


## RESEARCH ARTICLE

# Extended trajectory of spatial memory errors in typical and atypical development: The role of binding and precision

Maomiao Peng<sup>1</sup> | Annalysa Lovos<sup>1</sup>  | Kenneth Bottrill<sup>1</sup> | Katharine Hughes<sup>1</sup> |  
Miranda Sampsel<sup>1</sup> | Nancy Raitano Lee<sup>2</sup> | Leonard Abbeduto<sup>3,4</sup> |  
Angela John Thurman<sup>3,4</sup> | Jamie Edgin<sup>1</sup>

<sup>1</sup>Department of Psychology, The University of Arizona, Tucson, Arizona, USA

<sup>2</sup>Department of Psychological and Brain Sciences, Drexel University, Philadelphia, Pennsylvania, USA

<sup>3</sup>MIND Institute, University of California Davis Health, Sacramento, California, USA

<sup>4</sup>Department of Psychiatry and Behavioral Sciences, University of California Davis Health, Sacramento, California, USA

**Correspondence**

Jamie Edgin, Department of Psychology, The University of Arizona, Tucson, Arizona, USA.  
Email: [jedgin@arizona.edu](mailto:jedgin@arizona.edu)

**Funding information**

National Institutes of Health, Grant/Award Numbers: P50HD103526, R01HD088409

**Abstract**

Spatial reconstruction, a method for evaluating how individuals remember the placement of objects, has traditionally been evaluated through the aggregate estimation of placement errors. However, this approach may obscure the nature of task errors. Specifically, recent data has suggested the importance of examining the precision of responses, as well as absolute performance on item-context bindings. In contrast to traditional analysis approaches based on the distance between the target and the reconstructed item, in this study we further explored three types of errors (swap error, global error, and local distance) that may all contribute to the distance, with particular emphasis on swap errors and local distance due to their associations with item-context bindings and memory precision, respectively. We examined these errors in children aged 3–18 years, making comparisons between children with typical development (TD) and children with Down syndrome (DS), a population with known memory challenges. As expected, older children outperformed younger children in terms of overall memory accuracy. Of importance is that we measured uneven maturational trajectories of memory abilities across the various error types. Specifically, both remembered locations (irrespective of object identity) and swap errors (object-location binding errors) align with the overall memory accuracy. Memory precision, as measured by local distance in simpler set size 2 trials, mirrored overall memory accuracy. However, for more complex set size 3 trials, local distance remained stable before age 8 and showed age-related change thereafter. The group with DS showed reduced precision compared to a TD matched group, and measures of precision, and to a lesser extent binding errors, correlated with standard neuropsychological outcomes. Overall, our study contributed to a fine-grained understanding of developing spatial memory ability in a large sample of typical developing children and a memory impaired population. These findings contribute to a growing body of research examining precision as a key factor in memory performance.

**KEYWORDS**

down syndrome, hippocampus, memory development, precision, relational memory

## 1 | INTRODUCTION

For a human or an animal to navigate the physical world or predict an appropriate response in a social scenario, the brain needs access to bound representations of both objects and their corresponding spatial and temporal contexts. Much research supports the view that the hippocampus is involved in binding spatial and temporal relations (Cohen & Eichenbaum, 1993; Ekstrom & Yonelinas, 2020; Wang et al., 2015), and patients with hippocampal memory impairments have particular difficulty remembering arbitrary bindings (Konkel et al., 2008). Recent approaches have emphasized that the hippocampus may achieve this unique binding function by facilitating representational precision and the linking of cross-regional representations (Ekstrom & Yonelinas, 2020). When an imprecise binding happens, what factors are at play? A careful analysis of patterns of spatial reconstruction can help us to identify the mechanisms underlying binding deficits, including how typical children may develop binding ability and how atypically developing children (studies here in Down syndrome) may develop these skills differently.

For certain spatial reconstruction tasks, participants remember the locations of multiple items in the background context. After a delay, they place the items back into their original locations (Horecka et al., 2018; Watson et al., 2013). The most common way to evaluate reconstruction errors is by calculating the distance between the original and the reconstructed locations. However, this analysis method masks the reasons reconstructive errors may occur. For example, the participant may make consistent errors because of an altered view on a screen or consistently forget each object's identity and swap the locations of two or more objects. Traditional analyses ignore the types of errors made, which may conceal important between-group differences in populations with memory difficulties or advantages. Therefore, it is important to examine the nature of reconstructive errors in ways that allow for the division of error into distinct error types. One type of error, that is, a "swap error," involves the participant incorrectly putting one item into another item's location. This type of error is more commonly measured in working memory studies (Mathy & Varré, 2013), and recent studies have investigated its role in episodic memory (Blankenship & Kibbe, 2019). Another error involves a systematic spatial misplacement, perhaps due to the positioning of the screen in front of the participant, such as moving all the items toward one direction (e.g., to the left), biasing their placement toward the center, stretching them toward the perimeter, or rotating them to some degree, all named "global error." Yet another possibility is that they remember the approximate location but with poor precision, named "local distance" errors here (Horecka et al., 2018), measured by calculating the distance between the original and the reconstructed locations after adjusting the swap error and global error. Thus, this measurement differs from traditional methods that measure the distance directly between the original and reconstructed locations, as it takes into account the object-location binding, global misplacement, and precision, separately.

The neural basis for each of these errors may be distinct. Swap error is mainly related to the binding of objects with locations, which

is at the heart of hippocampal function according to relational binding theory (Cohen & Eichenbaum, 1993; Eichenbaum & Cohen, 2004; Horecka et al., 2018). Several possible explanations are possible for the occurrence of global errors. They may be related to the participant's use of a "unitized" representation via Gestalt perceptual features, or the processing of geometric information in the parahippocampal place area, or the shape information provided by the inferior temporal cortex (Corsi, 1972; Epstein, 2014; Horecka et al., 2018; Kessels et al., 2000; Perrett & Oram, 1993; Uttal & Chiong, 2004). Errors in local distance relate to the memory precision supported by the hippocampus and dorsolateral prefrontal cortex in conjunction (Stevenson et al., 2018). As the precision and binding model suggests (Yonelinas, 2013), the hippocampus supports high-resolution binding, and hippocampal damage leads to impaired performance on memory tasks that require high-resolution binding, but fewer impacts when tasks only require low-resolution binding. While all three types of error may be linked to medial temporal lobe function to some degree, we believe that object-location binding and precision are most closely related to the function of the hippocampus. Therefore, although our study considered global errors, the main focus was on object-location binding and precision.

The hippocampus undergoes protracted development after birth (Brown et al., 2012; Hu et al., 2013; Lee et al., 2014, 2017; Tamnes et al., 2013; Uematsu et al., 2012). However, the developmental trajectories of the hippocampal subregions and subfields, connected structures, and the larger hippocampal network show regional variation (Blankenship et al., 2017; Poldrack, 2010). The hippocampal tail undergoes development between 4 and 8 years (Canada et al., 2020) and remains stable from 9.5 to 12 years (Lee et al., 2020). Furthermore, Lee et al. (2020) found that item-space memory improves until 10.5 years and positively correlates with an increase in right hippocampal tail volume. Based on these findings, we hypothesized a consistent reduction in the swap error, which is associated with object-location binding, across our three typical development groups (mean age: 4, 6, and 12 years) and deficits in Down syndrome, which has had reports of decreased volume of the hippocampal tail (Koenig et al., 2021).

Ekstrom and Yonelinas (2020) suggest that precision is distributed across multiple brain regions, with the hippocampus playing a critical integrating role. Stevenson et al. (2018) found a correlation between increased gamma power in hippocampal CA1 and spatial memory retrieval precision, indicating its involvement in high-fidelity spatial memory representations. Recent work on subfield analysis of the hippocampus in Down syndrome has shown reductions in CA1, DG, and the hippocampal tail (Koenig et al., 2021), suggesting that this group may show difficulties across both processes. However, the specific developmental trajectory of precision remains unknown, suggesting a need for additional examination across development and in disorders of the hippocampus. Consequently, the present study aims to investigate the development of the ability to bind spatial relations and we assessed children's performance on swap errors, global errors, and local distance, a metric of precision. Many previous studies have found age-related development of relational binding, and high resolution allocentric search memory (Hassevoort et al., 2020; Lavenex & Lavenex, 2021; Lee et al., 2016; Ngo et al., 2019; Ribordy Lambert et al., 2015;

Yim et al., 2013), but few of these studies have attempted to simultaneously uncouple the processes involved in binding and precision in young children.

Down syndrome (DS) is the most common genetic form of intellectual disability, with most cases caused by an extra copy of chromosome 21 (Steingass et al., 2011), with a frequency of 1 in every 691 live births (Parker et al., 2010). One of the key symptoms in individuals with DS is developmental delays in learning and memory, especially those memory processes dependent on the hippocampus (Conners et al., 2011; Edgin, 2013; Godfrey & Lee, 2018; Nadel, 2003; Pennington et al., 2003). As mentioned previously, relational binding ability is one of the key functions of the hippocampus and could be measured with fine granularity with the spatial reconstruction paradigm, however we still know little about it in individuals with DS. Thus, in this study, we aimed to elucidate the relative strengths and weaknesses of specific subdomains that are measured by this paradigm in individuals with DS, as one way of examining how altered hippocampal function may manifest in different errors on this task.

Because there are limited data available for developmental and impaired populations using a paradigm that allows researchers to examine subtypes of errors, we turn to the literature on individuals with hippocampal damage to generate hypotheses about error types that may be found among individuals with DS. Two recent studies provided data from patients with hippocampal lesions that showed deficits in spatial-temporal binding. Watson et al. (2013) reported that three participants with substantial bilateral hippocampal damage showed errors caused by swapping the locations of the items, and that this significantly contributed to the overall performance difference between these patients and healthy controls. Horecka et al. (2018) also found that identity-related information (item-location binding) contributed to the primary differences between the patients and the healthy controls, whereas other types of relation bindings were not impaired. In other recent studies, binding precision in working memory assessments was assessed in autism and patients with specific developmental lesions to the hippocampus, finding no difficulties with precision on continuous outcomes, while the retrieval success of arbitrary bindings was impaired (Allen et al., 2022; Cooper et al., 2017). Therefore, more data are required regarding how binding and precision measures may vary across developmental stages and in atypically developing groups.

In summary, the goal of the current study is to examine the development of relational memory binding in TD children and children with DS using a fine-grained distinction of the different types of errors displayed during a spatial reconstruction task. We do so by analyzing these data with adjustment for systematic response biases that may obscure these metrics in these developing populations. Based on previous studies in hippocampal patients and developing groups, our hypotheses are that younger children may show more swap errors than older children, and children with DS may show similar deficit patterns to the hippocampal patients, specifically errors in binding (e.g., swap errors). Precision could also be impacted in this group given the extent of their hippocampal dysfunction. We present data on global hippocampal volumes in a small subset of participants to replicate past work showing volume reductions in this population as a basis for informing these analyses. We also

investigate whether those error types (precision, binding errors) are correlated with memory tasks that are often used to assess components of episodic and memory development, that is, spatial memory, pattern separation, memory updating and short-term memory. Through this analysis, we aim to shed light on shared neural mechanisms between these errors and various memory domains, given their distinct developmental trajectories. This analysis is important to highlight the extent that this new view of precision may relate to more traditionally constructed measures of memory that are often used as assessment tools.

