocampal maturation drives memory from generalization to specificity. *Trends in Cognitive Sciences*, *22*, 676–686. https://doi.org/10.1016/j.tics.2018.05.004

Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory*, *14*, 625–633. https://doi.org/10.1101/lm.663507

Kormi-Nouri, R., Moniri, S., & Nilsson, L.-G. (2003). Episodic and semantic memory in bilingual and monolingual children. *Scandinavian Journal of Psychology*, *44*, 47–54.

Koster, R., Chadwick, M. J., Chen, Y., Berron, D., Banino, A., Düzel, E., Hassabis, D., & Kumaran, D. (2018). Big-loop recurrence within the hippocampal system supports integration of information across episodes. *Neuron*, *99*, 1342–1354.e6. https://doi.org/10.1016/j.neuron.2018.08.009

Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: A model of the hippocampal system. *Psychological Review*, *119*, 573–616. https://doi.org/10.1037/a0028681

Kurdziel, L., Duclos, K., & Spencer, R. M. C. (2013). Sleep spindles in midday naps enhance learning in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 17267–17272. https://doi.org/10.1073/pnas.1306418110

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. *Sleep Medicine Reviews*, *18*, 531–541. https://doi.org/10.1016/j.smrv.2014.03.005

Lau, H., Alger, S. E., & Fishbein, W. (2011). Relational memory: A daytime nap facilitates the abstraction of general concepts. *PLoS One*, *6*, e27139. https://doi.org/10.1371/journal.pone.0027139
Lewis, P. A., & Durrant, S. J. (2011). Overlapping memory replay during sleep builds cognitive schemata. *Trends in Cognitive Sciences*, *15*, 343–351. https://doi.org/10.1016/j.tics.2011.06.004
Lindenberger, U., von Oertzen, T., Ghisletta, P., & Hertzog, C. (2011). Cross-sectional age variance extraction: What's change got to do with it? *Psychology and Aging*, *26*, 34–47. https://doi.org/10.1037/a0020525
Lindsay, D. S., Johnson, M. K., & Kwon, P. (1991). Developmental changes in memory source monitoring. *Journal of Experimental Child Psychology*, *52*, 297–318. https://doi.org/10.1016/0022-0965(91)90065-Z
Lloyd, M. E., Doydum, A. O., & Newcombe, N. S. (2009). Memory binding in early childhood: Evidence for a retrieval deficit. *Child Development*, *80*, 1321–1328. https://doi.org/10.1111/j.1467-8624.2009.01353.x
Lukowski, A. F., Wiebe, S. A., & Bauer, P. J. (2009). Going beyond the specifics: Generalization of single actions, but not temporal order, at 9 months. *Infant Behavior and Development*, *32*, 331–335. https://doi.org/10.1016/j.infbeh.2009.02.004
Marr, D. (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *262*, 23–81. https://doi.org/10.1098/rstb.1971.0078
McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457. https://doi.org/10.1037/0033-295X.102.3.419
Muehlroth, B. E., Sander, M. C., Fandakova, Y., Grandy, T. H., Rasch, B., Shing, Y. L., & Werkle-Bergner, M. (2020). Memory quality modulates the effect of aging on memory consolidation during sleep: Reduced maintenance but intact gain. *NeuroImage*, *209*, 116490. https://doi.org/10.1016/j.neuroimage.2019.116490
Natick. (2016). *MATLAB. version 2016b*. The MathWorks Inc.
Newcombe, N. S., Balcomb, F., Ferrara, K., Hansen, M., & Koski, J. (2014). Two rooms, two representations? Episodic-like memory in toddlers and preschoolers. *Developmental Science*, *17*, 743–756. https://doi.org/10.1111/desc.12162
Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development of episodic and autobiographical memory: A cognitive neuroscience perspective. In R. V. Kail (Ed.), (Hrsg.) *Advances in child development and behavior* (Bd. 35, S (pp. 37–85). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-009735-7.50007-4
Ngo, C. T., Benear, S. L., Popal, H., Olson, I. R., & Newcombe, N. S. (2021). Contingency of semantic generalization on episodic specificity varies across development. *Current Biology*, *31*, 2690–2697.e5.
Ngo, C. T., Lin, Y., Newcombe, N. S., & Olson, I. R. (2019). Building up and wearing down episodic memory: Mnemonic discrimination and relational binding. *Journal of Experimental Psychology: General*, *148*, 1463–1479. https://doi.org/10.1037/xge0000583
Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2018). The ontogeny of relational memory and pattern separation. *Developmental Science*, *21*. https://doi.org/10.1111/desc.12556
Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2019). Gain-loss framing enhances mnemonic discrimination in preschoolers. *Child Development*, *90*, 1569–1578. https://doi.org/10.1111/cdev.13297
Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, *110*, 611–646. https://doi.org/10.1037/0033-295X.110.4.611
Nyberg, L., Backman, L., Erngrund, K., Olofsson, U., & Nilsson, L.-G. (1996). Age differences in episodic memory, semantic memory, and priming: Relationships to demographic, intellectual, and biological factors. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *51B*, P234–P240. https://doi.org/10.1093/geronb/51B.4.P234
Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. https://aclanthology.org/volumes/D14-1/
Piaget, J., Grize, J.-B., Szeminska, A., & Bang, V. (1977). *Epistemology and psychology of functions (Bd. 23)*. Springer Science & Business Media.
Pillemer, D. B., & White, S. H. (1989). Childhood events recalled by children and adults. In H. W. Reese (Ed.), *Advances in child development and behavior* (Bd. 21, S (pp. 297–340). JAI. https://doi.org/10.1016/S0065-2407(08)60291-8
Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2021). *Nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–152*: https://CRAN.R-project.org/package=nlme
Price, H. L., & Connolly, D. A. (2006). BatMon II: Children's category norms for 33 categories. *Behavior Research Methods*, *38*, 529–531. https://doi.org/10.3758/BF03192808
Pudhiyidath, A., Roome, H. E., Coughlin, C., Nguyen, K. V., & Preston, A. R. (2020). Developmental differences in temporal schema acquisition impact reasoning decisions. *Cognitive Neuropsychology*, *37*, 25–45. https://doi.org/10.1080/02643294.2019.1667316
R Core Team. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing* [Software]. https://www.R-project.org/
Ramsaran, A. I., Schlichting, M. L., & Frankland, P. W. (2019). The ontogeny of memory persistence and specificity. *Developmental Cognitive Neuroscience*, *36*, 100591. https://doi.org/10.1016/j.dcn.2018.09.002
Ranganath, C. (2010). Binding items and contexts: The cognitive neuroscience of episodic memory. *Current Directions in Psychological Science*, *19*, 131–137. https://doi.org/10.1177/0963721410368805
Richards, B. A., & Frankland, P. W. (2017). The persistence and transience of memory. *Neuron*, *94*(6), 1071–1084. https://doi.org/10.1016/j.neuron.2017.04.037
Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature Neuroscience*, *17*, 981–986. https://doi.org/10.1038/nn.3736
Riggins, T. (2014). Longitudinal investigation of source memory reveals different developmental trajectories for item memory and binding. *Developmental Psychology*, *50*, 449–459. https://doi.org/10.1037/a0033622
Riggins, T., Geng, F., Botdorf, M., Canada, K., Cox, L., & Hancock, G. R. (2018). Protracted hippocampal development is associated with age-related improvements in memory during early childhood. *NeuroImage*, *174*, 127–137. https://doi.org/10.1016/j.neuroimage.2018.03.009
Rollins, L., & Cloude, E. B. (2018). Development of mnemonic discrimination during childhood. *Learning & Memory*, *25*, 294–297. https://doi.org/10.1101/lm.047142.117
Rovee-Collier, C. (1997). Dissociations in infant memory: Rethinking the development of implicit and explicit memory. *Psychological Review*, *104*, 467–498. https://doi.org/10.1037/0033-295X.104.3.467
RStudio Team. (2020). *RStudio: Integrated development environment for R*. RStudio, PBC.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. https://doi.org/10.1126/science.274.5294.1926
Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, *29*, 37–51. https://doi.org/10.1162/jocn_a_01028

