Developmental Review 72 (2024) 101119

Contents lists available at ScienceDirect

Developmental Review

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

Building a cumulative science of memory development

Chi T. Ngo [a,*], Elisa S. Buchberger [a], Phuc T.U. Nguyen [a], Nora S. Newcombe [b], Markus Werkle-Bergner [a,*]

[a] *Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany*
[b] *Department of Psychology, Temple University, Philadelphia, USA*

ARTICLE INFO

Keywords:
Mapping review
Methodology
Memory development
Cumulative science
Pattern separation

ABSTRACT

Cumulative science hinges on consolidating empirical evidence. However, both narrative reviews and *meta*-analyses often restrict integration by using construct labels in searches, which vary greatly across traditions and eras in psychology. A complementary approach is the mapping review, more common in other disciplines, which focuses on operational definitions and experimental design. Here, we pioneered using this approach in experimental psychology, applying it to memory development. We classified 506 publications in the last 5 decades using 16 design features and identified methodologically convergent work that originated from different theoretical frameworks. New insights emerge from this approach. First, characterizing the relative data density along each dimension of experimental design identifies *where* future research is most necessary. Second, we linked relevant evidence previously separated by nonoverlapping construct labels. Third, we illustrated the potential application of this technique as a precursor to subsequent research syntheses with an analysis of the development of pattern separation (also called mnemonic discrimination). To facilitate the process of literature integration and identification of methodological overlap, we have created a freely available interactive web application using the current database.

Introduction

By most standards, young children have less robust memory capacities than adults. But how do children develop to adult levels? What kinds of memory do children particularly struggle with at an early age? In what order do different aspects of the memory system emerge and how do they become coordinated? What does the development of memory tell us about a mature memory system? These questions have long fascinated developmental and cognitive psychologists alike. In the last five decades, using an array of different theories, researchers have made exciting discoveries regarding memory development. These theories include multi-system memory models that emphasize the dissociations between explicit and implicit memory (Gabrieli et al., 1995; Graf & Schacter, 1985), and between episodic and semantic memory (Tulving, 1972), as well as the source monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993), script theory (Farrar & Boyer-Pennington, 1999), and fuzzy-trace theory (Brainerd & Reyna, 1990). Most recently, developmental studies have found inspiration from computational models of memory, especially the Complementary Learning Systems theory (Norman & O'Reilly, 2003). Building a cumulative science of memory development hinges on our abilities to effectively synthesize existing knowledge across theories, such that new insights can be built upon an integrated body of knowledge (see also the

---

* Corresponding authors.
E-mail addresses: ngo@mpib-berlin.mpg.de (C.T. Ngo), werkle@mpib-berlin.mpg.de (M. Werkle-Bergner).

https://doi.org/10.1016/j.dr.2024.101119
Received 29 June 2023; Received in revised form 13 December 2023;

discussion in Curran, 2009). However, the use of different theoretical frameworks and terms tends to obscure the commonalities among experiments and impedes the cumulation of knowledge.

We propose a solution to this problem by conducting a *method-focused mapping review* to improve stock-taking of the available evidence on memory development. Instead of relying on construct labels (i.e., terminology), which evolve quickly, we developed a database that catalogues and collates experimental attributes of the literature to uncover the methodological convergence between research that stems from distinct theoretical frameworks. Unlike narrative reviews and *meta*-analyses that target specific conceptual questions, mapping reviews describe and quantify the empirical research activities in a particular area (Grant & Booth, 2009). With this database, we mapped the available evidence based on experimental methods found in the literature on memory development to identify trends, clusters, and gaps on the empirical basis. We then built on a unique feature of mapping reviews by estimating the method similarity between studies. A strength of mapping reviews is that they can widen a network of relevant research for more comprehensive narrative reviews and *meta*-analyses that follow (Grant & Booth, 2009).

We begin with a broad-stroke description of each theory and the representative paradigms employed to study children's long-term memory capacities to illustrate why gaining an aerial view of the existing methodologies is a necessary catalyst in building a cumulative science in this research area.

The memory systems account

The dissociation between explicit and implicit memory has been extremely influential in the learning and memory literature in general (Schacter et al., 1993). *Explicit memory* refers to the conscious, intentional recollection of the past including facts, concepts, and experiences. In contrast, *implicit memory* is the form of memory that facilitates subsequent performance on a given task via exposure, even in the absence of conscious recollection of such exposure (Gabrieli et al., 1995; Graf & Schacter, 1985). The dissociation of various memory systems was rooted in the observation that different forms of learning and memory rely on separable neural systems in the brain (e.g., Buckner et al., 1995; Thompson & Kim, 1996). The idea of multiple memory systems motivated researchers to test whether such systems develop differently. Children's explicit memory is measured in a wide range of tasks that require a conscious recollection of learned items (Drummey & Newcombe, 1995; Greenbaum & Graf, 1989; Gulya et al., 2002; Guttentag & Dunn, 2003), associations (Murphy et al., 2003), stories (Hömberg et al., 1993), or events (reviewed in Bauer, 2007) in a laboratory setting, or the retrieval of a child's personal events through interview techniques (Morris et al., 2010). Implicit memory is measured via a change in task performance as a function of prior experience with a given stimulus or a procedural skill, including sequence learning (Brown et al., 2010; Fischer et al., 2007; Thomas & Nelson, 2001), perceptual priming (Burden & Mitchell, 2005; Drummey & Newcombe, 1995; Murphy et al., 2003), conceptual priming (Murphy et al., 2003), and statistical learning tasks (Saffran et al., 1996).

Another aspect of multi-system memory theories is Tulving's proposal of a distinction between semantic and episodic memory systems (Tulving, 1972). Although both forms are explicit memory, *episodic memory* involves the recollection of a specific episode tied to its spatiotemporal context, whereas *semantic memory* refers to a dynamic collection of general knowledge of facts and concepts that are disconnected from the place and time of initial learning (Ghosh & Gilboa, 2014; Tulving, 1972). It is worth noting that supporting evidence for the division between episodic and semantic memory systems is the observation that these kinds of memory follow different developmental trajectories (Tulving, 1972). Children's episodic memory has been measured with a rich repertoire of tasks, ranging from associative memory between an item and its contexts (e.g., pairmate (Lee et al., 2020), co-occurring background (Lloyd et al., 2009; Ngo, Horner, et al., 2019; Ngo, Lin, et al., 2019; Ngo et a., 2017; Sluzenski et al., 2006), temporal order with other items (Bauer & Zelazo, 2013; Canada et al., 2020; Lee et al., 2016; Pathman et al., 2018; Pathman & Ghetti, 2014; Pathman et al., 2013), temporal memory (e.g., item-when associations (Cuevas et al., 2015; Hayne & Imuta, 2011; Lee et al., 2020; Lee et al., 2016; Mastrogiuseppe et al., 2019; Picard et al., 2012), temporal order of subunits within a sequence (Dikmen et al., 2014), a stage event (Bauer et al., 2012), a personal event (Pathman, Doydum, et al., 2013; Pathman, Larkina, et al., 2013), or items associated with both the "where" and "when" aspect of an event (Burns et al., 2015; Cheke & Clayton, 2015). In contrast, semantic memory is measured by testing existing conceptual and category knowledge, such as children's abilities to identify various items, their broader categories, inter-item conceptual relationships (Ceci et al., 2007; Unger et al., 2016; Unger et al., 2020), or schematic knowledge of events.

The source monitoring framework

Source Monitoring Framework posits that source monitoring – tracing the origin of information acquisition – involves complex attribution processes (Johnson et al., 1993). Specifically, source monitoring requires the evaluation of memory characteristics such as perception, contextual information, or emotional reactions through strategic decision making. Similar to some kinds of episodic memory assessments, source monitoring is tested in paradigms that require memory discrimination between sources of an item (Lindsay et al., 1991) (e.g., which person said the word "apple") or of an action (Drummey & Newcombe, 2002; Foley & Johnson, 1983; Foley et al., 1987; Foley et al., 1989; Foley & Johnson, 1985) (e.g., who baked the cake). This theory provides crucial insights into the conditions that lead to source memory confusions in early development. For example, findings showing that source similarity magnifies memory confusion, especially in children (Lindsay et al., 1991), have been taken to underscore the role of strategic processes in children's monitoring development (Shing et al., 2010).

Script theory

Relevant to semantic memory are studies that examined the development of scripts (Nelson & Gruendel, 1981). Scripts are generic representations of events, which contain organized knowledge of the interconnected event structures. According to script theories, we build general event representations from the repeated exposure of similar events (e.g., what happens at restaurants: Fivush et al., 1992; Hudson et al., 1992; Nelson & Gruendel, 1981; Schank & Abelson, 1977). This theory motivated developmental researchers to contrast children's knowledge for routine events (e.g., what usually happens at a grocery store) and their detailed memories for a specific instance of an event category (e.g., what happened last time you went to a grocery story) (Hudson & Mayhew, 2009). Children's script memory is also measured in semi-naturalistic event paradigms that manipulate event repetition and deviations across specific instances (Brubacher et al., 2018; Brubacher et al., 2012; Brubacher, Glisic, et al., 2011; Brubacher, Roberts, et al., 2011; Connolly & Price, 2006; Danby et al., 2017; Danby et al., 2017; Price & Connolly, 2013; Roberts et al., 2015; Roberts & Powell, 2006).

Fuzzy trace theory

Fuzzy trace theory proposes that there are parallel stored representations of each experience: a gist representation that entails the overall patterns and meanings, and a verbatim representation that preserves the precise surface form as it was presented (Brainerd & Reyna, 1990, 1998). This distinction is reminiscent of the semantic-episodic distinction. Developmental researchers in this tradition employ paradigms that dissociate behaviors associated with each kind of memory. For instance, verbatim memory is reflected by true memory for the exact items learned from a list, whereas gist memory is tied to false memory for the semantically related lures (i.e., the Deese-Roediger-McDermott [DRM] paradigm (Brainerd & Reyna, 1998; Carneiro et al., 2007; Dewhurst et al., 2007; Holliday et al., 2008; Howe, 2006, 2008; Howe & Wilkinson, 2011; McGeown et al., 2014; Metzger et al., 2008)). In another paradigm, verbatim memory is indexed by children's memories of the specific associations taught to them (e.g., A $>$ B, B $>$ C), whereas gist memory is indexed by their abilities to make transitive inference that transcend the learned associations (A $>$ C) (Brainerd & Reyna, 1993; Reyna & Brainerd, 1995).

Neurocomputational accounts

Most recently, insights from neurocomputational models of memory have ignited considerable interest among developmental researchers (Gómez & Edgin, 2016; Keresztes et al., 2018; Lee et al., 2017; Ramsaran et al., 2019). The Complementary Learning Systems (CLS) theory provides a useful framework for understanding the neural processes that support both the need for accumulating and applying generalizable knowledge, and the need to remember the specific instances (McClelland et al., 1995; Norman & O'Reilly, 2003). The balancing act between the two opposing learning schemes lies in a labor division between the neocortex and the hippocampus. The neocortex slowly learns the commonalities across experiences, thereby accumulating knowledge that can support *generalization*. In contrast, the hippocampus learns specific episodes via *pattern completion* and *pattern separation* processes. Pattern completion enables cued retrieval such that elements that co-occurred in the same event are retrieved together, thereby facilitating the holistic retrieval of a given experience from its partial cue (Norman & O'Reilly, 2003; Rolls, 2016). Pattern separation aids memory discrimination of similar experiences by reducing the degree of representational similarity among overlapping experiences (Norman & O'Reilly, 2003; Rolls, 2016).

To study the development of generalization, pattern completion, and pattern separation processes, some developmental researchers have adapted paradigms from the adult memory literature (reviewed in Liu, et al., 2016; Schapiro et al., 2017; Zeithamova & Bowman, 2020)). They focused especially on the paradigms in which the neural correlates aligned with putative signatures of these processes shown in adults. Children's generalization abilities have been estimated in a wide range of tasks, ranging from statistical learning paradigms that measure the detection of repeated patterns (Arciuli & Simpson, 2011; Jung et al., 2020; Pudhiyidath et al., 2020; Shufaniya & Arnon, 2018), to those that measure the application of such learned patterns to new situations, including transitive inference, categorization(Blanco & Sloutsky, 2019; Deng & Sloutsky, 2012), associative inference (Schlichting et al., 2017), and inference generation (Jung et al., 2020; Ngo et al., 2021; Pudhiyidath et al., 2020). The literature on pattern completion development is still thin, with only a handful of studies that operationalized this construct as holistic recollection by estimating retrieval success contingency among associations from the same episode (James et al., 2021; Ngo, Horner, et al., 2019). Within the past decade, there has been a surge of interest in the idea of pattern separation. Aligned with its operational definition in the adult memory literature (Lacy et al., 2011; Stark et al., 2013), previous studies typically measure children's abilities to discriminate similar memories by implementing a *similarity manipulation* that would interfere with the to-be-remembered information. A similarity manipulation requires participants to discriminate studied items from similar lures at test (e.g., a perceptually similar rubber duck to the one shown at encoding) (Mnemonic Similarity Task: Bouyeure et al., 2021; Canada et al., 2018; Hassevoort, et al., 2020; Keresztes et al., 2017; Ngo, Newcombe, & Olson, 2017; Rollins & Cloude, 2018).

Evaluating and integrating the *empirical* ***evidence***

This short review should highlight the challenge faced by reviewers of the memory development literature, namely, to consolidate empirical work that is saturated with terminologies, construct labels, and task names that will bias key word searches. Similarity among studies' experimental methodologies is often obscured by the discrepancy in terminologies across different frameworks, and consequently, disintegrate inherently related work. Adding another layer of complication is the fact that tasks that ostensibly measure overlapping processes can share minimal procedural similarity, especially in studies with developmental populations (e.g., reenactment of action sequences in children (Bauer et al., 2012; Brown & Murphy, 1975) versus temporal reconstruction of learned images in adults). Measuring children's memory requires a great deal of creativity, as it heavily relies on techniques that minimize the contamination of verbal abilities, boredom, and procedural demands. Such adaptation requires careful consideration of construct validity and construct equivalence across paradigms. The combination of terminology discrepancy and the variety of experimental paradigms creates a mammoth roadblock in our effort towards cumulative science (Moreau & Wiebels, 2022).

One solution to this problem is to systematically re-examine the literature solely through the lens of methodologies, disconnected from theories, task and construct labels. Given that operational definition lies at the heart of experimental psychology, methodology and experiments should spearhead the organization and consolidation of literature in our discipline. Building a cumulative science critically rests on *literature integration* that enables iterative reciprocity between theory and data. However, integration can only be effective when common terminology makes clear what topics are investigated and what phenomena are illuminated by each study. Although there are convergent aspects across different theories, as they build upon and complement one another, the construct labels and their interpretations vary remarkably. The discrepancy in construct labels can disconnect inherently related bodies of the literature and hamper our effort in fostering cumulative science (e.g., in subdisciplines within psychology: Bruce et al., 2014; Green et al., 2008; Iyer et al., 2021; Roberts et al., 2018; medical care: Roze des Ordons et al., 2018; psychiatry: Pawar & Spence, 2003). While narrative reviews and *meta*-analyses are extremely useful in synthesizing findings across studies and exposing open questions on a given topic, such syntheses are restricted within a research arena confined by boundaries created, in part, by construct labels.

In this paper, we conducted a mapping review centered on aspects of methodologies, devoid from their construct labels, to re-characterize the composition of the memory development literature. By breaking down the barriers that compartmentalize research areas within memory development, we can begin to identify convergence across seemingly disconnected literatures. Mapping review has been employed by other disciplines including intervention studies (Brigden et al., 2019; Cruwys et al., 2021; Fernández-Sotos et al., 2018; Peters & Wood, 2017), neurology (Biggin et al., 2020), health sciences (Paré & Kitsiou, 2016; Paré et al., 2015), educational research (Light & Smith, 1971), environmental sciences (James et al., 2016), software engineering (Petersen et al., 2008) and policy (Esdaile et al., 2019). Surprisingly, the application of this research synthesis approach has yet to appear in experimental cognitive psychology. Here, we conducted a multidimensional classification of prior empirical work on memory development based on methodologies to exemplify the utility of this approach.