## 2 | METHODS

### 2.1 | Participants

The participants in this study were 79 TD children (mean age = 6.76 years,  $SD = 3.55$ ; range = 3.25–17.47 years, 32 female), and 48 children with DS (mean age = 12.1 years,  $SD = 3.57$ ; range = 6.14–17.95 years, 24 female) recruited across three sites: the University of Arizona, Drexel University, and University of California, Davis. A subset of  $n = 5$  participants with DS received MRI structural imaging studies (for MRI acquisition and process parameters, see Appendix A), confirming a reduction in hippocampal volume after correction for total brain volume compared to an age-matched sample with TD (Left hippocampi:  $p = .002$ ; Right hippocampi:  $p = .034$ ), as has been reported in previous studies of DS (Carducci et al., 2013; Pinter et al., 2001; Smigielska-Kuzia et al., 2011). All children with DS were confirmed to have a genetic diagnosis of DS before entry into our study. All TD children were free from neurological or psychiatric disorders according to parent report. Informed consent was obtained from all parents prior to the experiment. This study was approved by the Institutional Review Boards of the participating universities.

To understand the developmental trends in TD children, we divided participants into three sub-groups according to age: younger (3.3–5.5 years,  $n = 39$ ), middle (5.5–7.5 years,  $n = 20$ ) and older (7.5–17.5 years,  $n = 20$ ) groups. The detailed demographic characteristics are presented in Table 1. To compare the TD and DS group, we first aimed to find the TD children matchable to the DS group via comparing the verbal and non-verbal raw scores of Kaufman Brief Intelligence Test (KBIT-2). We used the data from all the TD children whose verbal age equivalence was less than or equal to 7.5 years as a TD Matched group (Table 2), because that allowed for direct matching of the verbal and non-verbal raw scores with the DS group (non-verbal raw score: mean(TD) = 15.35, std(TD) = 5.07, mean(DS) = 12.64, std(DS) = 5.05,  $U = 1218$ ,  $p = .0889$ ; verbal raw score: mean(TD) = 28.13, std(TD) = 8.00, mean(DS) = 23.95, std(DS) = 12.49,  $U = 1219$ ,  $p = .0883$ ), and involved some older TDs in the comparison, thus enabling us to capture the potential differences that might exist in older youth. We did not divide the DS group into age groups due to the wide age range and small number of participants within each age category. Additionally, we did not have a priori hypotheses about specific age ranges that may be important for developmental change in the DS group.

## 2.2 | Alien object task

The alien object task (Figure 1a) is one task included in an NIH funded battery designed and validated for developmental and clinical memory

**TABLE 1** TD age groups.

Group	Age (years)	Gender		<i>n</i>
	Mean ± SD (range)	Female	Male	
Younger	4.45 ± 0.7 (3.3–5.5)	16	23	39
Middle	6.19 ± 0.58 (5.5–7.5)	10	10	20
Older	11.85 ± 3.43 (7.5–17.5)	6	14	20

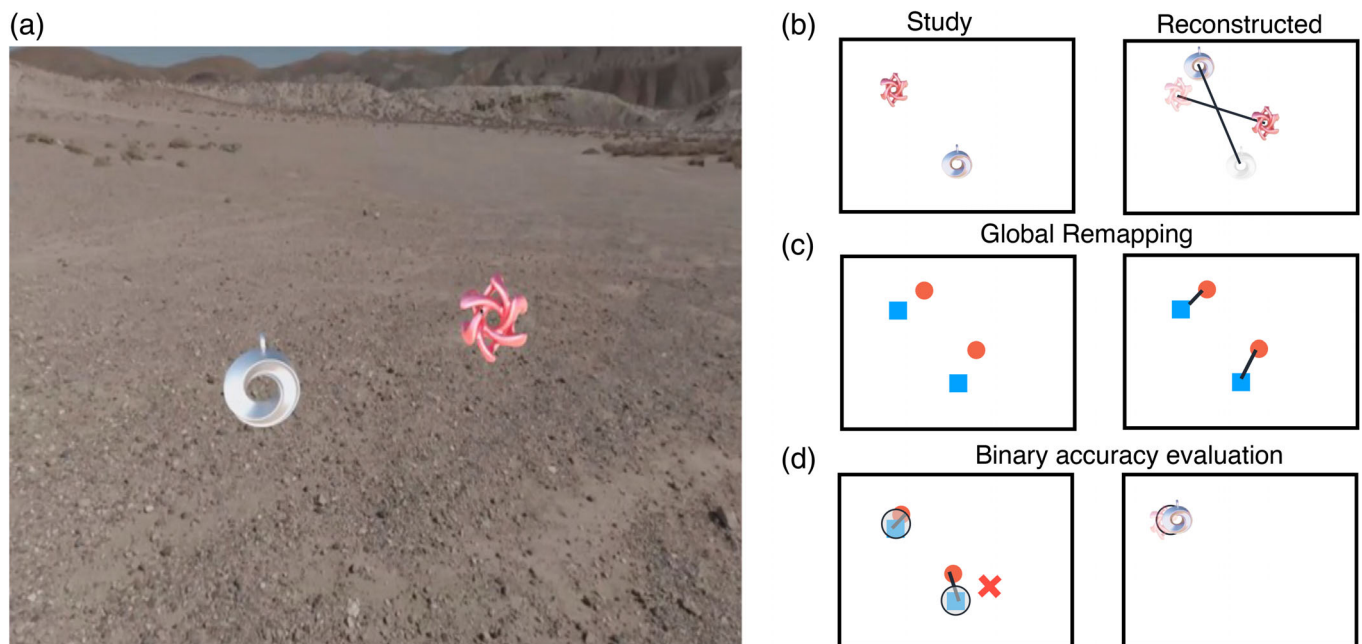
**TABLE 2** Mean raw scores (and standard deviation) on variables for the TD matched group and DS (*n* = 6 was dropped in DS because the missing of KBIT-2 data).

Group	TD matched	DS
Age	5.0 (1.1)	12.7 (3.4)
<i>n</i>	48	42
Original distance	22.5 (9.2)	27.6 (12.4)
Remembered locations	9.9 (3.6)	8.7 (3.9)
Swaps	3.8 (2.3)	3.9 (2.1)
Local distance	10.5 (2.8)	11.8 (2.3)

investigations aimed at studying young children and children with disability (the Arizona Memory Assessment for Preschoolers and Special Populations, or A-MAP; Edgin PI). This task was adapted from one commonly used spatial reconstruction task (Watson et al., 2013). The task was developed using Unity, which involved viewing the position of objects for 5 s in a naturalistic scene and then re-placing those objects in the scene on an iPad display. Specifically, the participants were first instructed to remember the location of two or three objects, which are parts of an alien spaceship on a desert scene picture. After a 5-s delay, participants were then instructed to put the objects back in their original location. The task storyline involved the reconstruction of an alien's spaceship after a crash. There were six trials in total. Half of the trials included two objects and half included three objects. The precise position is saved in Unity units, that is, virtual meters. At the beginning of the task, participants had a practice trial with two objects not used in the formal experiment.

## 2.3 | Measurements of errors in the A-MAP alien object task

Our analysis pipeline generated four main misplacement variables adapted from a recent study (Horecka et al., 2018) which are meant to allow for separate analysis of the ability to place objects in space based on location as well as the binding of the location and object identity—while



**FIGURE 1** Alien object task and the analysis procedure. (a) During the learning phase, participants were shown two or three objects and they were instructed to put the objects back in their original location after a five-second delay. (b) The studied locations (left) and the reconstructed locations (right). In the right panel, the studied locations are denoted by lighter colors, while the reconstructed locations are represented by solid colors. The black lines represent the distance between the studied locations and the reconstructed locations. (c) Stripping away the identity information (left) and re-pair the objects' original locations and reconstructed locations (right). The blue boxes represent the original locations, and the red circles represent the reconstructed locations. (d) Count the number of objects whose distance between the post-adjusted reconstructed locations and original locations is within the threshold (left; marked by a red 'x' if outside the threshold) and the objects that are both within the distance threshold and also with correct identity (right).

**TABLE 3** Main variables of the alien object task.

Variable name	Definition
Original distance	The mean of Euclidean distance between each object's studied location and reconstructed location.
Remembered location rate	The ratio of the number of item locations that fell within the threshold, regardless of identity (i.e., remembered locations), to the total number of locations. For details about the threshold see Section 2.3.
Swap rate	The ratio of the number of objects within the threshold but with a wrong identity to the number of remembered locations. For details about the threshold see Section 2.3.
Local distance	The mean of Euclidean distance between each object's studied location and reconstructed location after adjusting swap errors (set size both two and three trials), or after adjusting swap errors and global error (set size three trials).

accounting for systematic error. These four key indices were (1) original distance, (2) remembered location rate, (3) swap rate and (4) local distance (Table 3). Each of these variables will be introduced in following parts.

In order to evaluate the participant's overall performance on the task, we first measured the *original distance* between the reconstructed objects and studied objects for set size 2 and 3 trials separately (Figure 1b), which is defined by the mean of the Euclidean distance between each object's studied location and reconstructed location.

In Step 2, to adjust for potential errors caused by misplacing an item to another item's location, we performed a global remapping analysis. Specifically, we stripped away the identity information of the reconstructed objects and studied objects. As Figure 1c shows, the locations of the reconstructed objects are indicated as the red circles, and the locations of the studied objects are marked with the blue boxes. We re-paired the studied objects and the reconstructed objects by choosing the one that could minimize the sum of all misplacement errors of all pairs. This analysis allowed for an estimation of the participant's ability to remember the spatial positions without referring to object identity. After remapping, we re-computed the mean distance between the location of the reconstructed objects and the studied objects for set size 2 and 3 trials separately, defined as *local distance*.

In Step 3, the binary accuracy evaluation, we used a 95% confidence interval of all subjects' local distance as the threshold for both the TD and DS group. The location of reconstructed objects, after global remapping, was considered accurate if it fell within the threshold and inaccurate if it did not. Consequently, the *remembered location rate* was computed by calculating the number of item locations that fell within the threshold and dividing it by the total sum of locations separately for the set size 2 and set size 3 trials (Figure 1d). We further calculated the number of items' location that are within the threshold but with wrong item identity in Step 2; that is, participants remembered the locations but exchanged them with the locations of other items, which is defined as the number of swaps. To alleviate potential biases stemming from variations in the number of remembered locations across participants and ensure comparability between set size 2 and 3 trials, we calculated the *swap rate*. This involved dividing the number of swaps by the corresponding number of remembered locations for each participant, separately for set size 2 and set size 3 trials. Previous investigations had

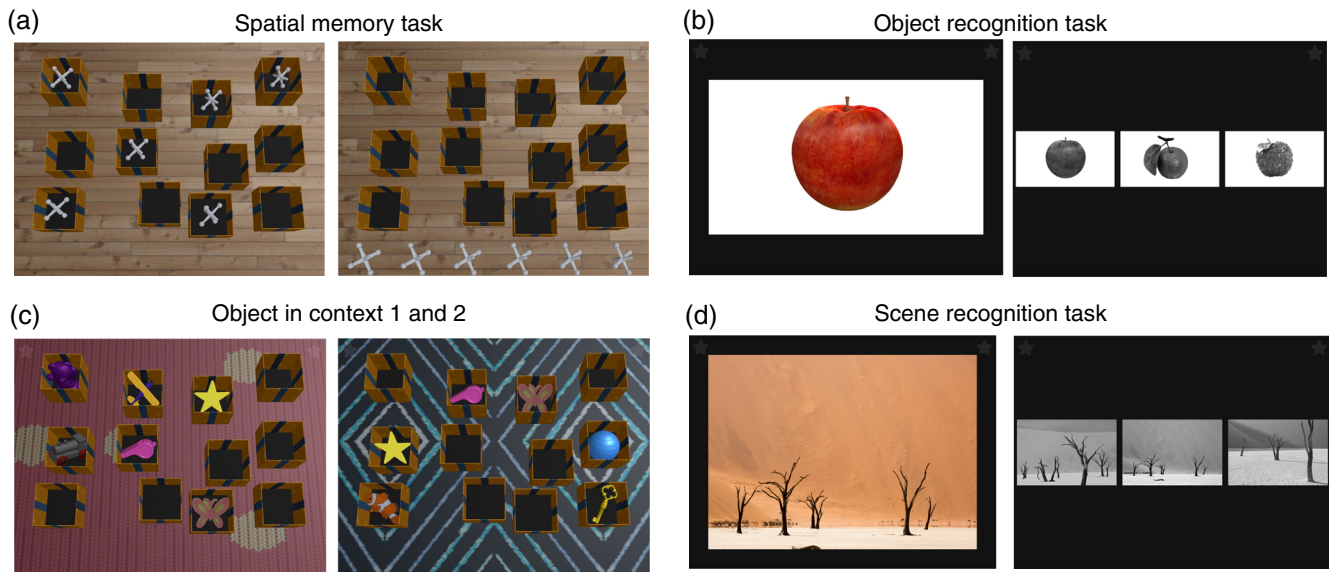
suggested that hippocampal patients were specifically impaired at swapping items (Horecka et al., 2018; Watson et al., 2013).