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, *60*, 378–389. https://doi.org/10.1016/j.neuron.2008.09.023

Sluzenski, J., Newcombe, N. S., & Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 89–100. https://doi.org/10.1037/0278-7393.32.1.89

Spencer, R. M. C., & Riggins, T. (2022). Contributions of memory and brain development to the bioregulation of naps and nap transitions in early childhood. *Proceedings of the National Academy of Sciences*, *119*, e2123415119. https://doi.org/10.1073/pnas.2123415119

Stark, S. M., Yassa Lacy, J. W., & Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia*, *51*, 2442–2449. https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Tompary, A., & Davachi, L. (2017). Consolidation promotes the emergence of representational overlap in the hippocampus and medial prefrontal cortex. *Neuron*, *96*, 228–241.e5. https://doi.org/10.1016/j.neuron.2017.09.005

Tompary, A., Zhou, W., & Davachi, L. (2020). Schematic memories develop quickly, but are not expressed unless necessary. *Scientific Reports*, *10*(1), 16968. https://doi.org/10.1038/s41598-020-73952-x

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U. (2014). Toward a unified framework for the study of between-person and within-person structures: Building a bridge between two research paradigms. *Multivariate Behavioral Research*, *49*, 193–213. https://doi.org/10.1080/00273171.2014.889593

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation—What can be learnt from children? *Neuroscience & Biobehavioral Reviews*, *36*, 1718–1728. https://doi.org/10.1016/j.neubiorev.2012.03.002

Witkowski, S., Noh, S., Lee, V., Grimaldi, D., Preston, A. R., & Paller, K. A. (2021). Does memory reactivation during sleep support generalization at the cost of memory specifics? *Neurobiology of Learning and Memory*, *182*, 107442. https://doi.org/10.1016/j.nlm.2021.107442

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*, 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Zeithamova, D., & Preston, A. R. (2010). Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *Journal of Neuroscience*, *30*, 14676–14684. https://doi.org/10.1523/JNEUROSCI.3250-10.2010

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: Building memories to navigate future decisions. *Frontiers in Human Neuroscience*, *6*, 1–14. https://doi.org/10.3389/fnhum.2012.00070

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Buchberger, E. S., Joechner, A.-K., Ngo, C. T., Lindenberger, U., & Werkle-Bergner, M. (2024). Age differences in generalization, memory specificity, and their overnight fate in childhood. *Child Development*, *95*, e270–e286. https://doi.org/10.1111/cdev.14089