The instrumentation of this approach enabled us to achieve three goals. First, we *quantitively* characterized the composition of the memory development literature based on a broad set of method factors, thereby demonstrating the relative coverage of existing data in

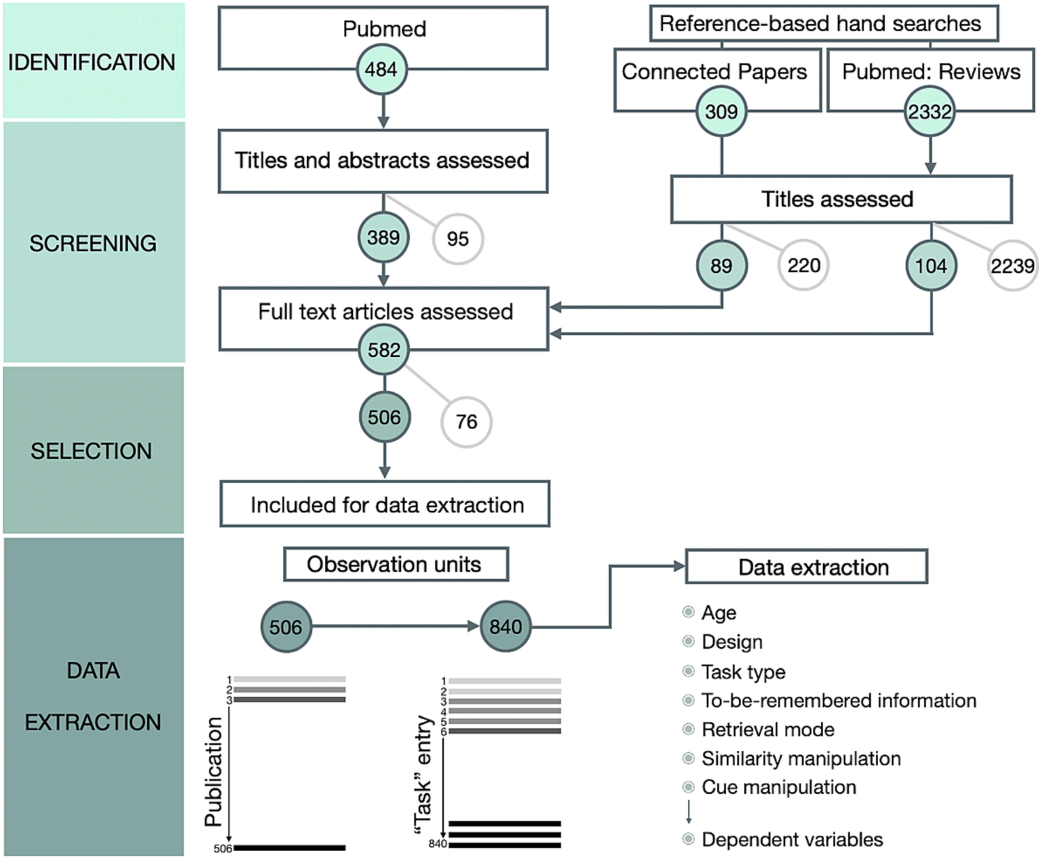

Fig. 1. A schematic pipeline of the four steps: (i) identification of relevant papers via PubMed and hand search strategies, (ii) the screening procedure based on a-priori inclusion and exclusion criteria, (iii) the selection of relevant papers, (iv) the application of the coding scheme on every task entry (observation units) from the selected papers.

various areas of investigation in memory development from the past 5 decades. Second, we relied on mapping approaches to connect publications strictly based on the convergence of their experimental designs, despite their remote theoretical inspirations and disjointed scholarship. Third, we explicated and exemplified the potential utility of this database by posing a specific question: *what additional evidence can be gained on the development of pattern separation across childhood*? We chose pattern separation because it stems from contemporary models of memory, ones that are propelling the current wave of interests in memory development research. We then contrasted the degree of the empirical evidence relevant to this question with and without the application of our approach, underscoring the theoretical stake in bridging across distant literatures.

Methods

Overall procedure

Our method entailed five main steps: (i) identification of relevant publications, (ii) screening of publications with a pre-determined set of inclusion and exclusion criteria, (iii) selection of relevant papers, (iv) development and application of the coding scheme to each selected publication, and (v) a visualization of each task entry and an aggregated map of the literature composition (see Fig. 1).

Identification

We identified relevant publications via two search strategies: PubMed and two reference-based hand searches.

PubMed. Our primary search strategy was based on PubMed, conducted in an iterative fashion spanning April 2020 to December 2021. All queries were done with the selection of “title/abstract”, and with a specification of “humans” under species. We carried out 5 total searches at three different time points:

a) Search term: “memory development”, spanning from 1970 to April 2020, which returned 311 papers. This step was conducted in April 2020.

Upon screening through the papers retrieved from this search, we noticed relatively poor coverage of three specific areas of memory development, namely explicit memory, implicit memory, and visual statistical learning. We performed three additional searches in July 2021:

b) Search terms: “development” AND “explicit memory”, spanning from 1970 to July 2021, which returned 65 papers. This step was conducted in July 2021.

c) Search terms: “development” AND “implicit memory”, spanning from 1970 to July 2021, which returned 61 papers. This step was conducted in July 2021.

d) Search terms: “development” AND “visual statistical learning”, spanning from 1970 to July 2021, which returned 9 papers. This step was conducted in July 2021. Note that we limited our search terms to “visual statistical learning” instead of “statistical learning” because statistical learning is a large literature encompassing language development, bilingualism, and computational linguistic studies.

Lastly, given that the literature review process spanned a period of 18 months, we repeated the original PubMed search (step a) on December 2021 to identify recent relevant publications.

e) “memory development”, from 2020 to December 2021, which returned 38 papers. A total of 484 publications resulted from these 5 searchers.

Hand search. To supplement the PubMed search and to increase the overall topical coverage of the literature, we conducted two hand searches. The first, and less systematic, hand search was done via Connected Papers, a webpage tool that provides exhaustive lists of retroactive and proactive references of the seed papers (https://www.connectedpapers.com/). We arbitrarily chose three published reviews and book chapters on specific topics within memory development, including implicit memory (Lloyd & Miller, 2014), personal memory (Peterson, 2002), eyewitness memories (Roberts, 2002). These topics, based on our judgment, appeared to show poor coverage from our original PubMed search. Each of the 3 reviews served as the seed paper. Based on the article titles of all three papers' references ($n = 309$), 96 papers were included in the screening step.

Second, we conducted a systematic hand search using PubMed. Again, we used the search term “memory development” within the period of 1970- September 2021, but this time with the article type selection of *meta*-analysis, review, and systematic review. We again set a specification of “human” under species. This search returned 75 papers. One member of the research team determined the topical relevance of each paper based on its title and the first paragraph. A total of 22 (21 reviews: Brown & Chiu, 2006; Cuevas & Sheya, 2019; Cycowicz, 2000; de Haan et al., 2006; Fivush, 2011; Gómez et al., 2018; Gómez & Edgin, 2016; Howe, 2013; Howe & Courage, 1993; Johnson et al., 2020; Mullally & Maguire, 2014; Nelson, 1998; Nelson, 2007; Ofen, 2012; Ofen et al., 2019; Ornstein, 1995; Pathman & Ghetti, 2016; Robinson et al., 2012; Rovee-Collier & Cuevas, 2009; Sander et al., 2012; Schneider & Ornstein, 2019) and 1 *meta*-analysis (Wu & Jobson, 2019) were identified as relevant to the topic of memory development in childhood. Next, a team of 4 researchers performed hand searches on all references ($n = 2332$) pooled across the 22 papers. Based on the article titles, 104 papers that did not overlap with the previously identified papers were included in the screening step.

Publication screening and selection

Inclusion criteria. The identified papers were subsequently screened by a team of 10 coders. There were five inclusion criteria. First, we only selected papers that included any age group within the range of age 3 to 10, excluding infants and adolescents. This decision was based on our primary interest in childhood. Second, we only selected papers that reported on a group or groups of typically-developing children. Third, we selected papers that reported empirical data, and thus, reviews, book chapters, or *meta*-analyses were excluded. Fourth, we only selected published papers from journals; thus, preprints and dissertations were excluded. Fifth, we only selected papers that were written in English.

Exclusion criteria. Papers with the following topical coverage or task types were excluded: (i) working memory, (ii) one-back memory task that requires retrieval immediately after the presentation trial (ii) *meta*-memory without memory veracity measures, (iii) prospective memory, (iv) forensically-oriented studies that include emotional or traumatic memory, (v) temporal duration judgment, (vi) theory of mind about false memory, and (vii) fear conditioning. Applying these sets of inclusion and exclusion criteria resulted in a total of 506 papers, published in 106 journals.

Coding scheme development and application

We developed a coding scheme that optimally captures the multi-dimensional design of a given memory task. The primary goal was to extract the relevant data on each methodological variable of each memory task, such that the combination of such variables encapsulates the key design ingredients. In our view, the most important empirical characteristics entail: (1) the basic design that captures age-differences or developmental change (i.e., cross-sectional vs. longitudinal design), (2) the broad classification of the memory task (i.e., task type), (3) the kind of memoranda being tested (i.e., to-be-remembered information). Importantly, inspired by the complementary neurocomputations propounded by contemporary models, we classified the experimental manipulations pertinent to pattern separation and pattern completion processes (i.e., similarity manipulations at (4) encoding or (5) retrieval, and (6) cue manipulations, respectively). Ancillary features including (7) participants' age, (8) encoding instruction, (9) encoding repetition, (10)

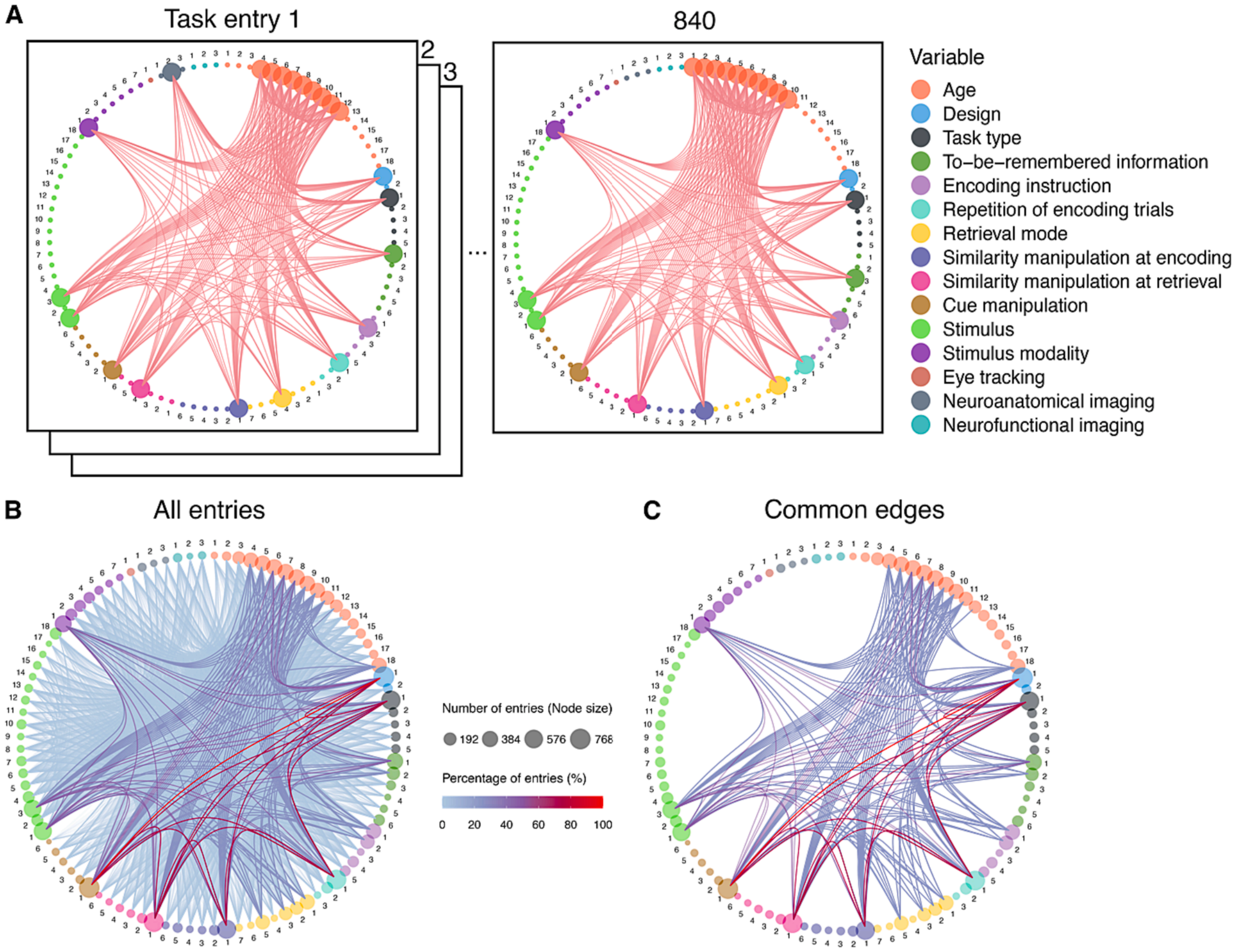

Fig. 2. (A) The methods employed in every task entry are visualized in a hierarchical edge bundling (HEB) plot by denoting which categories of a given variable are present, shown as enlarged nodes. For example, task entry 1 included ages 4 to 12, whereas task entry 840 included ages 1 to 10. An aggregated HEB of all task entries in the entire database. The number of entries in each category of a given variable is denoted by node size, with larger node sizes being greater number of entries. The color intensity of the edges denotes the frequency with which two categories co-occur in the same task entry. (C) A "trend" HEB that only includes the common edges, defined as those found in at least 20% of all task entries.

retrieval mode, (11) stimuli, (12) modality, (13) delay periods, and three physiological measures (14) eye tracking, (15) brain structural imaging, and (16) brain functional imaging were also recorded. Behavioral dependent variables were recorded but were not included in the similarity analyses. Based on this objective, our coding scheme contained 16 methodological variables. Two members of the research team (C. N. and E. B.) trained a team of seven research assistants on extracting the data from each memory task by exclusively relying on the methods section from each publication. Only the dependent variables were extracted from the results sections. The team continuously discussed the application of the coding scheme between April 2020 until December 2021 to resolve ambiguous cases. To estimate the extent of coding errors, we randomly selected over half (283 papers, 56 %) of the papers to be checked for coding accuracy by a second coder after the completion of the data extraction phase, and found that only 18 papers needed to be recoded.

Unit of observation. Departing from other mapping reviews (Fernández-Sotos et al., 2018; Glendøs & Berliner, 2017; McClain et al., 2021; McGrady & Hommel, 2016; Sadoughi et al., 2020; Symmank et al., 2017), our main unit of observation is *memory task* instead of publication. A single publication can include a large number of tasks across multiple experiments. Further, multiple tasks can be nested in one step-wise memory paradigm. For example, in an associative memory paradigm, participants are typically first tested on their memories for an individual item, followed by their memories of its associated context or pairmate. In this case, two task entries are needed to capture the targeted processes that support item memory versus associative memory. Therefore, task count, as opposed to publication count, better reflects the literature composition. The application of our coding scheme resulted in 840 task entries from the selected 506 publications.

Publication information. For each publication, we documented the authors' names, title, year of publication, publishing journal, location of data collection (if indicated), and keywords (if applicable). We assigned each paper with a unique identification which included the first author's last name and the publication year. For cases in which the same author's name and year applied for multiple publications, a letter in an alphabetical order was added (e.g., Bauer 2016b).

Variables

We defined different categories of each variable and classified them accordingly for each task entry (for full details, see Table S1).

Hierarchical edge bundling visualization method

Visualizing individual task entries. To convey the compound information from the variables in each task entry, we employed hierarchical edge bundling. Hierarchical Edge Bundling (HEB) refers to a visualization approach using graph hierarchies and edge bundles to cluster relations between hierarchically organized variables (Holten, 2006). Our HEB plot is organized in a hierarchy of 15 variables (e.g., age, design, encoding instruction), shown in distinct colors. Each node represents a category of a given variable. For example, node 1 of the age variable denotes age 1, whereas node 1 of the design variable denotes cross-sectional design. The nodes' numerical labeling align with the order of categories within a given variable specified in our coding scheme. When visualizing a single task entry (Fig. 2A), each enlarged node indicates that the corresponding category of a variable was employed, whereas the others were absent. Every two "activated" nodes were connected by an edge, simply denoting the co-occurrence of two categories in a task entry. This visualization technique provides an information-rich finger print for every individual task entry.