In addition, the Horecka et al. (2018) study used Umeyama's algorithm (Umeyama, 1991) to compute global errors shared by all objects that are caused by *rotation*, *translation* (moving all objects toward left or right, bottom or up), and *scaling* (squeezing all objects toward the center or stretching them toward the perimeter), after re-mapping the objects. We applied this same analysis (the global transformation step in Figure 5c), to compare the re-calculated remembered location rate, swap rate, and local distance after correction for global error between the TD age groups, and between the TD matched and DS groups, and also compare the variables (rotation angle, scaling, or translation magnitude) that depict the global correction process among the groups.

In the alien object task, there are two set sizes (two and three) and the global errors were only computed in the set size three condition. The reason is as follows: Umeyama's algorithm computes the similarity transformation parameters (rotation, translation, and scaling) that yields the least mean squared error between two sets of points (Umeyama, 1991). In our analysis, one set is the objects' studied locations, and the other is the reconstructed objects' locations after re-pairing identity in Step 2. In the set size two condition, the location of two objects could form a line segment; and in the set size three condition, the location of three objects could form a triangle. Since any two line segments are similar by definition, there is always a perfect similarity transformation between them, which would falsely take the local error (i.e., local distance) into shared global error. Thus, we only did this analysis in the set size three condition. After adjusting the global errors, we re-computed the distance between objects' studied locations and post-adjusted locations and obtained new local distances, remembered location rate, and swap rate. All steps of the above analysis were performed using Python 3.10.3.

## 2.4 | Neuropsychological tasks

To investigate whether there are shared neural mechanisms between these errors (original distance, remembered location rate, swap rate, and local distance) and the neural mechanisms indexed by neuropsychological tests, we measured the participants' performance on neuropsychological tasks from a novel battery (the A-MAP), including



**FIGURE 2** Neuropsychological tasks from the A-MAP (Arizona Memory Assessment for Preschoolers and Special Populations).

a digit span, spatial memory, object in context 1 and 2, object recognition, and scene recognition (Figure 2).

#### 2.4.1 | DAS-II digit span task

The digit span task from the Differential Ability Scales (DAS-II) Recall of Digits Forward task is frequently used for measuring working memory in research and clinical practice (Elliot, 1990). In this task, participants need to recall the digits sequentially as they were heard at a rate of 2 digits per second. The score is determined by the scoring method of the standardized assessment, which was correct blocks at each digit difficulty level, with score ranges from 0 to 38.

#### 2.4.2 | A-MAP spatial memory task

In this multi-trial task, the participants were shown the location of six of the same objects inside separate boxes, presented as birthday presents (12 in total). There were two practice trials that did not contribute to the score, and then either three or four trials where the participant had to remember the location of the six objects and correctly place them where they saw them. The total score for this phase ranges from 0 to 24, as it is possible to get six correct objects four times in a row. If the participant did not get at least a total of five objects correct in the first three trials, the phase ended before the fourth trial.

#### 2.4.3 | A-MAP object in context 1 and 2

This task has an identical set-up to the spatial memory task, but to measure the ability of memory updating, it included two phases in which

participants find objects in two overlapping configurations. In the object in context 1 task, six different objects were shown inside of 12 separate boxes on top of a rug background, and the participants were instructed to remember the location of the six objects and correctly place them where they saw them. In the object in context 2 task, the procedure was the same, but half of the objects are changed to new objects, and half of the objects are from the context 1 task but in different locations, thus, participants need to update their memory of these objects' location. Both the *Object in context 1 score* and *Object in context 2 score* ranged from 0 to 24 based on how many correct locations were selected in each context separately. As with the spatial memory task, this task terminated at the fourth trial if participants did not reach criterion.

#### 2.4.4 | A-MAP object recognition task

In a modified version of the Mnemonic similarity task, participants were shown 12 object pictures and then were instructed to select the image they previously saw among three options (target, 2 distractors). Half of the trials had choices quite similar looking to the target image, and the other half of the trials had choices with greater differences from the target image. We validated these categories with undergraduates and child pilot participants. The task includes a total of 12 trials in which the participant must select the image that they saw before, so the *Object recognition score* ranges from 0 to 12 based on how many correct images were selected.

#### 2.4.5 | A-MAP scene recognition task

In this task, participants were shown 12 scene pictures and then were instructed to select the image they previously saw among three options

(target, 2 distractors). As above, half of the trials had similar distractors, and the other half had more dissimilar choices, validated by ratings of similarity. There were a total of 12 trials in which the participant must select the image that they saw before, so the *Scene recognition score* ranged from 0 to 12 based on how many correct images were selected.

## 2.5 | Statistical analysis

To evaluate the influence of group and set size on the four misplacement measurements, we conducted two-way mixed ANOVAs on rank-transformed data. Rank transformation was employed due to deviations from normal distribution. Partial eta square ( $\eta^2$ ) values were computed as effect sizes for ANOVA effects: small ( $<0.06$ ), medium ( $0.06-0.14$ ), and large ( $\geq 0.14$ ). Post hoc analyses were conducted using the Bonferroni correction to compare the three subgroups within TD group.

Furthermore, we also utilized mixed Bayesian ANOVAs on rank-transformed data to explore potential null effects for non-significant findings identified in the two-way mixed ANOVAs. The analysis was performed using JASP software with default priors (JASP Team, 2023). In this analysis, the exclusion Bayes factor ( $BF_{\text{excl}}$ ), was computed as  $p(\text{models without that factor} | \text{data}) \div p(\text{models with that factor} | \text{data})$  (Keysers et al., 2020). For instance, a  $BF_{\text{excl}}$  value of 9 signifies that the models excluding the factor are nine times more likely than the models including it. This suggests that the factor does not exert a significant influence on the variance of the dependent variable. For the post hoc comparisons,  $BF_{01}$  was reported. Similarly, for instance, a  $BF_{01}$  value of 9 implies that it is nine times more likely for the two groups to exhibit no difference, compared to having a difference. In the analysis specifically involving set size 3 trials in Section 3.3, the same approach as described above was applied, replacing “set size” with “global transformation” (i.e., pre-global transformation or post-global transformation).

In an attempt to shed light on the potential common neural underpinnings between specific types of spatial memory errors calculated above and cognitive processes as determined by neuropsychological tests, and furthermore, to explore the possibility of spatial memory errors predicting cognitive task performance, we examined the relationship between swap rate, local distance and neuropsychological memory task scores. The neuropsychological memory task scores were fitted using the following three models:

- Model 1: neuropsychological memory task score  $\sim$  group + swap rate + group:swap rate
- Model 2: neuropsychological memory task score  $\sim$  group + local distance + group:local distance
- Model 3: neuropsychological memory task score  $\sim$  group + swap rate + local distance + group:swap rate + group:local distance

To determine which model best accounted for participants' performance, the three models were compared based on Akaike information criterion (AIC) and Bayesian information criterion (BIC). The variables were rescaled to fall between zero and one before entry into the models.

## 3 | RESULTS

### 3.1 | Performance of TD age groups on set size 2 and 3

#### 3.1.1 | Original distance

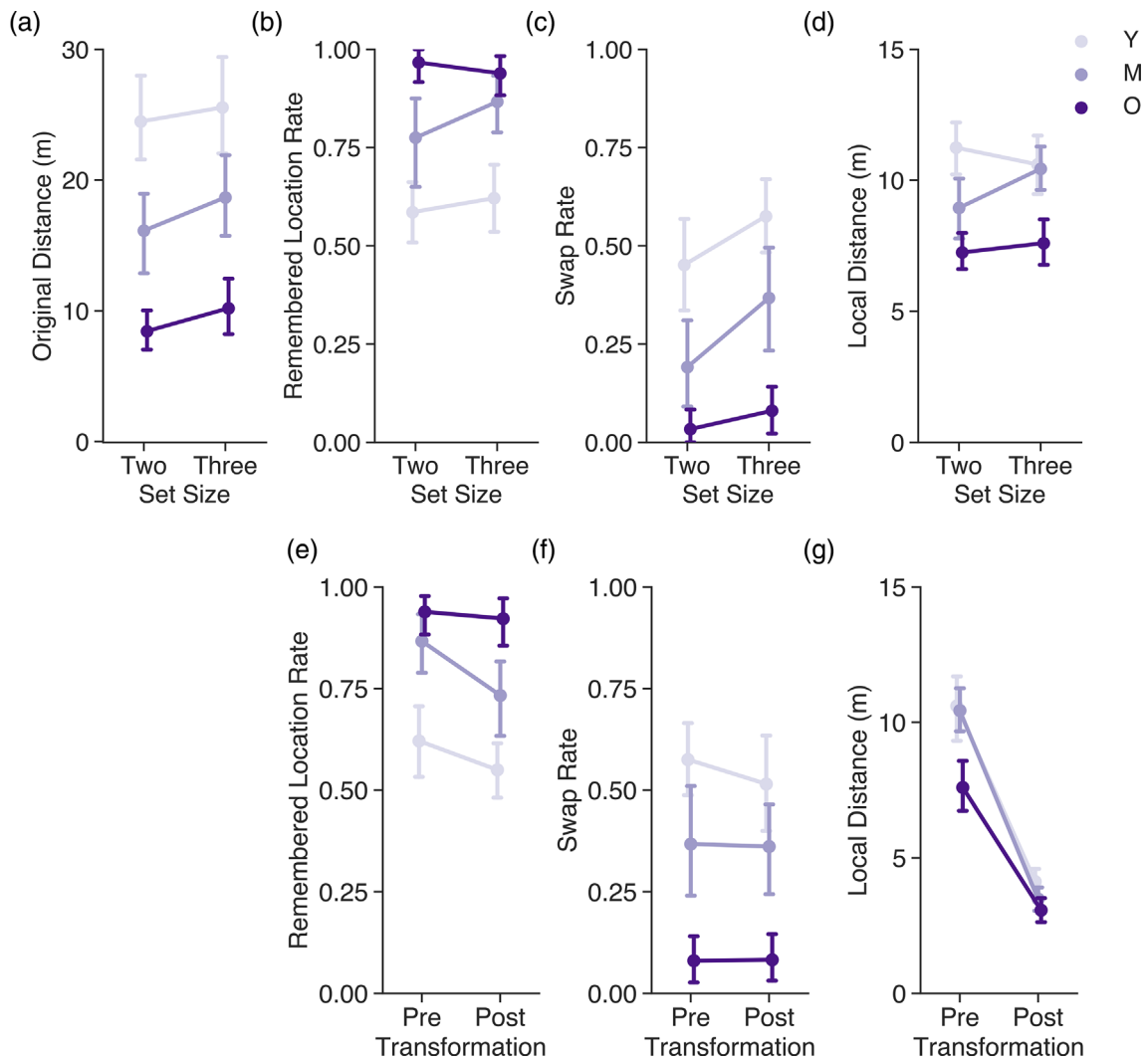
A 3 (age group)  $\times$  2 (set size) mixed ANOVA showed significant age group differences in original distances ( $F_{(2,76)} = 44.08$ ,  $p < .0001$ ,  $\eta^2 = 0.5371$ ; Figure 3a). For the set size effect, the data were inconclusive ( $F_{(1,76)} = 3.89$ ,  $p = .05$ ,  $\eta^2 = 0.0488$ ,  $BF_{\text{excl}} = 1.34$ ). Although the  $p$ -value approached the significance threshold for the alternative hypothesis, Bayesian analysis provided anecdotal evidence supporting a null effect. Furthermore, no significant effect of the interaction between age group and set size was found ( $F_{(2,76)} = 0.59$ ,  $p = .56$ ,  $\eta^2 = 0.0152$ ,  $BF_{\text{excl}} = 2.81$ ). Post hoc  $t$  tests on age groups used Bonferroni correction confirmed that older children showed reduced original distance than younger children (older  $<$  middle,  $t = -6.09$ ,  $p < .0001$ ; older  $<$  younger,  $t = -13.16$ ,  $p < .0001$ ), and middle children showed reduced original distance than younger children ( $t = -4.66$ ,  $p < .0001$ ).

#### 3.1.2 | Remembered location rate

A 3 (age group)  $\times$  2 (set size) mixed ANOVA showed significant age group differences in remembered location rate ( $F_{(2,76)} = 30.71$ ,  $p < .0001$ ,  $\eta^2 = 0.4470$ ; Figure 3b), but no significant effect of set size ( $F_{(1,76)} = 1.82$ ,  $p = .18$ ,  $\eta^2 = 0.0234$ ,  $BF_{\text{excl}} = 1.90$ ), and the effect of the interaction between age group and set size is inconclusive ( $F_{(2,76)} = 3.10$ ,  $p = .05$ ,  $\eta^2 = 0.0755$ ,  $BF_{\text{excl}} = 0.82$ ). Post hoc  $t$  tests used Bonferroni correction confirmed that older children remembered more locations than younger children (older  $>$  middle,  $t = 3.97$ ,  $p = .0005$ ; older  $>$  younger,  $t = 11.33$ ,  $p < .0001$ ), and middle children remembered more locations than younger children ( $t = 4.97$ ,  $p < .0001$ ).

#### 3.1.3 | Swap rate

A 3 (age group)  $\times$  2 (set size) mixed ANOVA showed significant age group differences ( $F_{(2,71)} = 40.84$ ,  $p < .0001$ ,  $\eta^2 = 0.5350$ ; Figure 3c) and set size differences ( $F_{(1,71)} = 12.09$ ,  $p = .0009$ ,  $\eta^2 = 0.1455$ ), but inconclusive evidence of the effect of the interaction between age group and set size ( $F_{(2,71)} = 0.84$ ,  $p = .4368$ ,  $\eta^2 = 0.0231$ ,  $BF_{\text{excl}} = 1.02$ ). Post hoc  $t$  tests with Bonferroni correction confirmed that older children showed fewer swap errors than younger ones (older  $<$  middle,  $t = -4.65$ ,  $p = .0001$ ; older  $<$  younger,  $t = -10.73$ ,  $p < .0001$ ), and middle children showed fewer swap errors than the younger children ( $t = -3.68$ ,  $p = .0013$ ). In addition, children showed more swap errors in set size 3 trials than in set size 2 trials ( $t = 3.49$ ,  $p = .0008$ ). Five participants, unable to place any items within the threshold for set size 2 (younger:  $n = 2$ ; middle:  $n = 1$ ) or set size



**FIGURE 3** Original distances (a), remembered location rate (b), swap rate (c), and local distance (d), remembered location rate after global transformation (e), swap rate after global transformation (f), and local distance after global transformation (g) by TD age groups. Panels (a)–(d) encompass data from set sizes 2 and 3, analyzed per the methodologies outlined in Figure 1, without global transformations. Panels (e)–(g) present set size 3 data after global transformation, obtained using the Umeyama algorithm (refer to Figure 5c), and set size 3 data prior to global transformation. Further details can be found in Section 2.3.

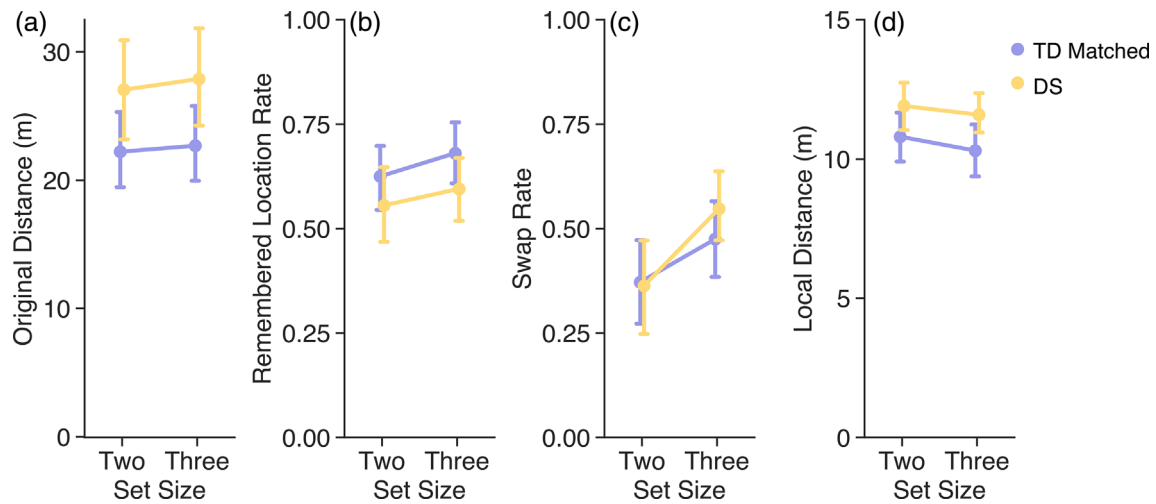
3 trials (younger:  $n = 2$ ), were omitted from the analysis due to incomputable swap rates and local errors.

### 3.1.4 | Local distance

A 3 (age group)  $\times$  2 (set size) mixed ANOVA showed significant age group differences in local distance ( $F_{(2,71)} = 24.31$ ,  $p < .0001$ ,  $\eta^2 = 0.4064$ ; Figure 3d), but no significant effect of set size ( $F_{(1,71)} = 0.62$ ,  $p = .4345$ ,  $\eta^2 = 0.0086$ ,  $BF_{\text{excl}} = 4.15$ ), and no interaction between age group and set size ( $F_{(2,71)} = 1.80$ ,  $p = .1729$ ,  $\eta^2 = 0.0482$ ,  $BF_{\text{excl}} = 2.34$ ). Post hoc  $t$  tests with Bonferroni correction confirmed that older children showed shorter local distance than younger groups (older < middle,  $t = -4.79$ ,  $p < .0001$ ; older < younger,  $t = -8.11$ ,  $p < .0001$ ). However, the

analysis yielded inconclusive evidence regarding the difference in local distance between middle children and younger children ( $t = -2.41$ ,  $p = .0544$ ,  $BF_{01} = 0.2903$ ). While the  $p$ -value favored the null hypothesis, the Bayesian analysis provided strong support for the alternative hypothesis (considering that the BF for the alternative hypothesis is the inverse of that for the null hypothesis, there is a 3.4 times higher chance of group differences over no difference). Given these findings, we conducted additional analyses to compare the local distance between the middle and younger children separately for set size 2 and 3 trials. These analyses revealed that the middle children showed reduced local distance in set size 2 trials compared with younger children ( $t = -2.88$ ,  $p = .0064$ ), but no difference between the middle children and younger children in set size 3 trials ( $t = -0.37$ ,  $p = .7117$ ,  $BF_{01} = 3.39$ ).





**FIGURE 4** Original distances (a), remembered location rate (b), swap rate (c), and local distance (d) for TD Matched and DS group. The data displayed in this figure are averages for set sizes 2 and 3, is analyzed according to the methodologies delineated in Figure 1, without global transformations.

### 3.2 | Comparing TD matched group with DS

For the original distance, a 2 (group)  $\times$  2 (set size) mixed ANOVA showed no effect of set size ( $F_{(1,88)} = 0.26, p = .6100, \eta^2 = 0.0030, BF_{\text{excl}} = 7.51$ ; Figure 4a), and no effect of the interaction between group and set size ( $F_{(1,88)} = 0.11, p = .7462, \eta^2 = 0.0012, BF_{\text{excl}} = 12.85$ ). The analysis provided inconclusive evidence regarding the difference in original distance between the TD matched group and DS group ( $F_{(1,88)} = 3.77, p = .0554, \eta^2 = 0.0411, BF_{\text{excl}} = 1.06$ ).

For the remembered location rate, a 2 (group)  $\times$  2 (set size) mixed ANOVA showed no effect of group ( $F_{(1,88)} = 1.96, p = .1654, \eta^2 = 0.0217, BF_{\text{excl}} = 2$ ; Figure 4b), and no effect of the interaction between group and set size ( $F_{(1,88)} = 0.11, p = .7388, \eta^2 = 0.0013, BF_{\text{excl}} = 5.5$ ), but inconclusive results of the set size effect ( $F_{(1,88)} = 3.96, p = .0498, \eta^2 = 0.0430, BF_{\text{excl}} = 1.39$ ).

For the swap rate, a 2 (group)  $\times$  2 (set size) mixed ANOVA showed that more swap errors occurred in set size 3 trials than in set size 2 trials across both groups ( $F_{(1,80)} = 10.91, p = .0014, \eta^2 = 0.1200$ ; Figure 4c), but no effect of group ( $F_{(1,80)} = 0.67, p = .4165, \eta^2 = 0.0083, BF_{\text{excl}} = 4.60$ ), and no effect of the interaction between group and set size ( $F_{(1,80)} = 0.22, p = .6412, \eta^2 = 0.0027, BF_{\text{excl}} = 4.73$ ). Eight participants, unable to place any items within the threshold for set size 2 (TD matched:  $n = 3$ ; DS:  $n = 2$ ) or set size 3 trials (TD matched:  $n = 2$ ; DS:  $n = 1$ ), were omitted from the analysis due to incomputable swap rates and local errors.

For the local error, a 2 (group)  $\times$  2 (set size) mixed ANOVA revealed that the DS group showed longer local distance than the TD matched group ( $F_{(1,80)} = 5.07, p = .0271, \eta^2 = 0.0596$ ), but no effect of set size ( $F_{(1,80)} = 1.54, p = .2187, \eta^2 = 0.0189, BF_{\text{excl}} = 3.74$ ), and no effect of the interaction between group and set size ( $F_{(1,80)} < 0.01, p = .9598, \eta^2 < 0.0001, BF_{\text{excl}} = 6.91$ ; Figure 4d).