Visualizing the corpus. Aggregating across all 840 individual task entries yields a corpus HEB, showing the relative frequency with which each method's category was employed by researchers in the field (see Fig. 2B). The number of task entries that possess a given

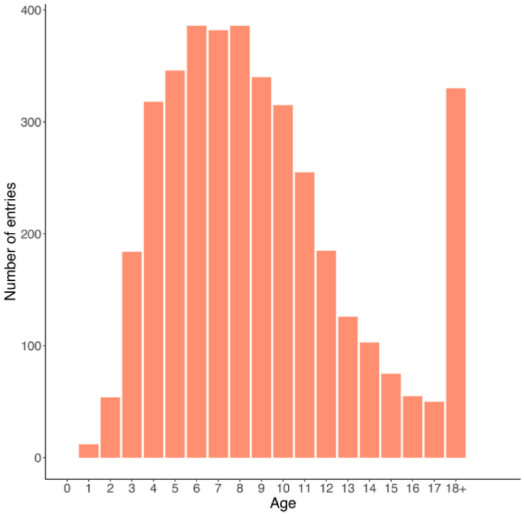

Fig. 3. The distribution of age groups in the database.

variable's category is denoted by its respective HEB node size. Pairwise connections between nodes, or edges, indicate the co-occurrence of two attributes in the same task entry. The color of the edges denotes the proportion of task entries that have the co-occurrence of two given attributes (Fig. 2B). Bundling and color-coding edges therefore highlight the trends in combinations of methods in a corpus. Trends in the literature are reflected by the "common" edges, defined as those found in at least 20 % of the task entries (Fig. 2C).

Results and discussion

Overall literature composition

First, we found that the locations of data collection were spread across 23 countries (see Figure S2), with the densest areas of data collection in North America and Europe. Second, to quantitively characterize the composition of the memory development literature, we amalgamated the whole corpus of all task entries in hierarchical edge bundling to simultaneously illustrate: (i) the relative proportion of each methodological variable, denoted as node sizes for each variable, and (ii) the relative density in the *combinations* of those variables in the same task entry, denoted as color intensity of the connections (see Fig. 2B). Selecting inter-node connections found in at least 20 % of whole corpus reveals methodological trends in the literature (see Fig. 2C).

Age. Given our age inclusion criteria, the age distribution in the database is most dense in the childhood period spanning age 3 to 12 (see Fig. 3). The high frequency of young adults indicates that testing young adults as a comparison group is a common practice in memory development. Note that all adults, whose age was 18 and above, are binned into one group.

Design. One striking finding from this work is that the memory development literature on childhood primarily consists of cross-sectional studies. Of the 840 task entries, only 9 % of the task entries (39 publications) employed a longitudinal design (see Figs. 4 and 5). The sparsity in empirical evidence on intra-person changes in memory development over time calls for a much greater emphasis on longitudinal studies in the future (Lindenberger et al., 2011). Cross-sectional age-comparative studies always present a mixture of effects due to age- and inter-individual differences and thus cannot reveal the effects of within-person changes due to development. In fact, some studies have flagged important differences in the age patterns observed from cross-sectional and longitudinal designs (Keresztes et al., 2022). Therefore, increasing our effort in directly testing developmental changes in mnemonic functions with longitudinal designs is necessary.

When exclusively viewing the existing longitudinal data, we identified a number of missing areas in memory development (see Fig. 4). Most notable, data on within-child changes are absent for statistical learning and the inference tasks from our database. Developmental changes in memory for what-where-when conjunctions also appears to be a knowledge gap. The combination of

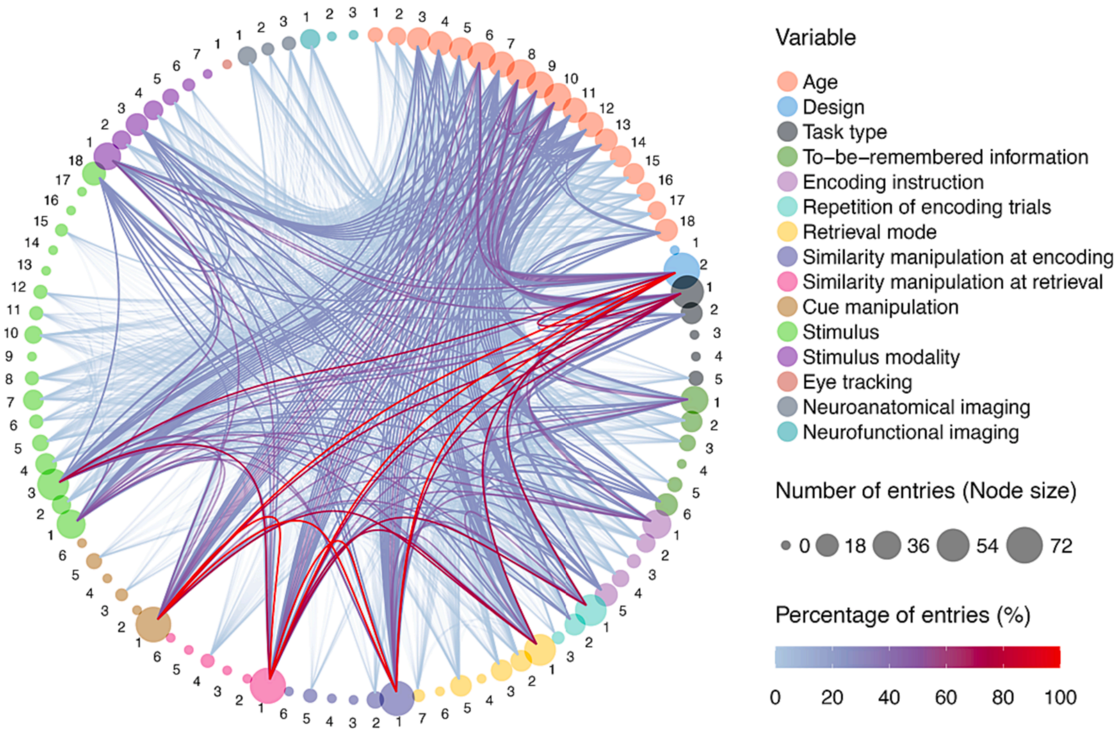

Fig. 4. An aggregated HEB of all task entries that used a longitudinal design. The number of entries in each category of a given variable is denoted by node size, with larger node sizes being greater number of entries. The color intensity of the edges denotes the frequency with which two categories co-occur in the same task entry.

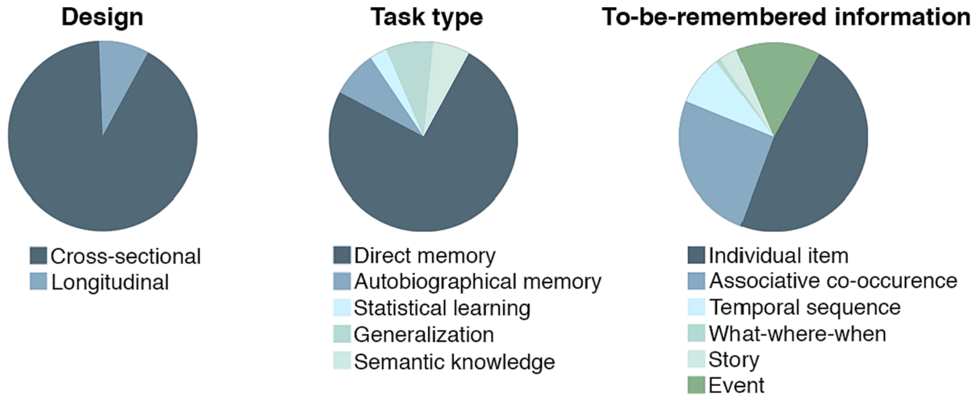

Fig. 5. Pie charts depicting the relative composition of the experimental design, task type, and to-be-remembered information variables.

longitudinal design and neuroimaging is extremely rare. Further, we lack longitudinal data on how children's memories are impacted by a cue manipulation or a similarity manipulation to test the development of pattern completion and separation processes.

Task type. Of the 840 task entries, 74.8 % (506 publications) employed direct memory tasks, 7.86 % (55 publications) employed personal memory interviews, 3 % (23 publications) employed statistical learning tasks, 8.10 % (63 publications) employed generalization tasks, and 6.31 % (42 publications) tested children's semantic knowledge (see Fig. 5). It is worth noting that there is likely an enormous body of research on statistical learning that intersects with language development. The relatively small number of task entries for statistical learning here is due to the fact that we restricted our search to "visual statistical learning" for feasibility purposes. In addition, the relatively lower number of semantic knowledge task type is likely due to the lower bound of our age criterion (age 3). We expect a larger literature on semantic knowledge development as vocabulary acquisition for a lower age range (Landau et al., 1998; Mareschal & Quinn, 2001).

To-be-remembered information. The breakdown of the to-be-remembered information is as follows: 47.7 % tested memory for individual items (278 publications), 25.5 % tested memory for associative co-occurrence (169 publications), 8.5 % tested temporal memory (58 publications), 0.8 % tested participants' memories for what-where-when associations (6 publications), 3.1 % tested children's memories of stories that were told to them (25 publications), and 14.4 % tested participants' event memories (101 publications) (see Fig. 5). It is heartening to note that slightly more than half of the task entries went beyond individual items, but the variety of methods in the studies that did so is also noteworthy.

Encoding instruction. The highest frequency of encoding instruction was intentional (39.4 %), followed by incidental encoding instruction (32.1 %). Only a few entries (2 %) employed both kinds of encoding instructions, limiting our knowledge of what difference intentionality makes for children. Approximately 15 % of papers did not specify whether participants were aware of an upcoming memory test, a notable omission.

Repetition of encoding trials. The number of encoding trial repetitions was coded as either a continuous variable or as a percentage for tasks that required a learning criterion (e.g., 80 %). For visualization and analytical purposes, we binned this variable into three categories: one-shot learning (items were shown once at encoding), multi-shot learning (items were shown more than once) and learning criterion (items were learned until a performance criterion is reached). The highest frequency of entries used one-shot learning, followed by multi-shot learning, followed by learning to a criterion. Whether this news is good or bad depends on the problem under study. For example, one-shot learning is best if the focus is on life events, whereas learning to a criterion may be essential for studies of forgetting rates.

Retrieval mode. Free recall, cued recall, old/new recognition, and alternative-forced-choice were the commonly employed retrieval modes, whereas remember/know responses, old/similar/new recognition, and re-enactment were less common (see Fig. 2A). Remember/know and old/similar/new judgments may be challenging for young children, whereas re-enactment may be odd for older participants.

Stimuli and modality. The most frequently used stimuli were objects and animals, and they were most often presented visually (see Fig. 2C).

Similarity manipulation. Most of the papers, 77 %, did not include a similarity manipulation. Among the 287 entries that were classified as having a similarity manipulation, 229 were implemented at encoding and 159 were implemented at retrieval. For both cases, the majority of these papers used conceptual similarity among to-be-remembered information at encoding (94 entries, 81 papers), or between the studied and tested items at retrieval (84 entries, 65 papers). This group of papers primarily employed the DRM paradigm (Deese, 1965; Roediger & McDermott, 1995) to examine how conceptual relatedness impacts the rate of false recognition for semantically related lures at test. We classified 33 entries from 26 papers and 35 entries from 29 papers that manipulated perceptual similarity at encoding and at retrieval, respectively. Seventeen entries and 34 entries were classified as manipulating both conceptual and perceptual similarity at encoding and retrieval, respectively. Seventy-three and 13 entries had an overlapping element at encoding and retrieval, respectively. Finally, 22 and 7 entries used repetition at encoding or at retrieval to manipulate similarity among studied items at encoding, or between studied and test items at retrieval. As we discuss below, a subgroup of these entries is pertinent for the examination of pattern separation development.

Cue manipulation. Approximately 94 (11 %) of the total task entries employed some form of a cue manipulation in their experimental designs. Nearly half of these entries used a binary manipulation, that is, a cue was either present or absent at test. Even fewer entries employed other forms of cue manipulations: 15 entries from 10 papers manipulated the degree of cue presence in a gradient manner, and 20 entries from 20 papers manipulated the type of cues to examine their relatively effectiveness in retrieval success.

Physiological measures. Physiological measures such as eye tracking and neuroimaging were rare. Eye tracking was only seen in 3 % of the task entries (15 papers). Further, we did not identify any papers on developmental changes in children's viewing behaviors; all 15 papers were cross-sectional (see Fig. 6A).

A slightly higher number of publications included a brain structural imaging component. A total of 52 task entries from 28 papers reported neuroanatomical indices (see Fig. 6B). Among these, 44 entries from 23 papers reported results from T1-weighted sequences. An additional 12 entries from 7 papers collected T2- or Proton Density-weighted sequences, which are suitable for hippocampal subfield segmentation protocols. Most notably, only 9 entries from 5 papers used diffusion-weighted imaging to report brain water diffusivity indices either in gray matter regions and or from reconstructed white matter pathways (for a detailed breakdown of each neuroanatomical imaging category, see Figure S3). Studies with a diffusion-weighted imaging component all used direct memory tests.

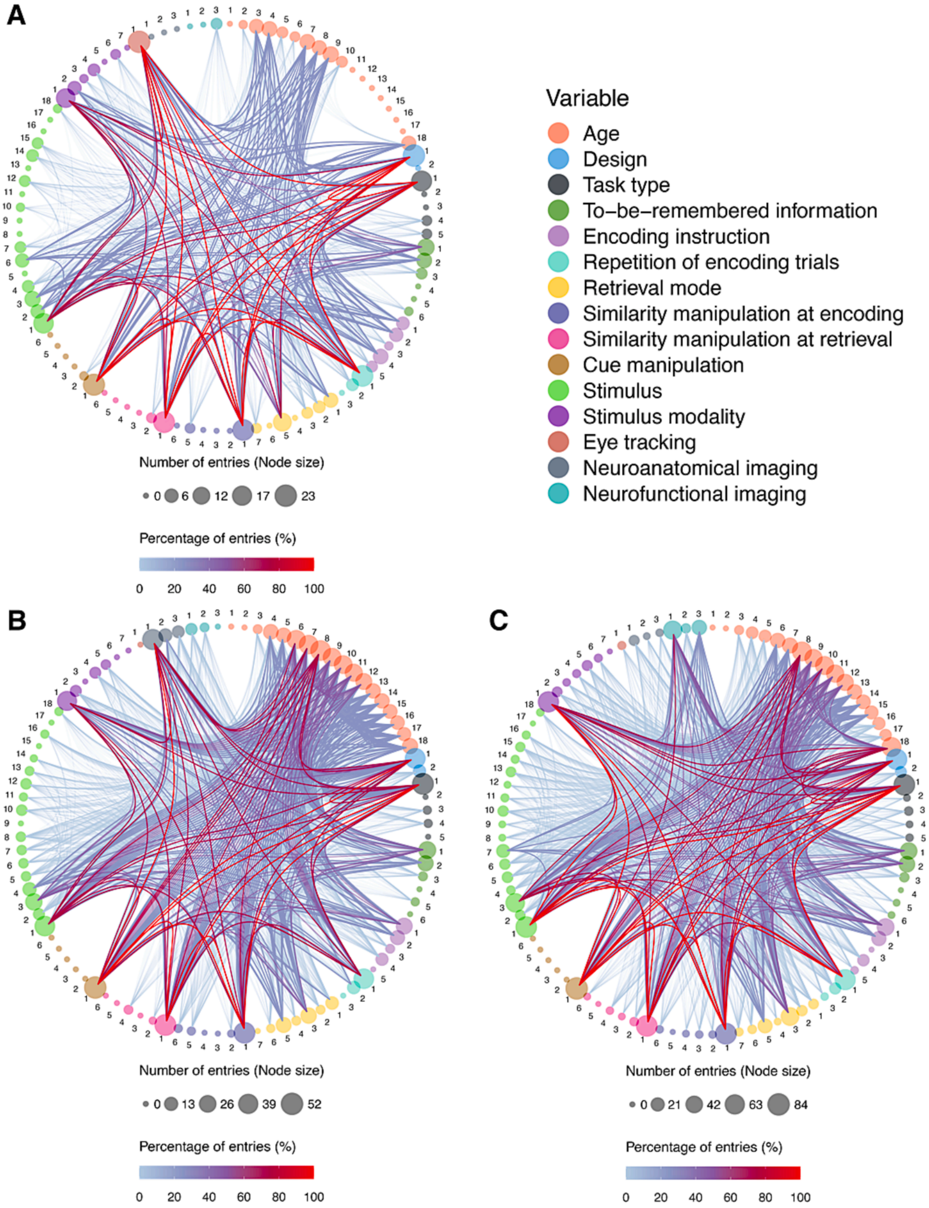

Fig. 6. Hierarchical edge bundling (HEB) plots that visualize the combination of methodological variables for task entries that included eye tracking (A), neuroanatomical measures (B), and neurofunctional measures (C). In each HEB, the proportion of task entries that employed a given category of a variable is denoted as node size, and the proportion of the combinations between two given variables is denoted as the color intensity of the edges.