### 3.3 | Set size 3 analysis post global transformation

We present an additional analysis of global errors in set size 3 trials only, because the adjustment for global errors is impossible to interpret for 2-object trials. In the set size 3 trials, we added a step between Step 2 and Step 3 to adjust global errors (rotation, scaling, and translation) shared by all objects, and re-calculated the remembered location rate, swap rate, and local distance, described in Section 2.3 and also shown in Figure 5.

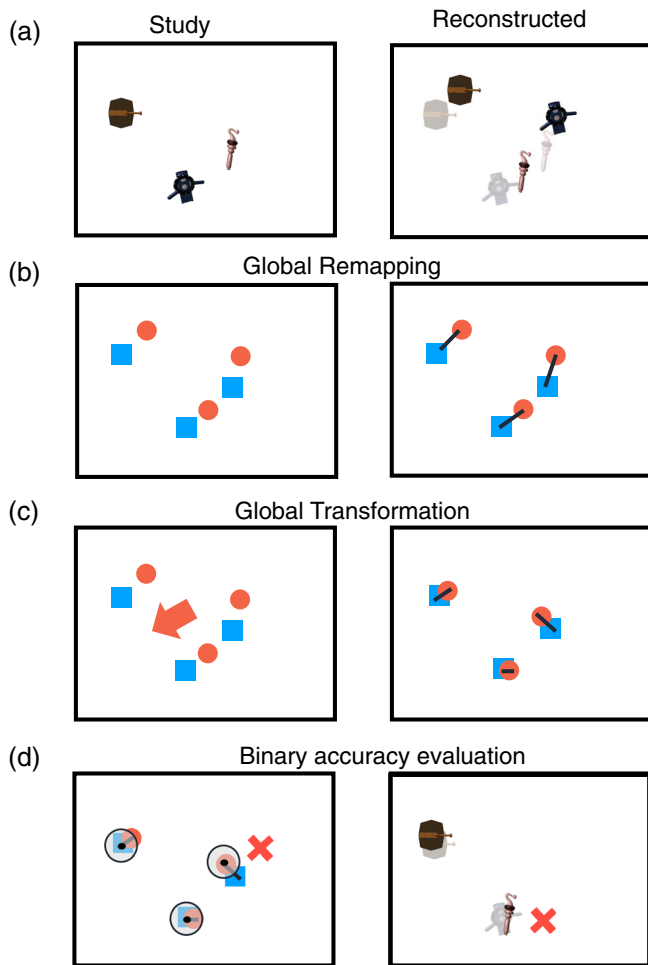
#### 3.3.1 | Performance of TD age groups

##### Global errors

A one-way ANOVA with rotation angle as dependent variable and age group as independent variable, showed that rotation angle differed between age groups ( $F_{(2,76)} = 11.93, p < .0001, \eta^2 = 0.2389$ ). Post hoc  $t$  tests with Bonferroni correction showed no significant differences between the older and middle groups ( $t = -1.63, p = .3315, BF_{01} = 1.14$ ), but the middle group showed smaller rotation angle than the younger age group ( $t = -2.89, p = .0179$ ). The older group also showed smaller rotation angle than the younger age group ( $t = -5.28, p < .0001$ ).

No significant differences were found between the age groups for the scaling ( $F_{(2,76)} = 3.05, p = .0532, \eta^2 = 0.0743, BF_{01} = 0.89$ ).

For the translation magnitude, we found a main effect of age ( $F_{(2,76)} = 10.98, p = .0001, \eta^2 = 0.2241$ ). Post hoc  $t$  tests with Bonferroni correction found that older children showed smaller translation magnitude than younger ones (older < middle,  $t = -3.65, p = .0024$ ; older < younger,  $t = -5.20, p < .0001$ ), but middle children showed no differences when compared to younger children ( $t = -1.44, p = .4671, BF_{01} = 1.54$ ).



**FIGURE 5** Set size 3 analysis with global transformation procedure. (a) The studied locations (left) and the reconstructed locations (right). In the right panel, the studied locations are denoted by lighter colors, while the reconstructed locations are represented by solid colors (b) Stripping away the identity information and re-pair the objects' original locations and reconstructed locations (right). The blue boxes represent the original locations, and the red circles represent the reconstructed locations. The black lines represent the distance between the studied locations and the reconstructed locations after re-pairing. (c) Adjusting the errors (rotation, scaling, translation) shared by all objects. The red circles were all moved toward the left and bottom directions. (d) Count the number of objects whose distance between the post-adjusted reconstructed locations and original locations is within the threshold (left; marked by a red 'x' if outside the threshold) and the objects that are both within the distance threshold and also with correct identity (right).

#### Remembered location rate

A 3 (age group)  $\times$  2 (global transformation) mixed ANOVA on the remembered location rate showed significant main effect of age group ( $F_{(2,76)} = 27.77, p < .0001, \eta^2 = 0.4222$ ; Figure 3e), global transformation ( $F_{(1,76)} = 162.34, p < .0001, \eta^2 = 0.6811$ ), and the interaction between age group and global transformation ( $F_{(2,76)} = 6.50, p = .0025, \eta^2 = 0.1461$ ). Subsequent Post hoc tests revealed that the older children demonstrated a significantly higher remembered location rate than the younger children, both before and after global

transformation (before global transformation:  $t = 6.68, p < .0001$ ; after global transformation:  $t = 8.17, p < .0001$ ). Similarly, the middle children also showed a significantly higher remembered location rate than the younger children, both before and after global transformation (before global transformation:  $t = 4.16, p = .0008$ ; after global transformation:  $t = 3.22, p = .0157$ ). Intriguingly, the previously non-significant difference in the remembered location rate before global transformation between older and middle children ( $t = 1.86, p = .42, BF_{01} = 0.84$ ) became significant after global transformation ( $t = 3.62, p = .0051$ ).

Furthermore, we calculated the difference scores of the remembered location rate before and after global transformation and explored their correlation with each global error. Surprisingly, the results indicated no significant correlation between the difference score and the global errors (rotation:  $r = -0.14, p = .2041, BF_{01} = 2.33$ ; scaling:  $r = -0.03, p = .7952, BF_{01} = 6.68$ ; translation:  $r = 0.13, p = .2572, BF_{01} = 3.46$ ).

The decrease in the remembered location rate observed from the pre to post global transformation, as depicted in Figure 5, can be attributed to the utilization of a more rigorous threshold in the analysis after global transformation. Specifically, the threshold was changed from 19.6 to 7.53. This alteration in the threshold is because the threshold for remembered locations is defined as a relative measurement based on the 95% confidence interval of all subjects' local distance, to avoid the ceiling effect. Consequently, the transition from an initially non-significant to a significant difference between older and middle children may be ascribed to the fact that older children exhibited shorter local distances, as delineated in Section 3.1.4. Hence, older children manifested a diminished susceptibility to the influence of the more stringent threshold, contributing to the observed findings.

#### Swap rate

A 3 (age group)  $\times$  2 (global transformation) mixed ANOVA on the swap rate showed significant main effect of age group ( $F_{(2,74)} = 22.37, p < .0001, \eta^2 = 0.3768$ ; Figure 3f), global transformation ( $F_{(1,74)} = 438.51, p < .0001, \eta^2 = 0.8556$ ), and the interaction between age group and global transformation ( $F_{(2,74)} = 25.37, p < .0001, \eta^2 = 0.4067$ ). Post hoc tests revealed inconclusive results on the difference between the middle children and the young children in the swap rate before global transformation ( $t = -2.41, p = .13, BF_{01} = 0.35$ ), and no significant difference after global transformation ( $t = -1.20, p = 1, BF_{01} = 2$ ). In addition, the older children demonstrated a significantly lower swap rate than the middle children both before and after global transformation (before global transformation:  $t = -3.94, p = .0020$ ; after global transformation:  $t = -4.01, p = .0017$ ). The older children also demonstrated a significantly lower swap rate than the younger children both before and after global transformation (before global transformation:  $t = -8.69, p < .0001$ ; after global transformation:  $t = -5.89, p < .0001$ ).

We calculated the difference scores of the swap rate before and after global transformation and explored their correlation with each global error. The results indicated no significant correlation between the difference score and the global errors (rotation:  $r = -0.06, p = .6,$

$BF_{01} = 5.87$ ; scaling:  $r = 0.1425, p = .2163, BF_{01} = 2.88$ ; translation:  $r = -0.17, p = .1428, BF_{01} = 2.09$ ).

*Local distance*

A 3 (age group)  $\times$  2 (global transformation) mixed ANOVA on the local distance showed significant main effect of age group ( $F_{(2,74)} = 13.04, p < .0001, \eta^2 = 0.2602$ ; Figure 3g), global transformation ( $F_{(1,74)} = 91.72, p < .0001, \eta^2 = 0.5535$ ), and the interaction between age group and global transformation ( $F_{(2,74)} = 6.41, p = .0027, \eta^2 = 0.1477$ ).

Post hoc tests revealed that the older children showed reduced local distance than the middle children before global transformation ( $t = -4.62, p = .0003$ ). However, after global transformation, this distinction became non-significant ( $t = -1.11, p = 1, BF_{01} = 2.00$ ). In addition, the older children demonstrated reduced local distance than the younger children both before and after global transformation (before global transformation:  $t = -4.84, p < .0001$ ; after global transformation:  $t = -3.32, p = .0112$ ). There was no significant difference in local distance between the middle children and the younger children before global transformation ( $t = -0.37, p = 1, BF_{01} = 3.40$ ), but inconclusive results of the difference between the two groups after global transformation ( $t = -1.97, p = .34, BF_{01} = 0.75$ ).

Additionally, we calculated the difference scores of the local distance before and after global transformation and explored their correlation with each global error. The results showed significant negative correlation between the difference score and the global errors

(rotation:  $r = -0.29, p = .0118$ ; scaling:  $r = -0.30, p = .0089$ ; translation:  $r = -0.56, p < .0001$ ).

Thus, the transition from an initially significant to a non-significant difference between older and middle children may be attributed to the fact the middle children showed greater translation error than the older children, as mentioned in Section 3.3.1.1. Consequently, the implementation of global transformation helped adjust the translation error, resulting in the disappearance of the observed difference.

3.3.2 | Comparing TD and DS groups

Three one-way ANOVA with each global error as dependent variable and group as independent variable separately was implemented to compare the global errors between TD matched and DS group.

No significant differences were found between the groups in any of the variables (rotation:  $F_{(1,88)} = 0.40, p = .5312, \eta^2 = 0.0045, BF_{excl} = 3.80$ ; scaling:  $F_{(1,88)} = 0.24, p = .6245, \eta^2 = 0.0027, BF_{excl} = 4.07$ ; translation:  $F_{(1,88)} = 1.96, p = .1655, \eta^2 = 0.0217, BF_{excl} = 1.92$ ).

The null effect we found for global errors (rotation, scaling, and translation) is consistent with our hypothesis that there is no difference between the TD and DS groups, as global errors are thought to be related to the functions of extra-hippocampal structures (Epstein, 2014; Horecka et al., 2018; Perrett & Oram, 1993). Moreover, since our primary focus is on object-location binding and precision, and no differences between the TD matched group and DS group were observed in these variables, we did not pursue further exploration of the variables re-calculated in Section 3.3 through regression analyses.

See Figure 6 for a schematic summary of the main results.

3.4 | Correlations with neuropsychological memory tasks

We investigated the relationship between neuropsychological memory tasks (digit span task score, spatial memory score, object in context 1 score, object in context 2 score, object recognition score, scene recognition score) and variables from the

Comparisons Between Groups

Comparing variables	Among TD age groups	TD Matched to DS
Original distance	Older < Middle < Younger	TD Matched $\approx$ DS
Remembered location rate	Older > Middle > Younger	TD Matched $\approx$ DS
Swap rate	Older < Middle < Younger	TD Matched $\approx$ DS
Local distance	Older < Middle In set size 2, Middle < Younger In set size 3, Middle $\approx$ Younger	TD Matched < DS

FIGURE 6 Schematic summary of the main results for both set sizes. Differences in local distance among TD age groups are presented separately for set sizes 2 and 3.

TABLE 4 Neuropsychological task performance for the TD group and DS.

Tasks	TD		DS		p
	n	Mean $\pm$ SD (range)	n	Mean $\pm$ SD (range)	
Digit span score	76	15.76 $\pm$ 6.34 (0-35)	38	8.45 $\pm$ 4.38 (0-20)	<.0001
Spatial memory score	79	17.94 $\pm$ 5.10 (6-24)	48	13.58 $\pm$ 5.65 (0-23)	<.0001
Object in context 1 score	78	14.82 $\pm$ 8.24 (0-24)	47	6.77 $\pm$ 7.38 (0-23)	<.0001
Object in context 2 score	78	12.06 $\pm$ 8.31 (0-24)	45	4.69 $\pm$ 4.76 (0-19)	<.0001
Object recognition score	79	9.89 $\pm$ 3.12 (0-12)	48	6.29 $\pm$ 3.20 (0-12)	<.0001
Scene recognition score	79	9.03 $\pm$ 2.90 (1-12)	48	6.33 $\pm$ 2.94 (1-12)	<.0001

**TABLE 5** (a) Model comparison; (b) digit span; (c) spatial memory, object in context 1, object in context 2, object recognition, and scene recognition.

(a)	Digit span			Spatial memory			Object in context 1			Object in context 2			Object recognition			Scene recognition		
	AIC	BIC	Estimate	AIC	BIC	p	AIC	BIC	p	AIC	BIC	p	AIC	BIC	p	AIC	BIC	p
Model 1	-111.2	-97.8		-38.4	-24.5		46.5	60.3		17.4	31.1		10.9	24.8		-2.0	11.9	
Model 2	-88.0	-74.3		-32.0	-17.7		63.2	77.4		40.0	54.1		20.8	35.0		17.1	31.4	
Model 3	-111.0	-92.2		-48.9	-29.4		33.3	52.7		8.5	27.7		3.8	23.2		-9.9	9.6	
<b>(b)</b>																		
Digit span																		
Estimate																		
Group	0.28									0.16, 0.39								
Swap	-0.12									-0.31, 0.07								
Group × swap	-0.22									-0.45, 0.01								
<b>(c)</b>																		
Group	Spatial memory			Object in context 1			Object in context 2			Object recognition			Scene recognition					
	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p			
Group	-0.18	-0.592, 0.233	.3896	0.25	-0.35, 0.85	.4050	0.61	0.06, 1.15	.0303	-0.32	-0.84, 0.19	.2174	-0.38	-0.87, 0.10	.1203			
Swaps	0.02	-0.224, 0.255	.8983	-0.29	-0.64, 0.05	.0968	-0.17	-0.51, 0.16	.3065	0.04	-0.26, 0.34	.7826	-0.33	-0.62, -0.05	.0205			
Local distance	-1.10	-1.707, -0.487	.0005	-1.11	-1.99, -0.23	.0138	-0.54	-1.34, 0.26	.1832	-1.28	-2.04, -0.52	.0012	-1.22	-1.94, -0.50	.0011			
Group × swaps	-0.40	-0.718, -0.079	.0150	-0.07	-0.53, 0.39	.7707	-0.33	-0.76, 0.10	.1301	-0.33	-0.73, 0.06	.0990	-0.07	-0.45, 0.31	.7158			
Group × local distance	0.78	0.023, 1.536	.0435	-0.05	-1.16, 1.07	.9335	-0.49	-1.50, 0.52	.3346	1.25	0.30, 2.19	.0101	1.01	0.12, 1.90	.0273			

alien object task, using three linear regression models (see Section 2.5 for details). Model selection is based on AIC and BIC (Table 5a).

For digit span, Model 1 (including group, swap rate, and the interaction between group and swap rate as predictors) demonstrated superior fit among the three models, as indicated by the lowest AIC and BIC values. For spatial memory task, objects in context 1, objects in context 2, object recognition, and scene recognition tasks, Model 3 (including group, swap rate, local distance, the interaction between group and swap rate, and the interaction between group and local distance as predictors), exhibited superior fit relative to the other two models. Subsequent analyses and interpretations are based on these selected models.

Furthermore, all neuropsychological tasks were evaluated for potential ceiling (accuracy  $\geq 0.95$ ) and floor effects (accuracy  $\leq 0.05$ ). The floor effect was most prominent in the object in context 1 and 2 tasks, affecting 8% and 15% of TD matched children and 40% and 21% of DS children, respectively. The ceiling effect was primarily found in the TD matched group across spatial memory (29%), the objects in context 1 (20%), object recognition (38%), and scene recognition (24%) tasks.

### 3.4.1 | Digit span

TD children outperformed DS children on the digit span task (Table 4,  $p < .0001$ ). Analysis of Model 1 showed that the group effect was found to be significant, while the other effects did not reach significance (Table 5b).

### 3.4.2 | Spatial memory

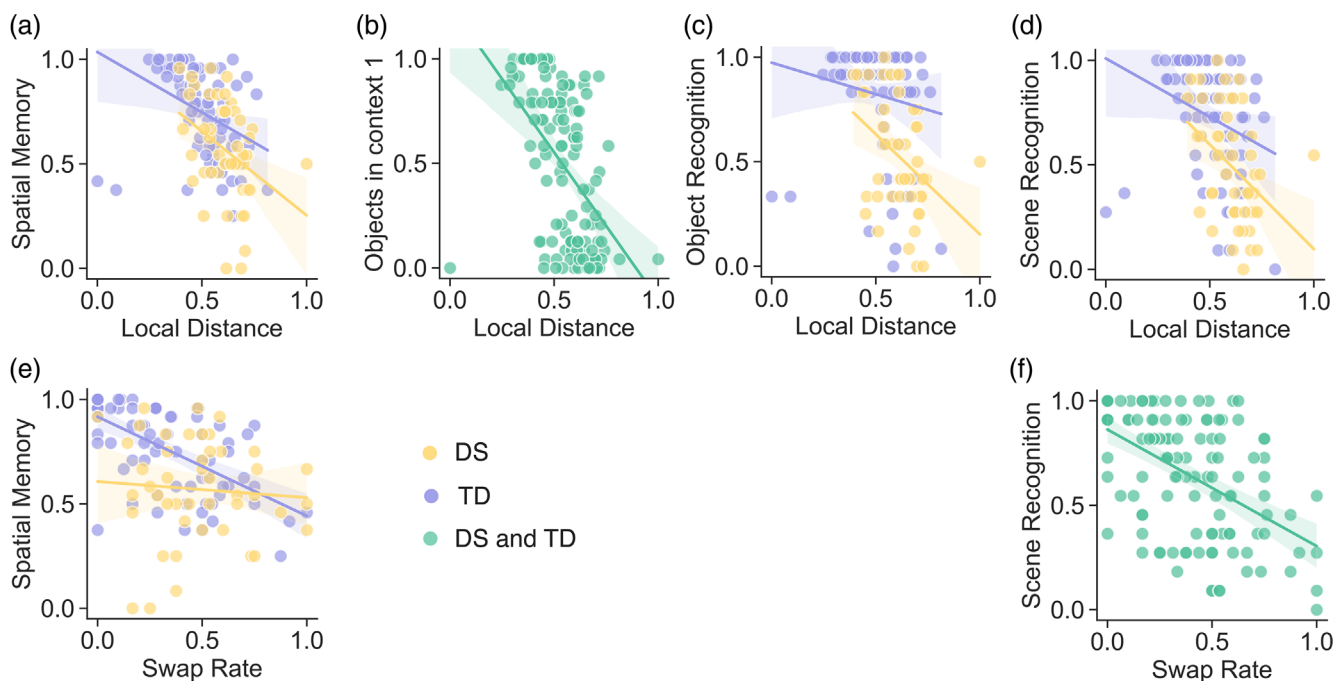
Analysis of Model 3 showed that lower local distance negatively correlated with better spatial memory performance. Examining the interaction terms, we found that in the TD group, a lower swap rate was associated with better spatial memory performance (simple slope:  $-0.38$ ,  $CI = [-0.59, -0.17]$ ), while no such relationship was found in the DS group (simple slope:  $0.02$ ,  $CI = [-0.22, 0.26]$ , Figure 7e). In the DS group, a lower local distance was associated with better spatial memory performance (simple slope:  $-1.10$ ,  $CI = [-1.71, -0.76]$ ), while no such relationship was found in the TD group (simple slope:  $-0.32$ ,  $CI = [-0.76, 0.19]$ , Table 5, Figure 7a).

### 3.4.3 | Objects in context 1

TD children exhibited significantly better performance on the objects in context 1 task compared to DS children (Table 4,  $p < .0001$ ). Results from Model 3 showed that lower local distance was associated with improved objects in context 1 task performance for both TD and DS groups (Table 5c, Figure 7b).

### 3.4.4 | Objects in context 2

TD children showed significantly better performance on the objects in context 2 task compared to DS children (Table 4,  $p < .0001$ ). Analysis of Model 3 also showed significant group effect, while the other effects did not reach significance (Table 5c).



**FIGURE 7** Swaps and local distance predict neuropsychological memory tasks. The data displayed in this figure are averaged across set sizes 2 and 3. Each dot represents one participant; the line shows a linear regression ( $\pm 95\%$  CI).

TABLE 6 Original distance.

	Digit span			Spatial memory			Object in context 1			Object in context 2			Object recognition			Scene recognition		
	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p	Estimate	95% CI	p
Group	0.399	0.272, 0.525	<.0001	0.301	0.136, 0.466	<.001	0.543	0.301, 0.785	<.0001	0.557	0.345, 0.77	<.0001	0.454	0.253, 0.656	<.0001	0.373	0.173, 0.574	<.001
Original distance	0.001	-0.234, 0.236	.992	-0.231	-0.52, 0.057	.115	-0.376	-0.794, 0.042	.078	-0.294	-0.661, 0.073	.116	-0.186	-0.538, 0.166	.297	-0.276	-0.626, 0.074	.121
Group × original distance	-0.638	-0.943, -0.333	<.0001	-0.510	-0.903, -0.117	.011	-0.854	-1.429, -0.279	.004	-0.960	-1.464, -0.456	<.001	-0.606	-1.085, -0.127	.014	-0.559	-1.036, -0.082	.022

### 3.4.5 | Object recognition

TD children exhibited significantly better performance on the object recognition task compared to DS children (Table 4,  $p < .0001$ ). Results from Model 3 showed that lower local distance was associated with improved object recognition task performance (Table 5c, Figure 7c). In addition, the interaction between group and local distance had a significant effect on performance. In the DS group, a reduced local distance was associated with better object recognition performance (simple slope:  $-1.28$ ,  $CI = [-2.04, -0.52]$ ), while no such relationship was found in the TD group (simple slope:  $-0.03$ ,  $CI = [-0.59, 0.53]$ ).