These findings reveal a gap in our knowledge of how brain microstructural properties in gray matter and white matter connectivity may relate to other aspects of memory development including personal memories, statistical learning, generalization, and semantic knowledge.

Considering the technical difficulty in collecting functional brain data in developmental samples, there has been considerable effort in measuring task-relevant brain activation in children from 8 years of age (see Fig. 6C). We found 54 entries from 33 papers that measured task-related BOLD signals using functional magnetic resonance imaging (fMRI). There are six entries from 3 papers that collected task-free fMRI data to quantify the temporal coactivation patterns in the brain during off-task periods. The remaining 28 entries from 21 papers measured event-related potentials with electroencephalogram (EEG) (for a detailed breakdown of each neuroanatomical imaging category, see Figure S4).

Delay period. The delay value can be a range instead of a single value; therefore we visualized this variable separately (see Fig. 7). The most common delay period was immediate test for direct memory, statistical learning, generalization, and semantic knowledge tasks. The longer delay periods that span multiple years were mostly seen with personal memory tasks. Thus, we lack information on maintenance of the former and immediate memory for the latter.

Potential applications of the database

In addition to taking stock of the existing evidence in research on memory development, our database generates new avenues to integrate findings in psychology—both contemporary and long-past—into a comprehensive network of literature. Using specific examples, we highlight two applications that facilitate the consolidation of existing empirical evidence and promote theory building. The first application is quantifying the similarity between entries based on the multidimensional classification of their method variables. This utility of the database discovers bridges connecting related work that can, and should, coalesce to shed light on common phenomena relevant for understanding the development of human memory. The second application is to identify publications whose methods align with the operational definitions of specific constructs of interest. This approach relies on a filtering method with a combination of method variables to identify potentially relevant entries. Such utility allows mapping reviews to serve as the precursor and a powerful springboard for subsequent systematic or narrative reviews and *meta*-analyses that target outcomes related to a set of questions (Grant & Booth, 2009). Here, we exemplified the implementation to one specific research question in memory development: *What is the developmental profile of pattern separation in childhood*?

Application 1: Quantifying inter-task similarity

Developing a multidimensional classification of empirical work enables researchers to estimate publications' methodological overlap using data-driven approaches, thereby identifying related (but disconnected) work. One way of quantifying inter-task similarity is to measure the proximity of method factors between two given entries. To this end, we first calculated the similarity score between two entries for each variable separately, and then averaged across all variables to yield one similarity score for every inter-task comparison. Given that the nature of the extracted data inherently differs across variables (continuous vs. categorical, single-value vs. multi-value), our comparison of categorical data was carried out on a variable-by-variable basis. For 7 out of 16 variables, we simply applied a binary method to comparing categorical variables. A match and a mismatch were assigned as 1 and 0 respectively (see

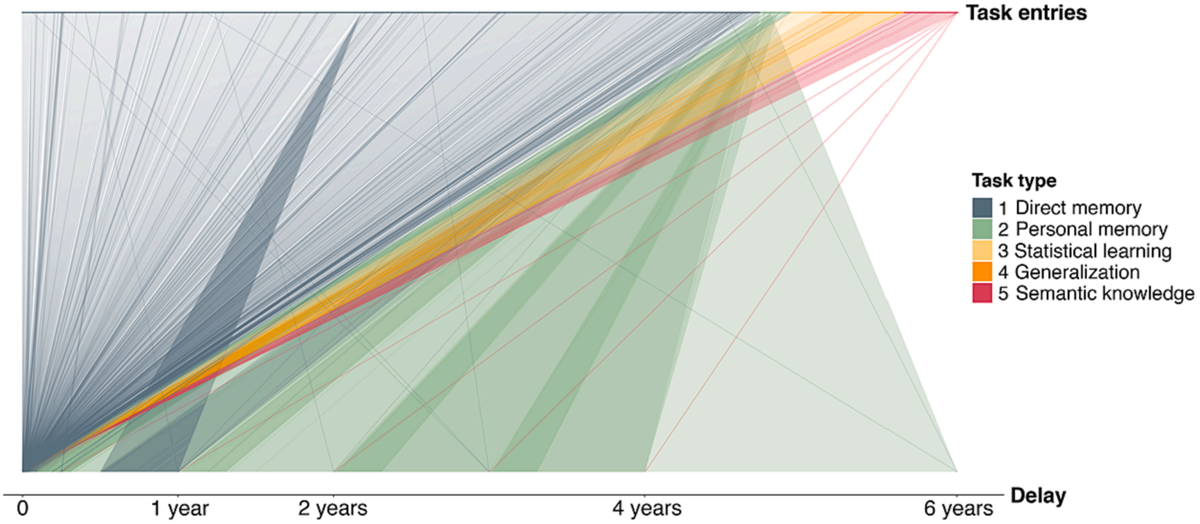

Fig. 7. Delay periods ranging from 0 to 6 years across the five task types and across all task entries. Every point on the top *horizontal*-axis denoting a given task entry is mapped onto its respective delay value(s) or range(s), shown on the bottom *horizontal*-axis. Entries with concrete delay values (e.g., one hour, two years) are shown as individual lines. Entries with delay values as ranges (e.g., 6- to 12-month delay) are shown as projected shades that cover the range of delay values on the bottom x-axis.

Fig. 6). For example, if entries A and B employed a cross-sectional and a longitudinal design, their similarity score on the design dimension would be 0. If both employed the design, their similarity score on this variable would equal 1. These 7 variables include: *design*, *task type*, *to-be-remembered information*, *encoding repetition, stimuli*, *retrieval mode, and eye-tracking.* For each of the remaining variables including age, encoding instruction, similarity manipulation, cue manipulation, delay period, neuroanatomical and functional imaging, we reasoned that are varying levels of similarity among the categories (see Table 1 and Fig. 8).

Next, we averaged the similarity score across 16 dimensions for each pairwise comparison, resulting in an 840 x 840 similarity matrix (see Fig. 9A). Visual inspection of the matrix shows that task entries of the same task type tend to employ overlapping methodological variables than those classified as different task types. This pattern is especially apparent for the personal memory entries, likely resulting from the fact that previous studies from this line of work mostly employ interviewing protocols. Our approach provided a quantitative estimation of the degree of overlap between any two task entries, ranging from very little to complete overlap (see Fig. 9B). Note that the calculation presented here is one among many possibilities of estimating inter-task similarity. In this instantiation, we specified that all 16 variables equally contributed to the similarity score and imposed assumptions about the degree of similarity between levels within a given variable. Alternatively, researchers could (i) apply all binary-based comparisons across all variables (see Figure S1), (ii) select a subset of the variables for the similarity analysis, or (iii) assign uneven weights across variables depending on the research questions and hypotheses.

Connecting publications. Our second goal was to connect publications based on methodological convergence, despite their remote theoretical inspirations and disjointed scholarship. Below we illustrate three examples of such convergence via the hierarchical edge bundling visualization tool, focusing on generalization (example 1), cue manipulations relevant to pattern completion (example 2), and on similarity manipulations relevant to pattern separation (example 3).

Convergence between cross-episode binding and relational binding flexibility. This example highlights a specific comparison of two entries from the generalization task type, which requires participants to extend memories to draw more general conclusions. One entry is from Schlichting and colleagues (2016), and the other is from Richmond and Pan (2013). Both studies employed an associative inference paradigm that measures the abilities to link separate episodes that shared a common element (see Fig. 10A).

Schlichting et al. (2016) examined the relation between hippocampal structure and associative inference from age 6 to age 18, motivated by the neurocognitive account of hippocampal engagement in cross-episode memory integration. Participants first learned a series of pseudo-object pairs after being instructed to create a visual or verbal story that link each pair. Importantly, every association

Table 1
Descriptions of our methods for comparing categorical variables for each method variable.

Variables	Inter-category similarity calculation
Age	Age was coded as a continuous variable, and thus, we normalized the similarity scale for age, ranging from a perfect match (e.g., age 3 versus age 3) to the furthest distance in our coding scheme (1 vs. 18 +). Given that age can have multiple values per task entry (e.g., a study that included children aged 3 and 4), we averaged across all nonduplicated pairwise comparisons for all age values between two given task entries to yield a single score of age similarity.
Encoding instruction	• This variable did not contribute to the inter-task similarity score for entries that were coded as “unspecified” (i.e., no extractable data).The rest of the values were treated in a binary fashion, with the exception that the entries coded as “both intentional and incidental” shared an intermediate similarity score (0.5) with those coded as either intentional or incidental encoding instruction.
Similarity manipulation	• We reasoned that task entries that had any form of similarity manipulations were more similar to each other than they are to those without a similarity manipulation. Thus, all 5 types of similarity manipulation shared an intermediate similarity score (0.5) to one another. Both the conceptual and the perceptual similarity manipulation is closer to the conceptual + perceptual manipulation than each one is to other types of manipulation. Therefore, these cases were assigned with a higher similarity score (0.75) compared to the rest.
Cue manipulation	• We treated task entries with any form of cue manipulation to be more similar to each other than they are to those without a cue manipulation. Thus, all 5 types of cue manipulations shared an intermediate similarity score (0.5) with one another.Two other assumptions were made. First, there were two types of binary cue manipulations: binary and binary cue provision contingent on recall success. We reasoned that these manipulations would be more similar to each other than they are to other types of cue manipulations. Second, there were two types of cue element manipulations: cue element and element order, both of which involved a manipulation of the elements from the same past experience. We again reasoned that these manipulations would be more similar to each other than they are to other types of cue manipulations. Therefore, they were assigned a higher similarity score (0.75) .
Delay period	Delay was coded as a continuous variable in the unit of minutes. Given a wide range in how delay periods were manipulated and reported, we allowed for single values (e.g., 10), multiple values per task entry, and ranges (e.g., 10080–14400). The delay periods from the selected papers ranged from 0 to 6 years. We first binned the delay periods into 13 levels ranging from immediate to 6 years, and then normalized the similarity scale across the 13 bins, ranging from a perfect match (e.g., 0 versus 0) to the furthest distance (0 vs. 2–6 years) (see Fig. 7). Admitted, our binning method was arbitrary and oversimplified the complexity of this method factor. Future investigations should consider theory-driven treatment for the delay variable when calculating inter-task similarity.
Neuroanatomical and functional imaging	• We considered entries that included any anatomical imaging data to be more similar to each other than to those without an anatomical imaging component. Thus, all 3 types of neuroanatomical imaging methods shared an intermediate similarity score (0.5) to one another. We applied the same logic to the brain functional imaging variable.To prevent the common absence of this method from erroneously inflating the similarity score, the physiological measure variables did not contribute to the inter-task similarity analysis for entries without such components.

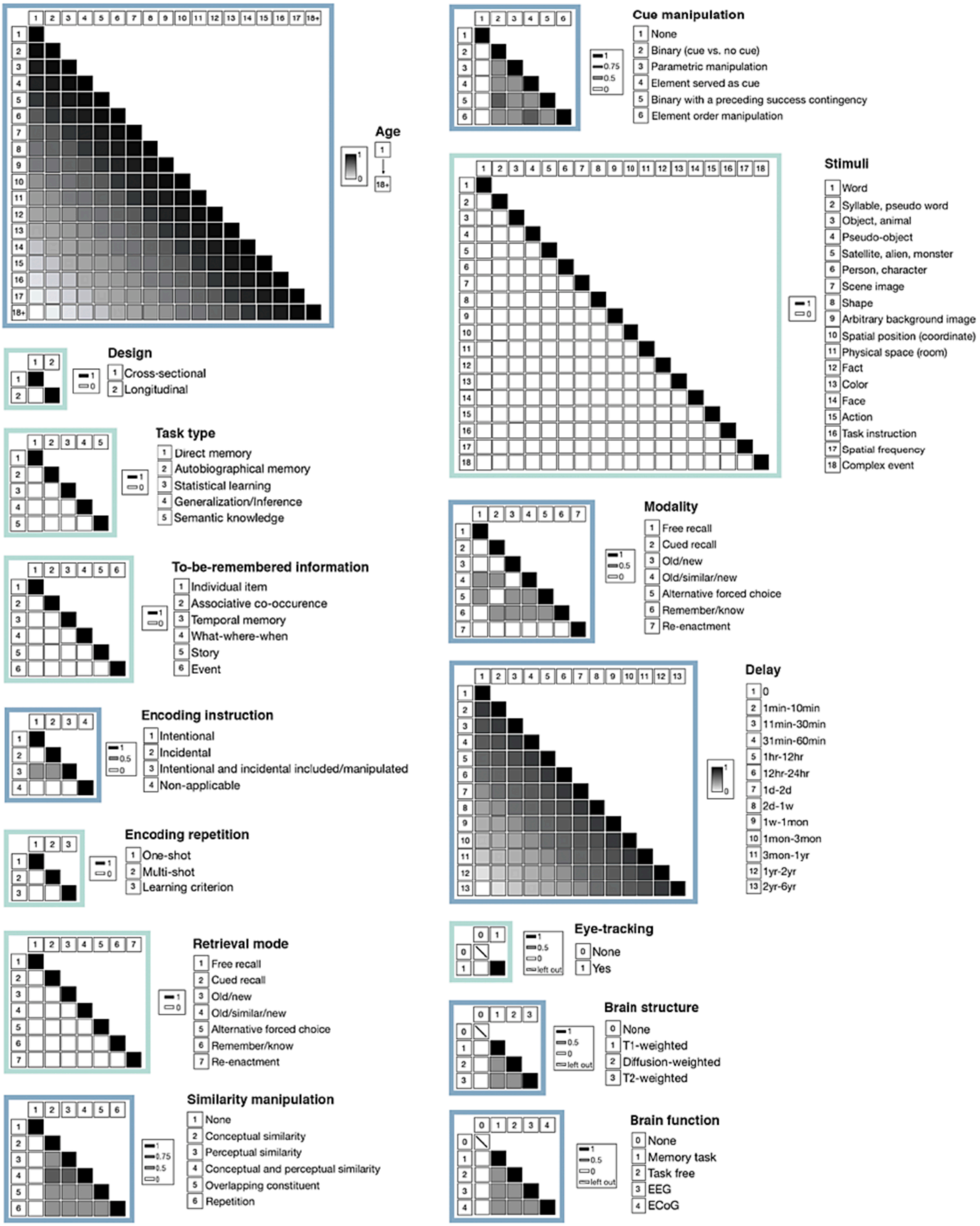

Fig. 8. A schematic depiction of the similarity scheme of the inter-task similarity analysis. The green and blue boxes denote binary and nonbinary variables, respectively. The gray scales denote the similarity value assigned to a specific pairwise comparison within each matrix. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