### 3.4.6 | Scene recognition

TD children exhibited significantly better performance on the scene recognition task compared to DS children (Table 4,  $p < .0001$ ). Results from Model 3 showed lower swap rate was also associated with improved scene recognition task performance for both TD and DS groups (Table 5c, Figure 7f). Additionally, lower local distance was linked to enhanced performance in the scene recognition task. The interaction between group and local distance had a significant impact on performance. Specifically, in the DS group, a reduced local distance was associated with better scene recognition performance (simple slope:  $-1.22$ ,  $CI = [-1.94, -0.50]$ ), while no such relationship was found in the TD group (simple slope:  $-0.21$ ,  $CI = [-0.74, 0.32]$ , Figure 7d).

### 3.4.7 | Correlations with original distance

To examine whether the original distance could be related to the neuropsychological tasks before any error transformation, we use the original distance to predict each neuropsychological memory task. Our findings indicate that there is a significant interaction between group and original distance in predicting neuropsychological memory tasks (Table 6), a shorter original distance was associated with better performance in each neuropsychological memory task in the TD group but not in the DS group (simple slope: digit span: TD:  $r = -0.64$ ,  $CI = [-0.83, -0.44]$ , DS:  $r = 0.001$ ,  $CI = [-0.23, 0.24]$ ; spatial memory: TD:  $r = -0.74$ ,  $CI = [-1.01, -0.47]$ , DS:  $r = -0.23$ ,  $CI = [-0.52, 0.06]$ ; objects in context 1 task: TD:  $r = -1.23$ ,  $CI = [-1.63, -0.83]$ , DS:  $r = -0.38$ ,  $CI = [-0.79, 0.04]$ ; objects in context 2 task: TD:  $r = -1.25$ ,  $CI = [-1.60, -0.91]$ , DS:  $r = -0.29$ ,  $CI = [-0.66, 0.07]$ ; object recognition: TD:  $r = -0.79$ ,  $CI = [-1.12, -0.47]$ , DS:  $r = -0.19$ ,  $CI = [-0.54, 0.17]$ ; scene recognition: TD:  $r = -0.84$ ,  $CI = [-1.16, -0.51]$ , DS:  $r = -0.28$ ,  $CI = [-0.63, 0.07]$ ).

## 4 | DISCUSSION

Previous studies have found that spatial memory ability develops across childhood and into adulthood, and that DS is associated with

spatial memory impairments relative to age expectations. We replicated those past findings here. Few previous studies have investigated the changes during development characterizing the spatial reconstruction errors contributing to inaccurate spatial memory across a wide age range, nor have these errors been compared between TD and DS groups, the latter constituting a population with known memory challenges and hippocampal dysfunction. In this study, we examined the performance of TD children (ages 3–18 years) and children with DS (6.8–17.9 years) on the spatial reconstruction task, differentiating the types of errors that contributed to inaccurate reconstruction. We found that in typically developing children, overall memory accuracy continually increases across the whole age range of our sample. However, examining the errors that contributed to the misplacement distances between the targets and the reconstructed objects, we found that although the ability to remember locations (i.e., regardless of identity) for both set size 2 and 3 trials and memory precision for the set size 2 trials increased with the similar pattern as the overall accuracy, the improvement in memory precision for set size 3 trials was only evident in TD children later than 7.5 years. In addition, the DS group showed the most difficulty with precision as compared to the KBIT-2 verbal and non-verbal raw score matched TD group in aggregated data across both 2 and 3 item trials. In total, these results suggest a protracted trajectory for the precision underlying spatial reconstruction memory.

The middle age of children (5.5–7.5 years) showed no differences with the younger children in set size 3 trials in local distance, suggesting that memory precision for more complex tasks might develop or mature later than 7 years old. These findings align with previous evidence suggesting that the hippocampal tail, which supports detailed memory, undergoes later development (Callaghan et al., 2021; Lee et al., 2020). As per the precision and binding model proposed by Yonelinas (2013), tasks necessitating complex and high-resolution bindings are predominantly dependent on the hippocampus. Our findings, using local distance as a memory precision metric, bolster this theoretical claim.

Previous studies of patients with hippocampal damage showed more swap errors relative to healthy controls (Horecka et al., 2018; Watson et al., 2013), and previous studies in DS have shown binding deficits (Nadel, 2003). However, we did not find that swap errors were significantly different in DS. We did find that the DS group showed greater local distance error than the TD matched group. This finding is consistent with the impaired high-resolution, but spared low resolution, place learning capacities documented previously in DS (Lavenex et al., 2015; Lavenex & Lavenex, 2021), and fits with previous studies that found that hippocampus damaged patients showed greater impaired precision than control participants (Horecka et al., 2018; Kolarik et al., 2016). The absence of swap error differences in this group may be due to the relative low demands and complexity of 2 versus 3 items in our task. Other metrics suggest that set-size differences were minimal and may have reflected the reduced binding demands of this task. As Ekstrom and Yonelinas (2020) suggested, many inconsistent results across memory studies may relate to the binding complexity and resolution required for a task, and here

these swap errors may demonstrate differences with other tasks administered in this group.

Our regression results showed that shorter local distance predicted better spatial memory, reflected by our spatial memory task within DS group and objects in context 1 task across both groups (the only difference between these two tasks is whether the objects to be remembered are the same or different). The non-significant effect of the correlation between the local distance and the spatial memory task in the TD group may stem from the fact that 29% of participants showed a ceiling effect, suggesting the spatial memory task, with its identical objects, might be less difficult for the TD children. In contrast, a more complex task—the objects in context 1, which involves more complex binding, displayed significant results across both groups despite 40% of DS children demonstrating a floor effect and 20% of TD children reaching a ceiling effect. Also, smaller local distance predicted better mnemonic discrimination capacity, hypothesized to be linked to pattern separation, reflected by the object recognition task and the scene recognition task in the DS group. It is worth noting that 38% and 24% of the TD children exhibited ceiling effects in the object recognition and scene recognition tasks respectively. This factor may account for the absence of a significant correlation between local distance and these two tasks within the TD group. Regardless, these correlations show that the precision measure is predictive across groups and tasks, providing additional evidence that spatial precision should be considered as a component when evaluating memory performance across development and in memory disordered groups.

Additionally, the swap rate inversely predicts scene recognition performance in both TD and DS groups, but not object recognition. The scene and object recognition tasks utilized in our study were designed to assess both pattern separation and memory recognition. This disparity between scene and object recognition results may be attributed to the intrinsic complexity of scene recognition, which requires the hippocampal-mediated binding of objects to their context, unlike object recognition tasks potentially engaging familiarity-based recognition regions such as the perirhinal cortex (Baxter & Murray, 2001; Brown & Aggleton, 2001; Yonelinas, 2013). Consequently, the observed correlation between swap rate and scene recognition further suggested the link between swap rate and the hippocampal binding. In addition, our study revealed a negative relationship between swap rate and spatial memory performance in the TD group, but not in the DS group. This finding suggests that individuals with a higher swap rate, indicative of increased binding errors, exhibit deficits in the spatial memory task. Given the uniformity of objects within the spatial memory task, this intriguing correlation merits further exploration.

There are several limitations of this study. First, we used cross-sectional data, which limits our ability to fully capture developmental trends. Future studies using a longitudinal design would provide more evidence for our findings. Another limitation of this study is the lack of full neuroimaging data to understand the relationship between the deficits and hippocampal network structure and function, because of the conjunctive reasons of COVID-19 and difficulty imaging pediatric groups. In addition, the wide age range of the old TD group of this study is another limitation, which constrains our ability to depict the

development trajectory in a fine-grained manner and resulted in some measurement difficulties, including some issues with ceiling effects in the oldest participants. Lastly, the set size manipulation incorporated here was restricted to 2 and 3 items, as larger set sizes could have caused more floor effects in the cohorts of young and moderately impaired children. This limitation might contribute to the inconclusive findings on the set size effect or the interaction effect involving set size. The question of how set size intersects with representational precision in memory performance begs for additional research, and particularly in pediatric groups and groups with memory impairment. Overall, our study sheds light on uneven maturational trajectories of memory abilities related to different kinds of error and future investigations should aim to understand links between hippocampal structure and hippocampal-cortical connectivity in these groups with tasks of this nature.

In summary, our findings indicate that TD children older than 7.5 years showed the most notable improvement in memory precision for set size 3 trials. Furthermore, the DS group experienced greater difficulties with precision compared to the matched TD group in both 2 and 3 item trials. The results of this study provide evidence of uneven developing spatial memory skills in the typically developing population. We have also captured divergent deficit patterns to those previously reported for hippocampal patients, that is, weak precision rather than the difficulty in remembering arbitrary identity, in comparing children with DS and TD. These studies add to our understanding of the basis for deficits in memory development in DS as well as what functions may continue to be refined across typical development. Mainly, here we see precision emerges as an important construct, both in our behavioral findings in DS as well as in correlations with other tasks. Future work should aim to understand the neural basis of these deficits, with one emerging hypothesis being that hippocampal-cortical communication may be a key driving force in this skill development.

#### AUTHOR CONTRIBUTIONS

Conceptualization: Maomiao Peng, Jamie Edgin. Data Collection: Kenneth Bottrill, Annalysa Lovos, Katharine Hughes, and Miranda Sampsel. Methodology, investigation, and analyses: Maomiao Peng, Jamie Edgin, Annalysa Lovos, and Kenneth Bottrill. Writing – original draft: Maomiao Peng. Writing – review & editing: Maomiao Peng, Jamie Edgin, Annalysa Lovos, Kenneth Bottrill, Nancy Raitano Lee, Leonard Abbeduto, and Angela John Thurman. All authors provided critical feedback for this manuscript. Supervision: Jamie Edgin. The data came from a large collaborative NIH-funded project led Jamie Edgin, Nancy Raitano Lee, Leonard Abbeduto, and Angela John Thurman.

#### ACKNOWLEDGMENTS

This work was funded by R01HD088409 (to Jamie Edgin) and P50HD103526 (Leonard Abbeduto, UC Davis). We are grateful to the families and participants that made this work possible. Ms. Peng also would like to thank Lynn Nadel and Ying-hui Chou for their participation on her MA committee. Some of this work was submitted as an MA thesis (Maomiao Peng).

#### CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

#### DATA AVAILABILITY STATEMENT

The data that support this study are available on request from the corresponding author ([jedgin@arizona.edu](mailto:jedgin@arizona.edu)) upon reasonable request.