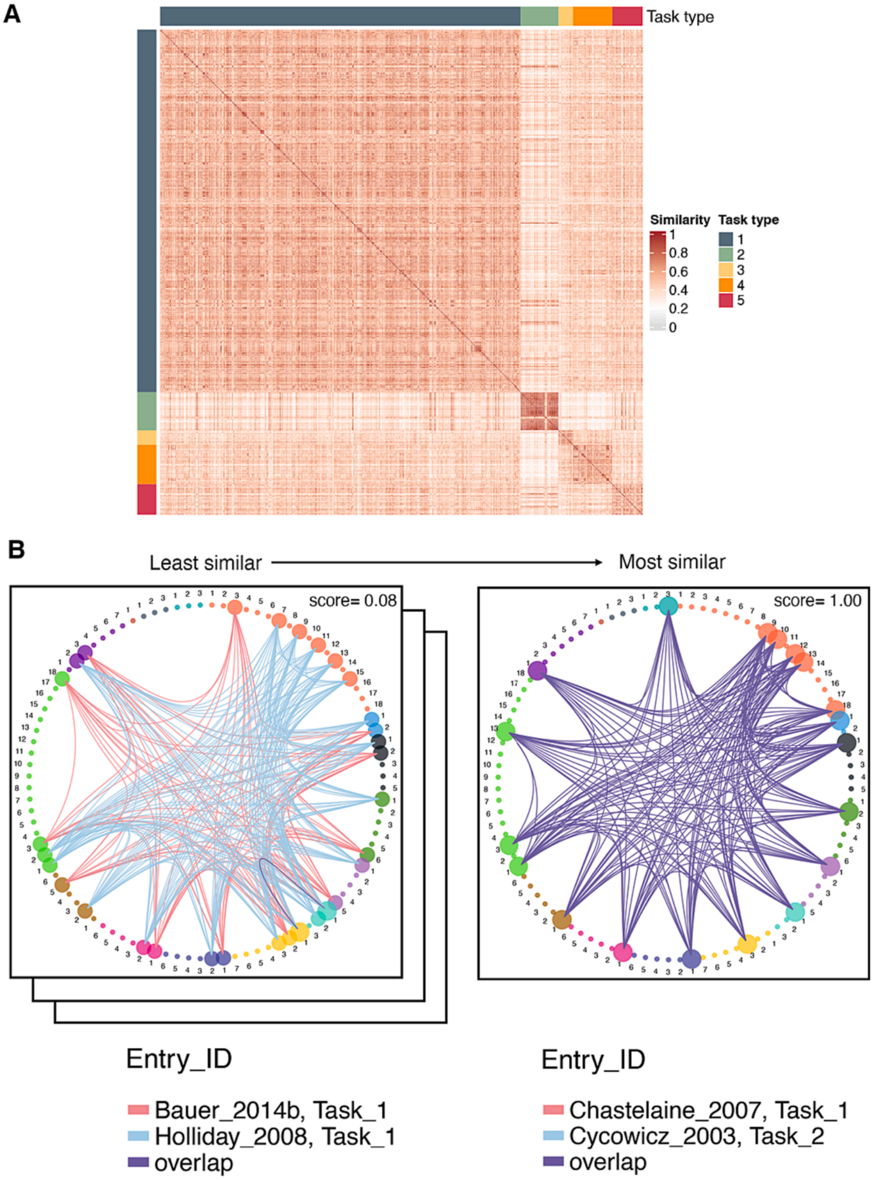

Fig. 9. A matrix of all pairwise similarity scores among the 840 task entries (A). Task entries are color coded by their classified task types along the rows and columns. Every cell represents an inter-entry similarity score, with darker colors denoting higher scores. Visual inspection of the matrix suggests that entries of the same task type tend to share greater methodological similarity with one another compared to those across different task types. This approach yields a gross estimation of similarity between two given task entries ranging from little to complete methodological overlap (B).

shares a pairmate with another association on the study list. After three consecutive encoding-test cycles of these direct associations, participants were given a surprise associative inference task that required the linking of two items, though never experienced together, on the basis of a shared pairmate (see Fig. 10A, top panel). The authors tested the prediction that hippocampal gray matter volume in a developmental sample would relate to their performances on associative inference.

On a seemingly different topic, the Richmond and Pan (2013) examined the association between relational memory and episodic future thinking from age 3 to age 5. Inspired by the “constructive episodic simulation hypothesis” (Atance & Neill, 2001), the authors tested whether children’s episodic future thinking is related to their ability to flexibly recombine elements from multiple past experiences. In this study, children were shown animal-place associations in story books. Unlike Schlichting and colleagues (2016), children were given a cover story that the place was the given animal’s favorite place. However, similar to Schlichting and colleagues’ study (2016), the associations overlapped such that two animals were paired with the same favorite place. After being tested their memories for the animal-place pairs, children were asked to infer a novel association between two animals who shared a favorite place (see Fig. 10A, bottom panel).

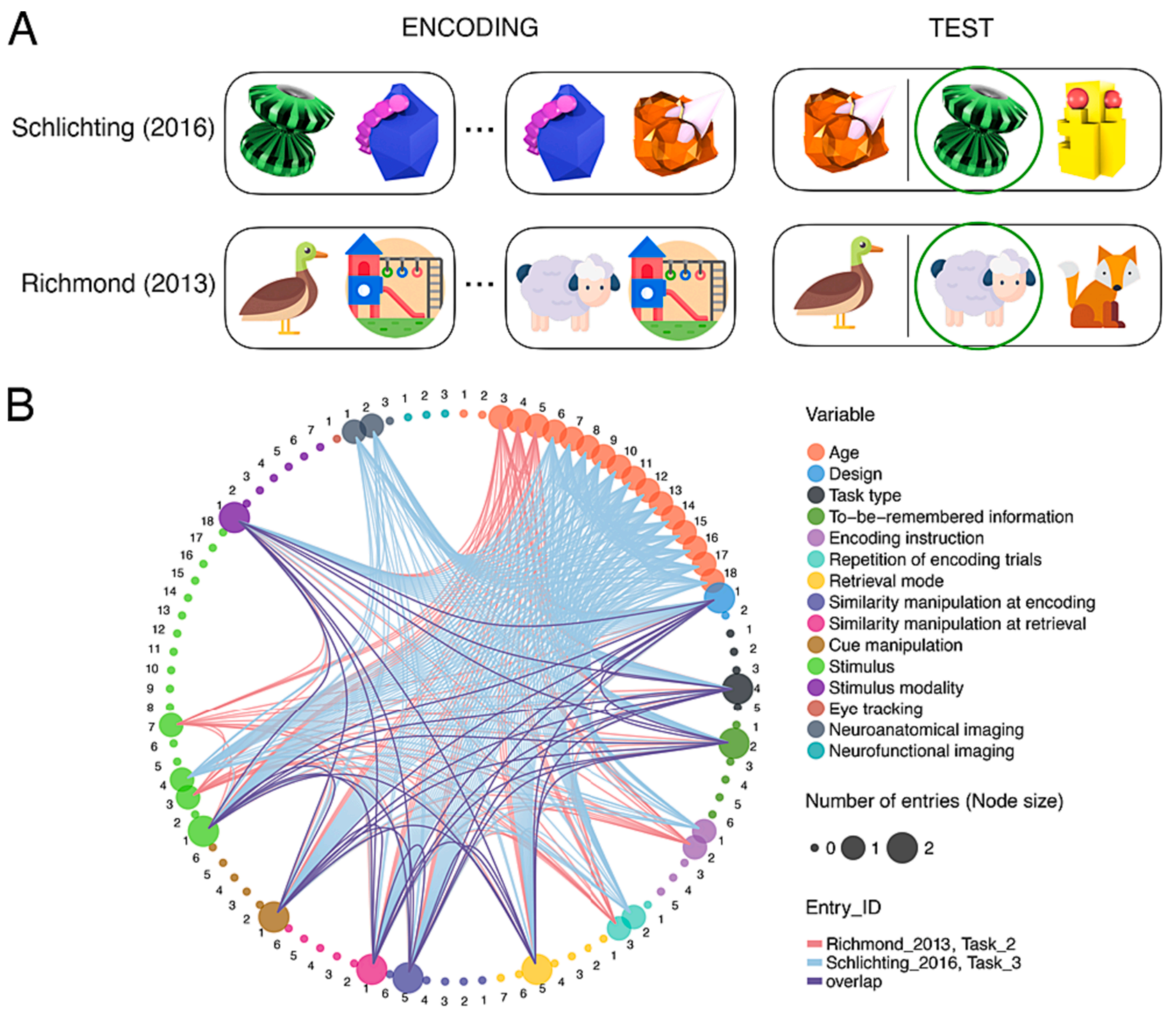

Fig. 10. An illustration of the experimental design for each of the two entries classified as generalization tasks (A) and their methodological convergence (B). In one entry (Schlichting et al., 2016) (A, top panel), participants' task was to infer that two items (e.g., the green and orange objects), though never experienced together, would be linked if they had shared a common pairmate. In another entry (Richmond & Pan, 2013) (A, bottom panel), after learning and being testing on a series of animal-place associations, children's task was to infer a novel association between two animals (e.g., the duck and the sheep) who had the same favorite place (e.g., the playground). The green circles indicate the correct answers. The overlay of the two entries in a hierarchical edge bundling demonstrates the shared method variables (purple connections) and those unique to only one entry (shown in their respective colored connections). Icons are from https://www.flaticon.com. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Our inter-task similarity metric assigns a score of 0.69, on a scale from 0 to 1. The overlay of their hierarchical edge bundling shows that they share the same task type, to-be-remembered information, and similarity manipulation. They differ in the age of the children, encoding instruction, stimuli, and whether they have neuroanatomical imaging variables (see Fig. 10B). Taken together, these studies show that associative inference performance is above chance level by age 5, although not at age 3 (Richmond & Pan, 2013), and that it continues to improve across the elementary school years (ages 6 to 12), thus providing a longer developmental trajectory than either study taken alone. One potential next step would be to use a common task, perhaps with the child-friendly stimuli used in Richmond & Pan (2013), on a wide age range that spans early to late childhood. A second kind of study would be to test the prediction that hippocampal maturation underpins the development of inferences about new (but warranted) associations based on past experiences, thereby enabling children to flexibly imagine possible futures.

Convergence between memory trace composition and pattern completion. In an example that relates to the construct of pattern completion, we compared two entries that employed the same kind of cue manipulation: one from Ceci and colleagues (1980) and another from Ngo and colleagues (2019) (see Fig. 11). By alternating which elements of a given experience serve as cues during retrieval, these two studies tested whether participants remembered aspects of a past experience as an integrated unit. The motivation and framing are different, however, with one targeting memory trace composition (Ceci et al., 1980), whereas the other targeting pattern completion (Ngo, Horner, et al., 2019).

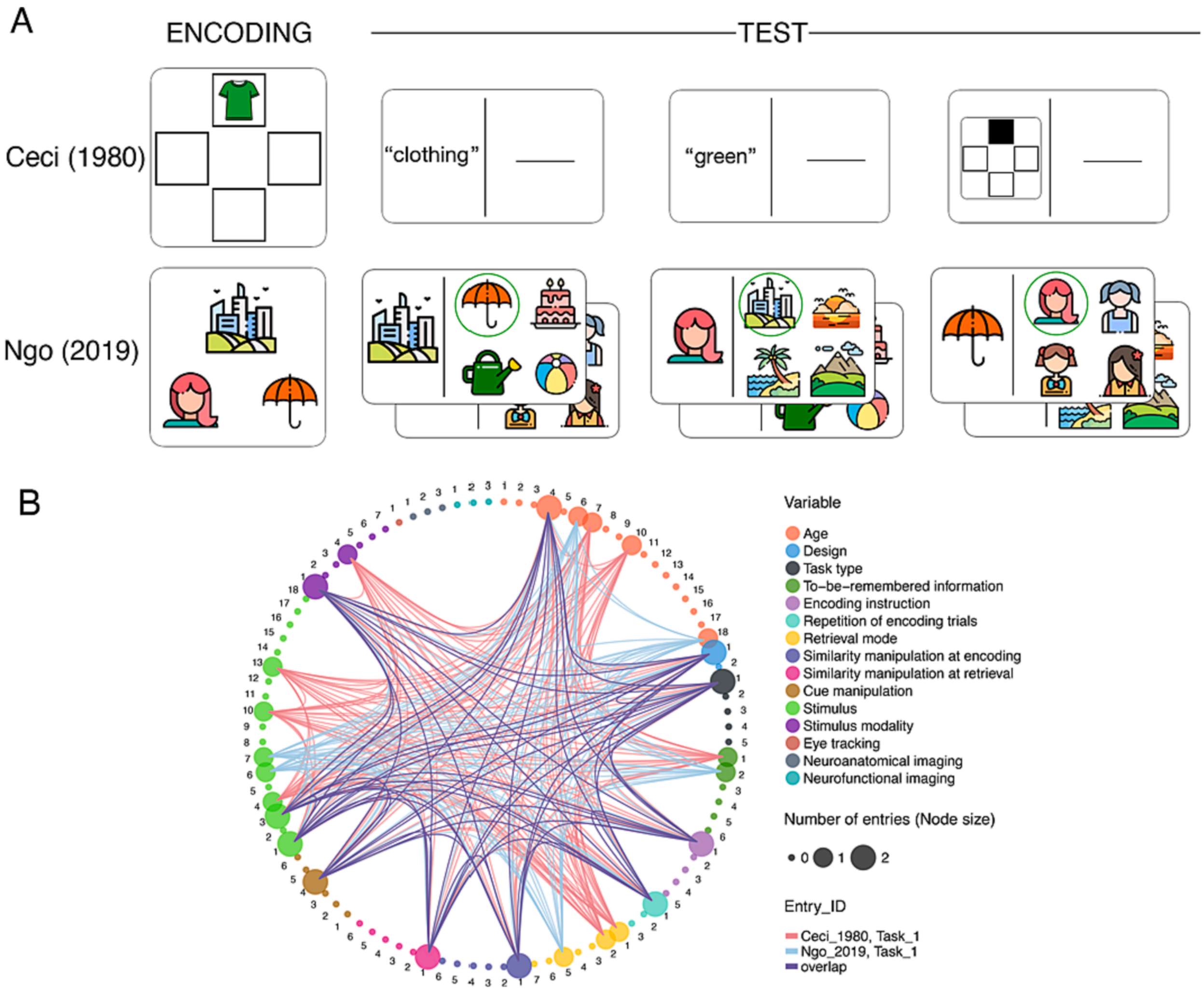

Fig. 11. An illustration of the experimental design for each of the two entries with a cue manipulation (A) and their methodological convergence (B). In one study (Ceci et al., 1980) (A, top panel), children first learned objects in different semantic categories that were presented in different colors and positions (a green shirt located in the top square). At test, they performed a cued recall task for the objects, with each attribute served as the cue. In another study (Ngo, Horner, et al., 2019) (A, bottom panel), participants learned a series of person-place-object episodes, and were tested on all bidirectional associations from each episode. The green circles indicate the correct answers. The overlay of the two entries in a hierarchical edge bundling demonstrates the shared method variables (purple connections) and those unique to only one entry (shown in their respective colored connections). Icons are from https://www.flaticon.com. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Ceci and colleagues (1980) tested the hypothesis of a developmental shift in memory retention from perceptual representation toward semantic representation of items in memory (Underwood, 1969). Children, aged 4, 7, and 10, first learned objects in different semantic categories that were presented in different colors and positions (e.g., a green shirt located in the top square, see Fig. 11A, top panel). At test, they performed a cued recall task for the objects, with each attribute serving as the cue (e.g., semantic category: clothing; color: green; or position: top square). The authors quantified the relative composition of an object memory trace by measuring the retrieval success when its different attributes were cued. They also measured whether retrieval dependency between different attributes of the same object.

Ngo and colleagues (Ngo, Horner, et al., 2019) investigated the development of holistic episodic retrieval, an index of pattern completion. Motivated by previous findings in young adults (Horner & Burgess, 2014), the authors tested the idea that if children's episodic memories are stored as an integrated representation, such that remembering one constituent of an episode should elicit the retrieval of its other elements. Four-year-olds, 6-year-old children, and young adults were shown a series of "stories", each made up of a place, a person, and an object. They were then given a cued associative recognition task in which each element took turn serving as the cue (see Fig. 11A, bottom panel). Participants were instructed to choose one among four options that belong to same story as the cue element. Similar to Ceci and colleagues (1980), the authors estimated the retrieval dependency between different associations of the same episode.

The similarity score of these two studies is 0.80 on a scale of 0 to 1. Despite the different theoretical motivations, both studies asked whether the probability of retrieving one element is independent of the probability of retrieving other elements from the same experience. The visualization from our mapping review show that the studies overlap in several important experimental variables including the task type and cue manipulation. They differ in the to-be-remembered information, stimuli, and retrieval format (see Fig. 11B). Findings from both studies show that as early as age 4, children already retrieve memories for multi-attribute items (Ceci et al., 1980) and multi-element episodes (Ngo, Horner, et al., 2019) in a holistic fashion. That is, even preschoolers correctly recollect all or none of the elements encountered in the same experience, not simply some subset, for both individual objects and complex episodes. However, these findings are somewhat surprising because previous findings showed that preschoolers' autobiographical memory recall seems fragmentary, with only bits and pieces of a past event successfully retrieved (Peterson, 2002). Future work should incorporate cue manipulation to test age-related differences in memory coherence across different types of memoranda: multi-attribute items, multi-element episodes, and autobiographical events.