#### ORCID

Annalysa Lovos  <https://orcid.org/0000-0002-9709-8360>

#### REFERENCES

- Allen, R. J., Atkinson, A. L., Vargha-Khadem, F., & Baddeley, A. D. (2022). Intact high-resolution working memory binding in a patient with developmental amnesia and selective hippocampal damage. *Hippocampus*, 32(8), 597–609.
- Baxter, M. G., & Murray, E. A. (2001). Opposite relationship of hippocampal and rhinal cortex damage to delayed nonmatching-to-sample deficits in monkeys. *Hippocampus*, 11(1), 61–71.
- Blankenship, S. L., Redcay, E., Dougherty, L. R., & Riggins, T. (2017). Development of hippocampal functional connectivity during childhood. *Human Brain Mapping*, 38(1), 182–201.
- Blankenship, T. L., & Kibbe, M. M. (2019). Examining the limits of memory-guided planning in 3- and 4-year olds. *Cognitive Development*, 52, 100820.
- Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, 2(1), 51–61.
- Brown, T. T., Kuperman, J. M., Chung, Y., Erhart, M., McCabe, C., Hagler, D. J., Venkatraman, V. K., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Casey, B. J., Chang, L., Ernst, T. M., Frazier, J. A., Gruen, J. R., Kaufmann, W. E., Kenet, T., Kennedy, D. N., Murray, S. S., ... Dale, A. M. (2012). Neuroanatomical assessment of biological maturity. *Current Biology: CB*, 22(18), 1693–1698.
- Callaghan, B., Gasser, C., Silvers, J. A., VanTieghem, M., Choy, T., O'Sullivan, K., Tomparly, A., Davachi, L., & Tottenham, N. (2021). Age-related increases in posterior hippocampal granularity are associated with remote detailed episodic memory in development. *Journal of Neuroscience*, 41(8), 1738–1754.
- Canada, K. L., Botdorf, M., & Riggins, T. (2020). Longitudinal development of hippocampal subregions from early- to mid-childhood. *Hippocampus*, 30(10), 1098–1111.
- Carducci, F., Onorati, P., Condoluci, C., Di Gennaro, G., Quarato, P. P., Pierallini, A., Sarà, M., Miano, S., Cornia, R., & Albertini, G. (2013). Whole-brain voxel-based morphometry study of children and adolescents with Down syndrome. *Functional Neurology*, 28(1), 19–28.
- Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. MIT Press.
- Conners, F. A., Moore, M. S., Loveall, S. J., & Merrill, E. C. (2011). Memory profiles of Down, Williams, and fragile X syndromes: Implications for reading development. *Journal of Developmental and Behavioral Pediatrics: JDBP*, 32(5), 405–417.
- Cooper, R. A., Richter, F. R., Bays, P. M., Plaisted-Grant, K. C., Baron-Cohen, S., & Simons, J. S. (2017). Reduced hippocampal functional connectivity during episodic memory retrieval in autism. *Cerebral Cortex*, 27(2), 888–902.
- Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. McGill University.
- Edgin, J. O. (2013). Cognition in Down syndrome: A developmental cognitive neuroscience perspective. *Wiley Interdisciplinary Reviews. Cognitive Science*, 4(3), 307–317.
- Eichenbaum, H., & Cohen, N. J. (2004). *From conditioning to conscious recollection: Memory systems of the brain*. Oxford University Press.



- Ekstrom, A. D., & Yonelinas, A. P. (2020). Precision, binding, and the hippocampus: Precisely what are we talking about? *Neuropsychologia*, *138*, 107341.
- Elliot, C. D. (1990). *Manual for the differential ability scales*. Psychological Corporation.
- Epstein, R. A. (2014). Neural systems for visual scene recognition. In K. Kveraga & M. Bar (Eds.), *Scene vision: Making sense of what we see* (pp. 105–134). MIT Press.
- Godfrey, M., & Lee, N. R. (2018). Memory profiles in down syndrome across development: A review of memory abilities through the lifespan. *Journal of Neurodevelopmental Disorders*, *10*(1), 5.
- Hassevoort, K. M., Khan, N. A., Hillman, C. H., & Cohen, N. J. (2020). Differential development of relational memory and pattern separation. *Hippocampus*, *30*(3), 210–219.
- Horecka, K. M., Dulas, M. R., Schwarb, H., Lucas, H. D., Duff, M., & Cohen, N. J. (2018). Reconstructing relational information. *Hippocampus*, *28*(2), 164–177.
- Hu, S., Pruessner, J. C., Coupé, P., & Collins, D. L. (2013). Volumetric analysis of medial temporal lobe structures in brain development from childhood to adolescence. *NeuroImage*, *74*, 276–287.
- JASP Team. (2023). JASP (Version 0.17.2.1) [Computer software].
- Kessels, R. P., van Zandvoort, M. J., Postma, A., Kappelle, L. J., & de Haan, E. H. (2000). The Corsi block-tapping task: Standardization and normative data. *Applied Neuropsychology*, *7*(4), 252–258.
- Keyser, C., Gazzola, V., & Wagenmakers, E. J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799.
- Koenig, K. A., Oh, S. H., Stasko, M. R., Roth, E. C., Taylor, H. G., Ruedrich, S., Wang, Z. I., Leverenz, J. B., & Costa, A. C. (2021). High resolution structural and functional MRI of the hippocampus in young adults with Down syndrome. *Brain Communications*, *3*(2), fcab088.
- Kohl, M. (2020). *MKinfer: Inferential statistics*. R package version 0.6. <http://www.stamats.de>
- Kolarik, B. S., Shahlaie, K., Hassan, A., Borders, A. A., Kaufman, K. C., Gurkoff, G., Yonelinas, A. P., & Ekstrom, A. D. (2016). Impairments in precision, rather than spatial strategy, characterize performance on the virtual Morris water maze: A case study. *Neuropsychologia*, *80*, 90–101.
- Konkel, A., Warren, D. E., Duff, M. C., Tranel, D. N., & Cohen, N. J. (2008). Hippocampal amnesia impairs all manner of relational memory. *Frontiers in Human Neuroscience*, *2*, 15.
- Lavenex, P. B., Bostelmann, M., Brandner, C., Costanzo, F., Fragnière, E., Kléncklen, G., Lavenex, P., Menghini, D., & Vicari, S. (2015). Allocentric spatial learning and memory deficits in Down syndrome. *Frontiers in Psychology*, *6*, 62.
- Lavenex, P. B., & Lavenex, P. (2021). A critical review of spatial abilities in down and Williams syndromes: Not all space is created equal. *Frontiers in Psychiatry*, *12*, 669320.
- Lee, J. K., Ekstrom, A. D., & Ghetti, S. (2014). Volume of hippocampal subfields and episodic memory in childhood and adolescence. *NeuroImage*, *94*, 162–171.
- Lee, J. K., Fandakova, Y., Johnson, E. G., Cohen, N. J., Bunge, S. A., & Ghetti, S. (2020). Changes in anterior and posterior hippocampus differentially predict item-space, item-time, and item-item memory improvement. *Developmental Cognitive Neuroscience*, *41*, 100741.
- Lee, J. K., Johnson, E. G., & Ghetti, S. (2017). Hippocampal development: Structure, function and implications. In D. E. Hannula & M. C. Duff (Eds.), *The hippocampus from cells to systems: Structure, connectivity, and functional contributions to memory and flexible cognition* (pp. 141–166). Springer International Publishing.
- Lee, J. K., Wendelken, C., Bunge, S. A., & Ghetti, S. (2016). A time and place for everything: Developmental differences in the building blocks of episodic memory. *Child Development*, *87*(1), 194–210.
- Mathy, F., & Varré, J. S. (2013). Retention-error patterns in complex alphanumeric serial-recall tasks. *Memory*, *21*(8), 945–968.
- Nadel, L. (2003). Down's syndrome: A genetic disorder in biobehavioral perspective. *Genes, Brain, and Behavior*, *2*(3), 156–166.
- 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*(9), 1463–1479.
- Parker, S. E., Mai, C. T., Canfield, M. A., Rickard, R., Wang, Y., Meyer, R. E., Anderson, P., Mason, C. A., Collins, J. S., Kirby, R. S., Correa, A., & National Birth Defects Prevention Network. (2010). Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004–2006. *Birth Defects Research. Part A, Clinical and Molecular Teratology*, *88*(12), 1008–1016.
- Pennington, B. F., Moon, J., Edgin, J., Stedron, J., & Nadel, L. (2003). The neuropsychology of Down syndrome: Evidence for hippocampal dysfunction. *Child Development*, *74*(1), 75–93.
- Perrett, D. I., & Oram, M. W. (1993). Neurophysiology of shape processing. *Image and Vision Computing*, *11*(6), 317–333.
- Pinter, J. D., Brown, W. E., Eliez, S., Schmitt, J. E., Capone, G. T., & Reiss, A. L. (2001). Amygdala and hippocampal volumes in children with Down syndrome: A high-resolution MRI study. *Neurology*, *56*(7), 972–974.
- Poldrack, R. A. (2010). Interpreting developmental changes in neuroimaging signals. *Human Brain Mapping*, *31*(6), 872–878.
- Ribordy Lambert, F., Lavenex, P., & Banta, L. P. (2015). Improvement of allocentric spatial memory resolution in children from 2 to 4 years of age. *International Journal of Behavioral Development*, *39*, 318–331.
- Smigielska-Kuzia, J., Bockowski, L., Sobaniec, W., Sendrowski, K., Olchowik, B., Cholewa, M., Lukasiewicz, A., & Lebkowska, U. (2011). A volumetric magnetic resonance imaging study of brain structures in children with Down syndrome. *Neurologia i Neurochirurgia Polska*, *45*(4), 363–369.
- Steingass, K. J., Chicoine, B., McGuire, D., & Roizen, N. J. (2011). Developmental disabilities grown up: Down syndrome. *Journal of Developmental and Behavioral Pediatrics: JDBP*, *32*(7), 548–558.
- Stevenson, R. F., Zheng, J., Mnatsakanyan, L., Vadera, S., Knight, R. T., Lin, J. J., & Yassa, M. A. (2018). Hippocampal CA1 gamma power predicts the precision of spatial memory judgments. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(40), 10148–10153.
- Tamnes, C. K., Walhovd, K. B., Dale, A. M., Østby, Y., Grydeland, H., Richardson, G., Westlye, L. T., Roddey, J. C., Hagler, D. J., Jr., Due-Tønnessen, P., Holland, D., Fjell, A. M., & Alzheimer's Disease Neuroimaging Initiative. (2013). Brain development and aging: Overlapping and unique patterns of change. *NeuroImage*, *68*, 63–74.
- Uematsu, A., Matsui, M., Tanaka, C., Takahashi, T., Noguchi, K., Suzuki, M., & Nishijo, H. (2012). Developmental trajectories of amygdala and hippocampus from infancy to early adulthood in healthy individuals. *PLoS One*, *7*(10), e46970.
- Umeyama, S. (1991). Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*, 376–380.
- Uttal, D. H., & Chiong, C. (2004). Seeing space in more than one way: Children's use of higher order patterns in spatial memory and cognition. In G. L. Allen (Ed.), *Human memory: Remembering where* (pp. 125–142). Erlbaum.
- Wang, J. X., Cohen, N. J., & Voss, J. L. (2015). Covert rapid action-memory simulation (CRAMS): A hypothesis of hippocampal–prefrontal interactions for adaptive behavior. *Neurobiology of Learning and Memory*, *117*, 22–33.
- Watson, P. D., Voss, J. L., Warren, D. E., Tranel, D., & Cohen, N. J. (2013). Spatial reconstruction by patients with hippocampal damage is dominated by relational memory errors. *Hippocampus*, *23*(7), 570–580.
- Yim, H., Dennis, S. J., & Sloutsky, V. M. (2013). The development of episodic memory: Items, contexts, and relations. *Psychological Science*, *24*(11), 2163–2172.
- Yonelinas, A. P. (2013). The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. *Behavioural Brain Research*, *254*, 34–44.

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45-57.

**How to cite this article:** Peng, M., Lovos, A., Bottrill, K., Hughes, K., Sampsel, M., Lee, N. R., Abbeduto, L., Thurman, A. J., & Edgin, J. (2023). Extended trajectory of spatial memory errors in typical and atypical development: The role of binding and precision. *Hippocampus*, 1-18. <https://doi.org/10.1002/hipo.23576>

## APPENDIX A: MRI methods

MR images were acquired on a 3 T Siemens Magnetom Skyra scanner using a 32-channel head coil. Structural images were acquired with a T1 MPRAGE sequence (resolution 1.0 mm<sup>3</sup>, 208 slices, 256 × 256 acquisition matrix, flip angle = 9°, FOV = 256 mm, TR 2300 ms). The T1-weighted images were processed using FSLFAST for volumetric estimates (Zhang et al., 2001). Bootstrapped t-tests were performed in RStudio 2021.09.0.351 using the MKinfer package (Kohl, 2020).