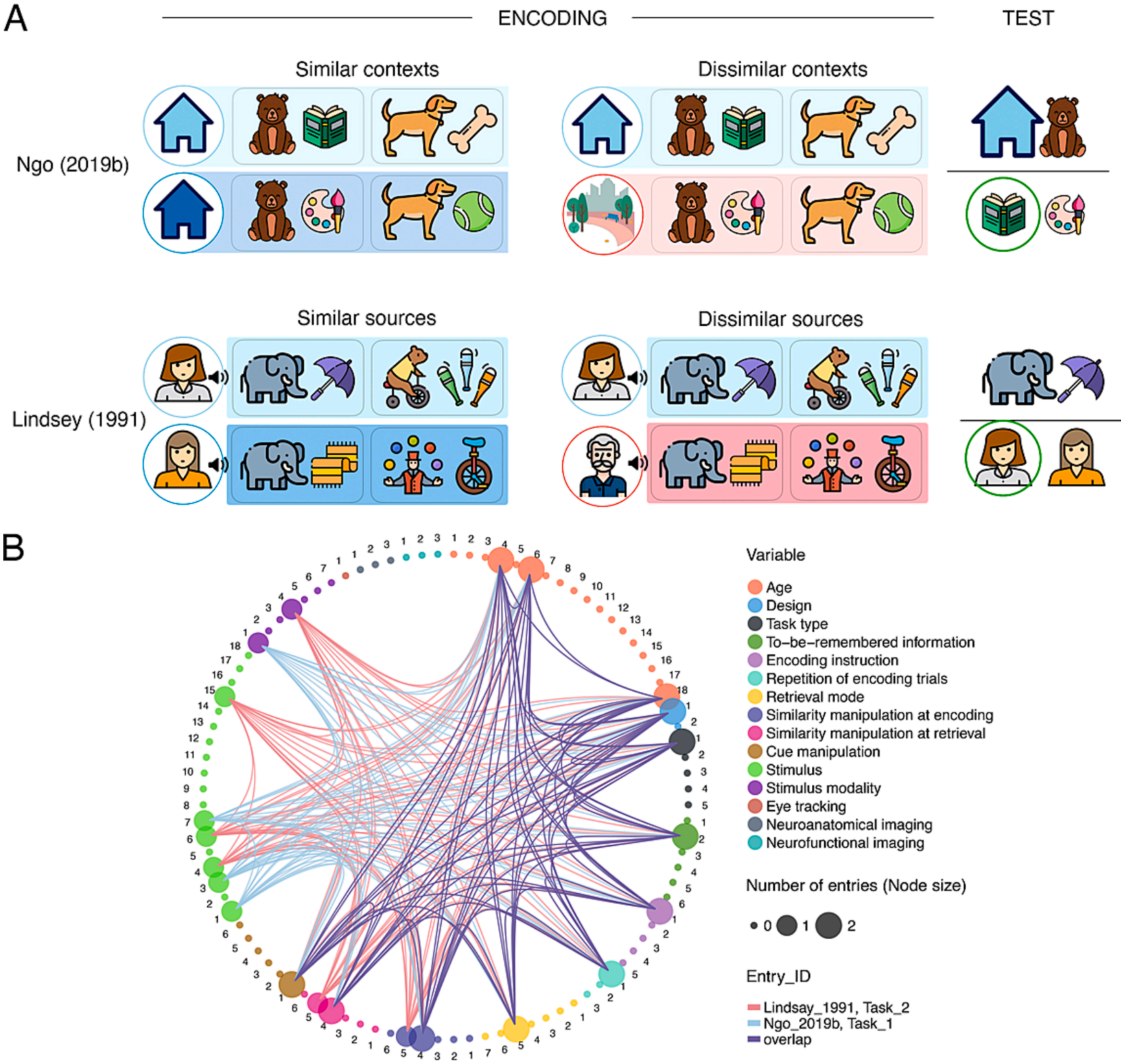

Fig. 12. An illustration of the experimental design for each of the two entries with a similarity manipulation (A) and their methodological convergence (B). In one entry (Ngo, Lin, et al., 2019) (A, top panel), participants learned a series of associations in either two similar (two blue houses), or two dissimilar contexts (a house versus a park). In another study (Lindsay et al., 1991) (A, bottom panel), participant were told two stories from two storytellers who were either similar (two female teenagers) to dissimilar (a female teenager and a male older adult) from each other. Some parts of the stories overlapped (the elephant), whereas others were unique to only one story (the purple umbrella, the blanket). The green circles indicate the correct answers. The overlay of the two entries in a hierarchical edge bundling demonstrates the shared method variables (purple connections) and those unique to only one entry (shown in their respective colored connections). Icons are from https://www.flaticon.com. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Convergence between pattern separation and source memory. In one example that pertains to the construct of pattern separation, we compared a study by Ngo and colleagues (Ngo, Lin, et al., 2019) versus one by Lindsay and colleagues (1991). Both studies were classified as implementing a similarity manipulation at encoding and retrieval. Inspired by the CLS theory, Ngo and colleagues (2019) examined the impact of context similarity on memory discrimination of overlapping associations in 4-year-olds, 6-year-olds, and young adults. Most prior research on pattern separation development has used similar object exemplars, which shows that by age 6, catastrophic interference is diminished. Here, the authors aimed to test whether the age-related differences in memory discrimination differs when pattern separation demand was placed on contexts versus objects. Participants learned a series of associations in either two similar (two blue houses), or two dissimilar contexts (a house versus a park) (see Fig. 12A, top panel).

Lindsay and colleagues (1991) charted the age patterns in source monitoring accuracy. Previous findings on children's reality judgments showed that by age 6, children can reliably discriminate between imagined versus actual events (e.g., Foley & Johnson, 1983). Lindsay and colleagues (1991) asked whether preschoolers had specific difficulty with reality monitoring or whether they had a general difficulty with source discrimination for similar sources. To test these hypotheses, they examined the impact of source similarity on source monitoring for overlapping as well as nonoverlapping associations in 4- and 6-year-old children and young adults (see Fig. 12A). Participants were told two stories from two storytellers who were either similar (two female teenagers) to dissimilar (a female teenager and a male older adult) from each other in their appearances. Some parts of the stories overlapped (the elephant), whereas others were unique to only one story (the purple umbrella, the blanket) (Fig. 12A, bottom panel).

The inter-task similarity score is 0.81. The visualization from our mapping review shows the differences between the two studies are the stimuli and modality, but there is overlap for important variables including age, design, task type, to-be-remembered information, and similarity manipulations (see Fig. 12B). Findings from the two studies agree in showing that the ability to discern overlapping episodes encountered in similar contexts or sources is much below the adult level at age 6. This developmental trajectory differs from both object memory discrimination (Ngo et al., 2017) and reality monitoring (Lindsay et al., 1991) in the length of time to reach adult-like levels. The convergence of the findings underscores the importance of characterizing how pattern separation interacts with memory representations in development.

As these examples show, a mapping review unveils overlap in methods, despite differences in theoretical frameworks, construct labels, surface-level differences in experimental materials, and importantly, reference to each other. In order to facilitate the process of literature integration and identification of methodological overlap within the existing publications, we have created an interactive web application (https://memdev.shinyapps.io/litreview_io/) that provides full access to the database and enables users to select and visualize the methodological variables from any individual task entry or an overlay of multiple task entries with HEB. Importantly, our application enables users to identify related publications (e.g., top 10 most similar entries) based on the similarity scores for a given seed entry. Unique to our methods-focused database, this data-driven search feature introduces a powerful tool for literature search that is currently not available with the other databases.

Application 2: Re-evaluating the development of pattern separation as an example case

We retroactively identified and characterized prior work that employed experimental manipulations germane to pattern separation. One key finding of the mapping review was that, although the terms pattern separation and completion only emerged in the memory development literature within the last decade, such manipulations have been employed from the 1970 s, but under different labels. We applied a filtering method to identify potentially relevant entries for pattern separation with a combination of method variables. First, they are entries that measure memories for a specific experience (task types 1 and 2: direct memory and personal memory), and contain either (i) any type of similarity manipulation at retrieval because this would induce interference, or (ii) any types of similarity manipulation at encoding if intrusions are measured. This filtering method resulted in a much larger set of papers, totaling 131 entries from 101 papers. Below, we summarized the key findings from the small number of pattern separation studies. Following that, we discuss findings from related topics based on our analysis to demonstrate that this approach adds valuable information on the role of pattern separation demands on memory discrimination during development.

Current understanding of pattern separation development in childhood. According to neurocomputational models of learning and memory, pattern separation refers to the process of reducing the representational similarity between overlapping inputs via the dentate gyrus' sparse coding scheme, thereby enabling separate storages of memories albeit high similarity (Norman & O'Reilly, 2003). This process relies on the dentate gyrus and CA3 subfields of the hippocampus—regions that are late-developing compared to other hippocampal subfields (Lavenex & Banta Lavenex, 2013). These observations invite the hypothesis that pattern separation follows a prolonged development in childhood, and that such development accounts for an improvement in memory specificity (Canada et al., 2018; Keresztes et al., 2017; Ngo et al., 2017). By our count, the number of studies that examined "pattern separation development" in childhood, as explicitly indicated by the authors either in the titles, abstracts, and/or introductions, totaled 15 task entries from 12 papers (Bouyeure et al., 2021; Canada et al., 2019; Hassevoort et al., 2020; Keresztes et al., 2017, 2020; Lambert et al., 2015; Ngo et al., 2018, 2019; Ngo, Lin, et al., 2019; Ngo et al., 2021; Rollins & Cloude, 2018; Sommer et al., 2021). Aligned with the operational definition by previous studies conducted with young (and sometimes older) adults (Bennett et al., 2015; Kirwan & Stark, 2007; Lacy et al., 2011; Stark & Stark, 2017; Stark et al., 2013; Toner et al., 2012), the majority of this group of developmental papers also estimated pattern separation using lure discrimination between object exemplars (e.g., perceptually similar chairs; Mnemonic Similarity Task, reviewed in Yassa & Stark, 2011). The most consistent age pattern from these papers shows that compared to school-aged children (aged 6 and above), preschoolers have a higher tendency to erroneously endorse similar test items as studied items – a hallmark of pattern separation failure (Ngo et al., 2018, 2019; Ngo, Lin, et al., 2019). When age is treated continuously, lure discrimination and age is associated linearly from early childhood (ages 3–4) to late childhood (ages 8–10) (Bouyeure et al., 2021; Canada et al., 2018; Ngo et al., 2021), or quadratically from age 6 to young adulthood (Keresztes et al., 2017). The age difference in lure discrimination between middle childhood (ages 6–7) and young adulthood is fainter, with most studies reporting a nonsignificant difference (Hassevoort et al., 2020; Ngo, Lin, et al., 2019; Ngo et al., 2021, 2017; Sommer et al., 2021), except for one study (Rollins & Cloude, 2018). Taken together, these findings show pronounced age-related improvements in pattern separation from early to middle childhood and continued strengthening into adolescence.

The evidence on pattern separation development thus far heavily draws on data from children's discrimination of object exemplars,

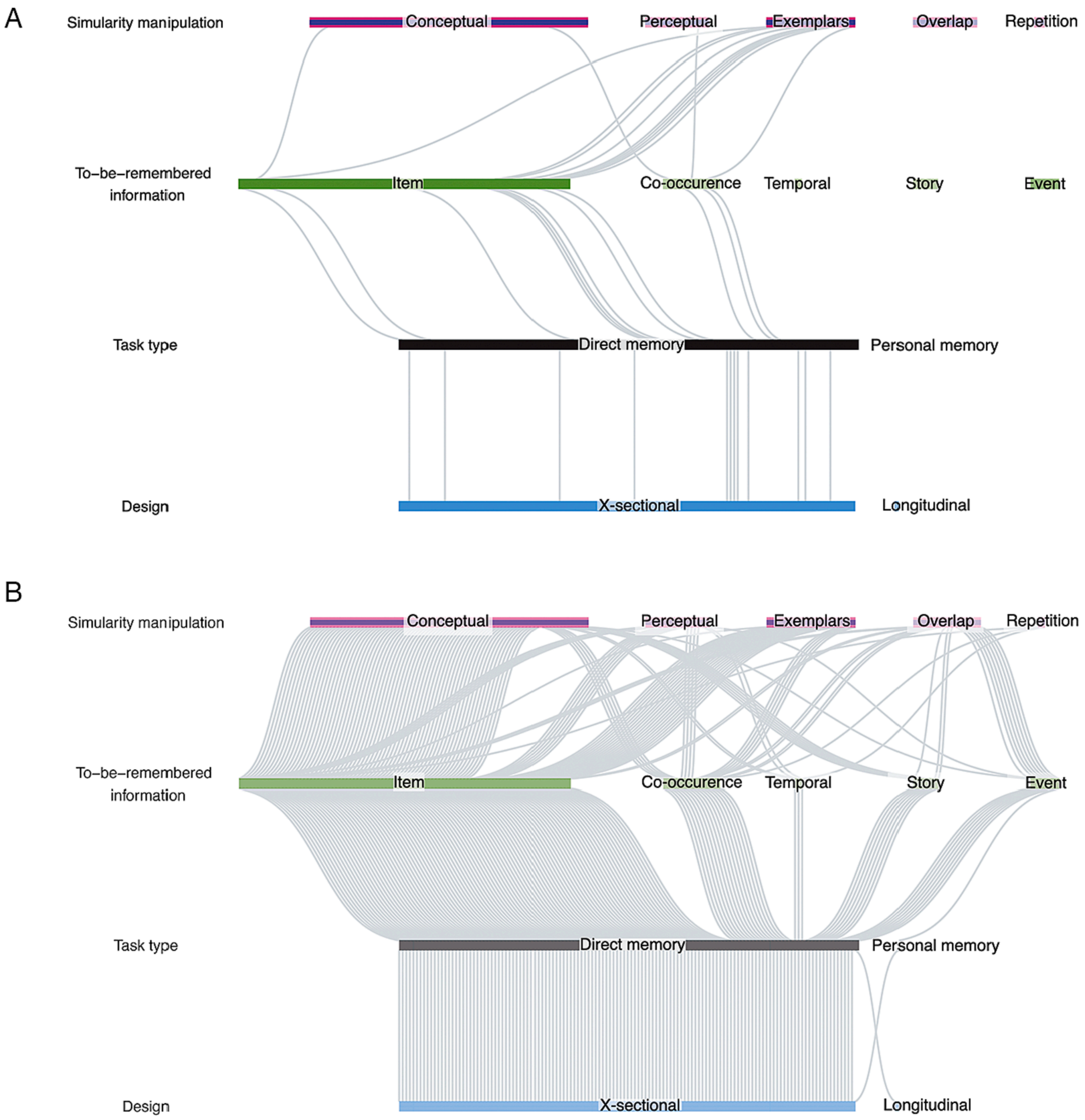

Fig. 13. Parallel-cord plots depicting the methods of the tasks explicitly identified as relevant for "pattern separation" in the publications (A) versus those using the method variables relevant to pattern separation in our database (B). For visualization purposes, we selected 4 variables: design, task type, to-be-remembered information, and similarity manipulation (shown as four levels in each parallel-cord plot). Each line denotes a single task entry, and is organized into its respective design, task type, to-be-remembered information, and similarity manipulation (collapsed between encoding and retrieval). Panel A shows that, up to the day of writing this review, all pattern separation papers are based on cross-sectional studies, which all tested memory retrieval of experimental stimuli. The majority of them tested individual items, with a few exceptions that examined memory for co-occurrences. Most of these papers used exemplars to induced similarity to test pattern separation. Panel B shows a much wider range in experimental design, including two longitudinal studies, and a fuller coverage of different to-be-remembered information. The kind of similarity manipulation also spans a broad range, including conceptual, perceptual, exemplars, overlapping, and repetition variations.

except for three cases. First, one study showed robust development of *allocentric spatial memory* resolution in children between 1.5 and 4 years of age (Lambert et al., 2015). Second, one study showed that children's lure discrimination for *semantically* related objects (without perceptual overlap) significantly improved from 3 to 8 years of age (Ngo et al., 2021). Third, one study examined children's memory discrimination for associations that occurred in similar versus dissimilar contexts. Here, the authors found that memory discrimination for associations learned in similar contexts appear to be equally challenging for 4- and 6-year-old children, and that adults far outperformed both groups of children (Ngo, Lin, et al., 2019). In contrast to studies with object exemplars which showed fainter age differences in pattern separation after age 6, this study tells a different story for memories of associations in similar contexts. These results prompt the consideration that age trends in pattern separation depend on the *kinds of memoranda* (e.g., individual objects versus associations in contexts). However, limited conclusions and hence literature integration, can be drawn from single-study findings, given that they each sit in isolation within the space of pattern separation development research.

Insights into pattern separation development from related topics. The key ingredient in the behavioral tasks that approximate pattern separation is the *similarity manipulation* between studied and test items. However, examining the impact of similarity-inducing interference on children's memory is not unique to pattern separation studies. In fact, our analyses revealed a total of 131 task entries (from 101 papers) that have employed a similarity manipulation either at the encoding or retrieval phase. Besides the 11 studies included in Fig. 13A, the remaining studies did not address, or even mention, pattern separation in the first place, but rather studied children's source monitoring (Day et al., 1998; Foley et al., 1994; Foley et al., 1993; Foley et al., 1987; Foley & Johnson, 1985a; Foley & Ratner, 1998; Johnson et al., 1979; Lindsay et al., 1991; Markham, 1991; Roberts & Blades, 1995, 1998; Roberts et al., 2016), event memory (Hudson, 1990b; Hudson et al., 1992; Hudson & Nelson, 1983; Mccartney & Nelson, 1981; Powell et al., 1999), verbatim versus gist memories (Anastasi & Rhodes, 2008; Brainerd et al., 2006; Brainerd, Reyna, & Forrest, 2002; Brainerd et al., 2018; Brainerd & Reyna, 2001; Carneiro et al., 2007; Dewhurst & Robinson, 2004; Ghetti et al., 2002; Holliday & Weekes, 2006; Howe, 2008, 2006; Howe et al., 2009), or iconic memory representation (Brown & Scott, 1971). Nevertheless, the task characteristics show tremendous overlap with typical pattern separation studies, as revealed by our analysis. Coalescing publications that contained a similarity manipulation enriches our understanding of pattern separation development (see Fig. 13B).

The role of event similarity in pattern separation development. Studies within the source monitoring framework echo the current understanding in pattern separation development based on findings that source *similarity* matters. Children's memory confusion is amplified when the sources closely resemble each other (Day et al., 1998; Lindsay et al., 1991; Markham, 1991). Three examples highlight that age-related differences in source discrimination increase as the degree of event overlap increases. First, 4-year-old children had more difficulty remembering the source of spoken words if the two speakers were of the same sex compared to when the speakers were of different sexes. In contrast, young adults' source discrimination was unaffected by source similarity (Lindsay et al., 1991). In the second example, 4-, 6-year-olds, and young adults watched videotapes of two storytellers, each narrating a story (Lindsay et al., 1991). There were two different similarity manipulations: (i) source similarity: same-sex versus different-sex storytellers, and (ii) story content overlap: some parts of the stories were shared, whereas others were unique to each story. Both source similarity and content overlap worsened memory discrimination for all age groups, but they disproportionately affected the younger children. When the similarity level was low, with dissimilar sources and nonoverlapping story contents, four-year-old children performed above chance and on-par with the 6-year-olds. When the similarity level was intermediate, such that either the sources were similar or the story content overlapped, 4-year-olds performed at chance and were less accurate than 6-year-olds. When similarity level was high, such that the sources were similar and the story content overlap, even the six-year-olds performed at chance. The third example comes from studies on children's reality monitoring and shows the same pattern. It was previously thought that by age 6, children reached adult-level proficiency in source discriminating between actual and imagined events (Foley & Johnson, 1985; Johnson et al., 1979). However, age differences persisted beyond age 6 if the events were carried out by the same person compared to when they were carried out by different people. In sum, age differences in source misattribution depend on the degree of similarity between the event contents and their contexts.

These findings resonate with a few studies that examined the impact of phonological similarity (Jarrold et al., 2015) and semantic similarity (Koppenol-Gonzalez et al., 2014) in children's *sequence memory*. Sequence memory of both phonologically similar items and semantically related items improves from age 5 to 12 (Koppenol-Gonzalez et al., 2014). However, there was an interactive effect between age and phonological similarity (e.g., horn, corn) on children's sequence memory. Older children outperformed younger children for sequences that contained phonologically dissimilar words (Jarrold et al., 2015) or pictures (Koppenol-Gonzalez et al., 2014). In contrast, when the sequence contained younger phonologically similar words, 8- and 9-year-olds performed on par with their younger counterparts, aged 5 and 6. Again, these findings suggest that the impact of pattern separation demand modulates the relationship between memory discrimination and age.

Children's failure in source monitoring has been attributed to low cognitive flexibility because such flexibility is needed for judging relevant cues during *meta*-mnemonic judgments at retrieval (Foley et al., 1993). This explanation does not fully account for why source similarity matters and how cognitive flexibility may override memory interference. An alternative interpretation of these findings is that pattern separation demand modulates the age effect on source discrimination. Given the on-going hippocampal development in childhood, younger children are less able to create and store distinct representations for similar experiences. And thus, as pattern separation demand increases, age-related differences in memory discrimination become more apparent. Nonetheless, insights from this literature point to factors that influence children's pattern separation abilities. The source monitoring group of papers further show that the effects of similarity on memory discrimination extend beyond perceptual and semantic similarity between sources. They also apply to actions and their functional overlap between perceptually dissimilar objects (Day et al., 1998; Foley et al., 1994). This literature also identifies factors that impact children's abilities to discriminate similar memories, including the involvement of self (Foley & Ratner, 1998; Roberts & Blades, 1998), retention periods (Roberts & Blades, 1995), and task instructions that draw attention to the differentiating sources during encoding (Roberts et al., 2016). These insights are valuable but not yet available solely from the small literature labeled as pattern separation development.

Another relevant line of work to pattern separation development is research on children's memory for repeating events from Script theory. At first glance, this seems to be a separate research area that primarily focuses on children's memory reconstructive processes, patterns of false memory, and script-based event memories. With our analyses, the similarity manipulation variable identified groups of papers that employed a *repetition* manipulation to investigate script-based memory intrusions (Hudson, 1990b; Hudson et al., 1992; Hudson & Nelson, 1983, 1986; Mccartney & Nelson, 1981; Powell et al., 1999). Children who have experienced recurring events find it difficult to retrieve detail from a particular episode, such that they often mixed up which details were entailed in which instances (Hudson, 1990a; Hudson et al., 1992; Powell et al., 1999). Preschoolers and early school-aged children can rely on generalized event representations to recall stories, but first graders remembered discrepant or unexpected information better than preschool children (Hudson & Nelson, 1983). This line of work suggests that younger children may have disproportionately greater difficulty with retrieving event details against a backdrop of similar events, compared to their older counterparts. From the perspective of CLS, recurring events place a high demand on pattern separation due to their overlap. These findings are in line with the idea that the ongoing development of the pattern separation in early childhood may explain preschoolers' lower abilities in remembering specific instances distinctively from other similar experiences. Further, research in this area shows that scripts are earlier developing and thus younger children overly rely on script knowledge when asked to recall specific instances (Hudson et al., 1992). Such findings expose the need to investigate the co-development between complementary memory processes: those that support the accumulation of regularities across related experiences, and those that support memory specificity for individual events. One viable prediction is that the heterogenous maturational rates among different hippocampal subfields may explain why children's generalized knowledge is apparent earlier than their memories for specific events (Keresztes et al., 2018).

The role of pre-existing semantic knowledge in pattern separation development. A wealth of studies on children's verbatim and gist memories from Fuzzy Trace theory elucidates how *semantic proximity* between concepts may interfere with retrieving a specific memory. This idea is commonly tested using the DRM paradigms, in which participants are first exposed to a list of semantically related words (e.g., bed, blanket), and are later tested on their accurate memory of the learned items as well as the frequency with which they falsely recall or recognize semantic lures (e.g., sleep). Findings on the age difference in false memory for semantic lures are mixed in their directions. Many studies found a "developmental reversal" effect, with false memory for semantic lure increases with age (Anastasi & Rhodes, 2008; Brainerd et al., 2006, 2002; Dewhurst & Robinson, 2004; Dewhurst et al., 2007; Holliday & Weekes, 2006; Howe et al., 2009). One study instead found an age-related decrease in false alarm (Ghetti et al., 2002) – a pattern that is also observed in studies that examined verbatim sentence recognition (Chiu 2006; Reyna & Kiernan, 1994). Some researchers suggest that the age effects on false memory for semantically-related lures may be sensitive to stimulus types (e.g., pictures, words, background distinctiveness (Howe, 2006, 2008), child-normed versus adult-normed lists (Carneiro et al., 2007)). More importantly, two studies show that although false memory for a semantically related lure tends to increase with age, false memory for phonologically related lures decreases with age (Dewhurst & Robinson, 2004; Holliday & Weekes, 2006), underscoring the importance in considering the different dimensions of memory interferences.

Outside of the DRM type of studies, two studies that tested the effect of suggestibility and metamemory in children further qualified the effect of semantic similarity on memory discrimination. In one study, 4-year-old and 9-year-old children were first told a story and later given misinformation about the stories with semantically related lures two days after encoding (Ceci et al., 2010). One week after the misinformation session, children had to discriminate between targets from the original story from those that were introduced during the delay period. Semantic distances for target-lure pairs was were estimated by an independent group of children in another study (Ceci et al., 2007). Discrimination performance interacted with semantic proximity on children's suggestibility. Younger children were worse at discriminating targets from lures than older children, but only on pairs that were more semantically connected for their age group. In fact, they even outperformed older children on pairs that were judged to be semantically far apart by their same-aged peers. These findings converge with studies from the Fuzzy Trace Theory and substantiate the idea that semantic proximity itself is age dependent. Integrating this line of work into the CLS framework tells us that participants' semantic network may, at times, steer the direction of the age effects on memory interference. Thus, we need to be thoughtful in our selection and characterization of the behavioral proxies for pattern separation, especially in the developmental context.

Summary. Bringing together these literatures unveils the common themes that cut cross many different theories and research questions. At the same time, it reveals a complex developmental story. Within the CLS framework, we learn that pattern separation generally improves across childhood, but many factors are at play, including the dimensionality (conceptual versus perceptual) and the complexity (e.g., object, context, sequence, event) of the memoranda. Further, children's pattern separation needs to be taken into consideration with other concurrent developmental stages of existing knowledge and prior exposure. Most importantly, this exercise shows the availability of a much richer dataset for a close examination of a given construct if we search unrestricted from construct labels but based on methodological and experimental factors.

Conclusions

The goal of this mapping review was to descriptively characterize and integrate past literature in memory development from over five decades. In devising a complementary approach to literature integration in experimental psychology, we place a premium on method factors as a gateway to mapping the interconnected webs across different literature within memory development. Our perspective is neutral as we do not intend to espouse of position. Our coverage is intended to be comprehensive and representative within the scope of our focus. The organization of the review places a strong primacy on methodologies, such that works employing similar methods are grouped together (Cooper & Hedges, 2009). This approach helps us bring order to an array of findings, scattered across a large literature of memory development. It has been argued that psychological science suffers from a lack of an overarching theoretical framework, and that the failure to operate in such frameworks is central to both the replication crisis and to building a more cumulative science (Muthukrishna & Henrich, 2019). The formulation of such overarching theoretical frameworks hinges on our abilities to identify the commonalities and differences in the empirical toolkits employed by different fields of literature.

Here, we suggest that it is a systematic framework for combining evidence by commonly studied processes, that may lead the field a step forward. We used memory development as a case study, but of course this general approach could and should be applied to other areas of psychological science. In addition, this work should be viewed as the initial step in the general approach and attitude towards understanding the common threads of methods in unifying findings – one that can be improved, expanded, and upscaled. Crucially, the scalability of this approach rests on researchers' willingness to participate in a stock-taking effort, at least at first. Stock-taking in a large set of publications is necessary in building the first version of such a mapping review analytical pipeline for a given field within psychological science. In fact, this genre of research has been a vibrant discourse in other disciplines, most notably in the field of information systems (Templier & Paré, 2015; Wagner et al., 2021; Webster & Watson, 2002), medical and health sciences (Anderson et al., 2008; Chalmers et al., 2002; Paré & Kitsiou, 2016). Surprising, to our knowledge, cognitive psychology has rarely utilized mapping reviews to cumulate existing evidence and to identify gaps that can guide future research.

Limitations

In this review, we aimed for comprehensiveness and representativeness in the publication identification step. We used a combination of two broad search methods by first applying a relatively broad search terms (e.g., "memory development") and handsearching the references from existing reviews and *meta*-analyses on memory development. Nonetheless, we could not achieve a true exhaustive search strategy that encompasses all of the existing memory development. Thus, the data presented here should not be taken to mean the *absolute* and definitive composition of the literature on memory development.

In an attempt to estimate the publication number that would fulfill our inclusion criteria with an even wider search net on PubMed, we used "memory" AND "development" as separate terms, with an "*and*" operation between them. Using the same years and species specifications as described in our methods, this search results in 17,181 hits. Note, that this is a much more liberal search because it only requires the terms 'memory' and 'development' to be contained in the title or abstract. For example, a first result of such search strategy is titled "Effector and memory T-cell differentiation: implications for vaccine development". In order to grasp a coarse estimate, we performed our identification step with 20 % of the 17,181 results on PubMed by screening for inclusion with the titles and abstracts. Given that Pubmed results can be sorted in a "best match" order, we screened the top 10 % and the bottom 10 % of this list, totaling 3436 titles and abstracts. Note that Pubmed only exports up to 10,000 publications, thus our lowest 10 % selection from this list is not the true bottom 10 % of the full 17181-paper list. Among the 3436 titles and abstracts, a very small percentage (6 %; 230 papers) was identified as relevant according our inclusion and exclusion criteria, and over half of them (53 %) were in our database. This alternative search suggests that our database captures a sizable portion of the literature, although there is room for expansion. The list of relevant but nonoverlapping publications is publicly available on our web application.

Second, our topical coverage leaves out an important and active research area of working memory development (for a thorough review on this topic, see Cowan, 2022). Thus far, our inclusion criteria target assessments of children's long-term memory capacities. Incorporating empirical work on working memory development would be a valuable expansion to the current database. Third, the coding scheme was designed with our 'lens' through which we view the literature. Given our interests, we placed an uneven attention towards different aspects of the memory development, namely the relevant processes for memory purported by neurocomputational models. Thus, the data generated from the coding scheme inevitably contain a biased view on the literature. For example, we placed a primacy on two experimental manipulations, similarity and cue manipulation, that can inform us about the processes of pattern separation, pattern completion, and generalization. Here, we carefully characterized how each manipulation was implemented (e.g., on the conceptual or perceptual dimension). Other design features including the delay period, manipulation of prior knowledge on the memory task, encoding instructions, metacognition assessments, can benefit from further expansion of the current coding scheme. Relatedly, we did not critically assess the selected publications based on methodological quality, logical rigor, completeness or breadth of explanation if theories are involved. Thus, this review does not include a commentary on the quality of the reported research activities.

Final remarks

Building a cumulative psychology requires that we move beyond single-sample studies to the synthesis of findings drawn from multiple studies (discussed in Curran, 2009). Pooled-study analysis and *meta*-analysis can offer solutions to combat the replication crisis in psychology. Adding to our toolkits for literature integration, mapping reviews offer a quantitative way of assessing whether multi-study integration is warranted based on their methodological overlap. There are several take-home points with respect to our outlook. This enterprise will significantly benefit from an active uptake from the scientific community; such joint effort will enhance the accuracy and expansion in the coding scheme. The creation of an interactive web application, we hope, will serve as a launching pad for such uptake and expansion activities. In order to accumulate resources for expansion, our web application incorporates a "suggestion" tab that allows users to input publications that they believe to meet our inclusion criteria but are not yet included in the database. However, the sustainability of an up-to-date database likely requires an automated pipeline. Such pipeline should entail the identification of publications that fit the inclusion criteria. Importantly, automatization of the method classification using large language models would significantly enhance the practicality of mapping reviews in experimental psychology. Fortunately, our database can readily serve as the training ground for such models for the field of memory development. Lastly, building on this work by implementing approaches that focus on findings will considerably shorten the distance between our discipline and a true cumulative science.

CRediT authorship contribution statement

Chi T. Ngo: Conceptualization, Data curation, Writing - original draft, Investigation, Validation, Formal analysis, Methodology, Project administration, Funding acquisition. **Elisa S. Buchberger:** Conceptualization, Data curation, Investigation, Methodology, Project administration. **Phuc T. U. Nguyen:** Visualization, Data curation, Formal analysis, Methodology, Software. **Nora S. Newcombe:** Writing - review & editing. **Markus Werkle-Bergner:** Conceptualization, Supervision, Resources, Writing - review and editing, Funding acquisition.

Declaration of competing interest

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

Data availability

All data and codes are publicly available on our interactive web application (memdev.shinyapps.io/litreview_io/) and github (github.com/memdev2022/mapping_analysis).

Acknowledgments

We would like to thank the coding team of research assistants: Isabel Gerber, Marie Schmidt, Tessa Meyer, Tydings McClary, Wilhelm Voigt, Hannes Steinrüch, Carla Stein, and Viola Werling. C. T. N. was supported by the German Research Foundation, Germany (NG-191/2-1) and the Jacobs Foundation, Switzerland (2021-1417-99). M.W.-B. received financial support from the Jacobs Foundation, Switzerland, through an Early Career Research Fellowship. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dr.2024.101119.

References

Anastasi, J. S., & Rhodes, M. G. (2008). Examining Differences in the Levels of False Memories in Children and Adults Using Child-Normed Lists. *Developmental Psychology, 44*(3), 889–894. https://doi.org/10.1037/0012-1649.44.3.889

Anderson, S., Allen, P., Peckham, S., & Goodwin, N. (2008). Asking the right questions: Scoping studies in the commissioning of research on the organisation and delivery of health services. *Health Research Policy and Systems, 6*, 1–12. https://doi.org/10.1186/1478-4505-6-7

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science, 14*(3), 464–473. https://doi.org/10.1111/j.1467-7687.2009.00937.x

Atance, C. M., & Neill, D. K. O. (2001). Episodic future thinking. Retrieved from *Trends in Cognitive Sciences, 5*(12), 533–539 http://tics.trends.com1364.

Bauer, P., Doydum, A., Pathman, T., Larkina, M., Güler, E., & Burch, M. (2012). It's all about location, location, location: Children's memory for the "where" of personally experienced events. *Journal of Experimental Child Psychology, 113*(4), 510–522. https://doi.org/10.1111/j.1743-6109.2008.01122.x.Endothelial

Bauer, P. J. (2007). Event memory. In W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology* (pp. 373–425). Hoboken, NJ: John Wiley and Sons.

Bauer, P. J., Larkina, M., & Doydum, A. O. (2012). Explaining variance in long-term recall in 3- and 4-year-old children: The importance of post-encoding processes. *Journal of Experimental Child Psychology, 113*(2), 195–210. https://doi.org/10.1016/j.jecp.2012.05.006

Bauer, P. J., & Zelazo, P. D. (2013). NIH toolbox cognition battery (CB): Summary, conclusions, and implications for cognitive development. *Monographs of the Society for Research in Child Development, 78*(4), 133–146. https://doi.org/10.1111/mono.12039

Bennett, I. J., Huffman, D. J., & Stark, C. E. L. (2015). Limbic tract integrity contributes to pattern separation performance across the lifespan. *Cerebral Cortex, 25*(9), 2988–2999. https://doi.org/10.1093/cercor/bhu093

Biggin, F., Emsley, H. C. A., & Knight, J. (2020). Routinely collected patient data in neurology research: A systematic mapping review. *BMC Neurology, 20*(1), 1–9. https://doi.org/10.1186/s12883-020-01993-w

Blanco, N. J., & Sloutsky, V. M. (2019). Adaptive Flexibility in Category Learning? Young Children Exhibit Smaller Costs of Selective Attention Than Adults. *Developmental Psychology, 55*(10), 2060–2076. https://doi.org/10.1037/dev0000777

Bouyeure, A., Patil, S., Mauconduit, F., Poiret, C., Isai, D., & Noulhiane, M. (2021). Hippocampal subfield volumes and memory discrimination in the developing brain. *Hippocampus, 1–13*. https://doi.org/10.1002/hipo.23385

Brainerd, C. J., Forrest, T. J., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology, 42*(5), 962–979. https://doi.org/10.1037/0012-1649.42.5.962

Brainerd, C. J., & Reyna, V. F. (1990). Gist is the grist: Fuzzy-trace theory new intuitionism. *Developmental Review, 10*, 3–47.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100*(1), 42–67. https://doi.org/10.1037/0033-295X.100.1.42

Brainerd, C. J., & Reyna, V. F. (1998). Fuzzy-Trace Theory and Children's False Memories. *Journal of Experimental Child Psychology, 71*(2), 81–129. https://doi.org/10.1006/jecp.1998.2464

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory:Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.), *Advances in Child Development and Behavior* (pp. 41–100). https://doi.org/10.1016/S0065-2407(02)80062-3.

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development, 73*(5), 1363–1377. https://doi.org/10.1111/1467-8624.00477

Brainerd, C. J., Reyna, V. F., & Holliday, R. E. (2018). Developmental reversals in false memory: Development is complementary, not compensatory. *Developmental Psychology, 54*(9), 1773–1784. https://doi.org/10.1037/dev0000554

Brigden, A., Parslow, R. M., Linney, C., Higson-Sweeney, N., Read, R., Loades, M., & Crawley, E. (2019). How are behavioural interventions delivered to children (5–11 years old): A systematic mapping review. *BMJ Paediatrics Open, 3*, e000543.

Brown, A. L., & Murphy, M. D. (1975). Reconstruction of arbitrary versus logical sequences by preschool children. *Journal of Experimental Child Psychology, 20*(2), 307–326. https://doi.org/10.1016/0022-0965(75)90106-X

Brown, A. L., & Scott, M. S. (1971). Recognition memory for pictures in preschool children. *Journal of Experimental Child Psychology, 11*(3), 401–412. https://doi.org/10.1016/0022-0965(71)90045-2

Brown, J., Aczel, B., Jiménez, L., Kaufman, S. B., & Grant, K. P. (2010). Intact implicit learning in autism spectrum conditions. *Quarterly Journal of Experimental Psychology, 63*(9), 1789–1812. https://doi.org/10.1080/17470210903536910

Brown, R. D., & Chiu, C. Y. P. (2006). Neural correlates of memory development and learning: Combining neuroimaging and behavioral measures to understand cognitive and developmental processes. *Developmental Neuropsychology, 29*(2), 279–291. https://doi.org/10.1207/s15326942dn2902

Brubacher, S. P., Earhart, B., Roberts, K. P., & Powell, M. B. (2018). Effects of label training and recall order on children's reports of a repeated event. *Applied Cognitive Psychology, 32*(5), 600–609. https://doi.org/10.1002/acp.3440

Brubacher, S. P., Glisic, U., Roberts, K. P., & Powell, M. (2011). Children's ability to recall unique aspects of one occurrence of a repeated event. *Applied Cognitive Psychology, 25*(3), 351–358. https://doi.org/10.1002/acp.1696

Brubacher, S. P., Roberts, K. P., & Powell, M. (2011). Effects of practicing episodic versus scripted recall on children's subsequent narratives of a repeated event. *Psychology, Public Policy, and Law, 17*(2), 286–314. https://doi.org/10.1037/a0022793

Brubacher, S. P., Roberts, K. P., & Powell, M. (2012). Retrieval of episodic versus generic information: Does the order of recall affect the amount and accuracy of details reported by children about repeated events? *Developmental Psychology, 48*(1), 111–122. https://doi.org/10.1037/a0025864

Bruce, C. R., Delgado, E., Kostick, K., Grogan, S., Ashrith, G., Trachtenberg, B., & Blumenthal-Barby, J. S. (2014). Ventricular assist devices: A review of psychosocial risk factors and their impact on outcomes. *Journal of Cardiac Failure, 20*(12), 996–1003. https://doi.org/10.1016/j.cardfail.2014.09.006

Buckner, R. L., Petersen, S. E., Ojemann, J. G., Miezin, F. M., Squire, L. R., & Raichle, M. E. (1995). Functional anatomical studies of explicit and implicit memory retrieval tasks. *Journal of Neuroscience, 15*(1 I), 12–29. https://doi.org/10.1523/jneurosci.15-01-00012.1995

Burden, M. J., & Mitchell, D. B. (2005). Implicit memory development in school-aged children with attention deficit hyperactivity disorder (ADHD): Conceptual priming deficit? *Developmental Neuropsychology, 28*(3), 779–807. https://doi.org/10.1207/s15326942dn2803

Burns, P., Russell, C., & Russell, J. (2015). Preschool children's proto-episodic memory assessed by deferred imitation. *Memory, 23*(8), 1172–1192. https://doi.org/10.1080/09658211.2014.963625

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*(6), 723–736. https://doi.org/10.1002/dev.21946

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*(8). https://doi.org/10.1093/cercor/bhy211

Canada, K., Ngo, C., Newcombe, N., Geng, F., & Riggins, T. (2018). It's all in the details: Relations between young children's developing pattern separation abilities and hippocampal subfield volumes. *Cerebral Cortex*.

Carneiro, P., Fernandez, A., Albuquerque, P., & Esteves, F. (2007). Analyzing false memories in children with associative lists specific for their age. *Child Development, 78*(4), 1171–1185.

Ceci, S. J., Fitneva, S. A., & Williams, W. M. (2010). Representational Constraints on the Development of Memory and Metamemory: A Developmental-Representational Theory. *Psychological Review, 117*(2), 464–495. https://doi.org/10.1037/a0019067

Ceci, S. J., Lea, S. E., & Howe, M. J. (1980). Structural analysis of memory traces in children from 4 to 10 years of age. *Developmental Psychology, 16*(3), 203–212. https://doi.org/10.1037/0012-1649.16.3.203

Ceci, S. J., Papierno, P. B., & Kulkofsky, S. (2007). Representational constraints on children's suggestibility: Research article. *Psychological Science, 18*(6), 503–509. https://doi.org/10.1111/j.1467-9280.2007.01930.x

Chalmers, I., Hedges, L. V., & Cooper, H. (2002). A brief history of research synthesis. *Evaluation and the Health Professions, 25*(1), 12–37. https://doi.org/10.1177/0163278702025001003

Cheke, L. G., & Clayton, N. S. (2015). The six blind men and the elephant: Are episodic memory tasks tests of different things or different tests of the same thing? *Journal of Experimental Child Psychology, 137*, 164–171. https://doi.org/10.1016/j.jecp.2015.03.006

Connolly, D. A., & Price, H. L. (2006). Children's suggestibility for an instance of a repeated event versus a unique event: The effect of degree of association between variable details. *Journal of Experimental Child Psychology, 93*(3), 207–223. https://doi.org/10.1016/j.jecp.2005.06.004

Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of reserach synthesis and meta-analysis* (pp. 3–17). New York: Russell Sage Foundation.

Cowan, N. (2022). Working memory development: A 50-year assessment of research and underlying theories. *Cognition, 224*(March), Article 105075. https://doi.org/10.1016/j.cognition.2022.105075

Cruwys, T., Brossard, B., Zhou, H., Helleren-Simpson, G., Klik, K. A., Van Rooy, D., & Calear, A. L. (2021). Disciplinary differences in the study of the relationship between social variables and mental health: A systematic mapping review. *Health*. https://doi.org/10.1177/13634593211063049

Cuevas, K., Rajan, V., Morasch, K. C., & Bell, M. A. (2015). Episodic memory and future thinking during early childhood: Linking the past and future. *Developmental Psychobiology, 57*(5), 552–565. https://doi.org/10.1002/dev.21307

Cuevas, K., & Sheya, A. (2019). Ontogenesis of learning and memory: Biopsychosocial and dynamical systems perspectives. *Developmental Psychobiology, 61*(3), 402–415. https://doi.org/10.1002/dev.21817

Curran, P. J. (2009). The seemingly quixotic pursuit of a cumulative psychological science: Introduction to the special issue. *Psychological Methods, 14*(2), 77–80. https://doi.org/10.1037/a0015972

Cycowicz, Y. M. (2000). Memory development and event-related brain potentials in children. *Biological Psychology, 54*(1–3), 145–174. https://doi.org/10.1016/S0301-0511(00)00055-7

Danby, M. C., Brubacher, S. P., Sharman, S. J., Powell, M. B., & Roberts, K. P. (2017). Children's Reasoning About Which Episode of a Repeated Event is Best Remembered. *Applied Cognitive Psychology, 31*(1), 99–108. https://doi.org/10.1002/acp.3306

Danby, M. C., Sharman, S. J., Brubacher, S. P., Powell, M. B., & Roberts, K. P. (2017). Differential effects of general versus cued invitations on children's reports of a repeated event episode. *Psychology, Crime and Law, 23*(8), 794–811. https://doi.org/10.1080/1068316X.2017.1324028

Day, K., Howie, P., & Markham, R. (1998). The role of similarity in developmental differences in source monitoring. *British Journal of Developmental Psychology, 16*, 219–232. https://doi.org/10.1111/j.2044-835X.1998.tb00920.x

de Haan, M., Mishkin, M., Baldeweg, T., & Vargha-Khadem, F. (2006). Human memory development and its dysfunction after early hippocampal injury. *Trends in Neurosciences, 29*(7), 374–381. https://doi.org/10.1016/j.tins.2006.05.008

Deese, J. (1965). *The structure of associations in language and thought*. Baltimore: The Johns Hopkins Press.

Deng, W., & Sloutsky, V. M. (2012). Carrot eaters or moving heads: Inductive inference is better supported by salient features than by category labels. *Psychological Science, 23*(2), 178–186. https://doi.org/10.1177/0956797611429133

Dewhurst, S. A., Pursglove, R. C., & Lewis, C. (2007). Story contexts increase susceptibility to the DRM illusion in 5-year-olds. *Developmental Science, 10*(3), 374–378. https://doi.org/10.1111/j.1467-7687.2007.00592.x

Dewhurst, S. A., & Robinson, C. A. (2004). *False Memories in Children. Society, 15*(11), 782–787.

Dikmen, S. S., Bauer, P. J., Weintraub, S., Mungas, D., Slotkin, J., Beaumont, J. L., & Heaton, R. K. (2014). Measuring episodic memory across the lifespan: NIH toolbox picture sequence memory test. *Journal of the International Neuropsychological Society, 20*(6), 611–619. https://doi.org/10.1017/S1355617714000460

Drummey, A. B., & Newcombe, N. S. (1995). Drummey 1995.pdf. *Journal of Experimental Child Psychology, 59*, 549–565.

Drummey, A. B., & Newcombe, N. S. (2002). Developmental changes in source memory. *Developmental Science, 5*(4), 502–513. https://doi.org/10.1111/1467-7687.00243

Esdaile, E., Thow, A. M., Gill, T., Sacks, G., Golley, R., Love, P., & Rissel, C. (2019). National policies to prevent obesity in early childhood: Using policy mapping to compare policy lessons for Australia with six developed countries. *Obesity Reviews, 20*(11), 1542–1556. https://doi.org/10.1111/obr.12925

Farrar, M. J., & Boyer-Pennington, M. E. (1999). Remembering Specific Episodes of a Scripted Event. *Journal of Experimental Child Psychology, 73*(4), 266–288. https://doi.org/10.1006/jecp.1999.2507

Fernández-Sotos, P., Torio, I., Fernández-Caballero, A., Navarro, E., González, P., Dompablo, M., & Rodriguez-Jimenez, R. (2018). Social cognition remediation interventions: A systematic mapping review. *PLoS ONE, 14*(6), 1–20. https://doi.org/10.1371/journal.pone.0218720

Fischer, S., Wilhelm, I., & Born, J. (2007). Developmental differences in sleep's role for implicit off-line learning: Comparing children with adults. *Journal of Cognitive Neuroscience, 19*(2), 214–227. https://doi.org/10.1162/jocn.2007.19.2.214

Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology, 62*, 559–582. https://doi.org/10.1146/annurev.psych.121208.131702

Foley, M. A., & Johnson, M. K. (1983). Age-related changes in confusion between memories for thoughts and memories for speech. *Child Development, 54*, 51–60.

Foley, M. A., Aman, C., & Gutch, D. (1987). Discriminating between action memories: Children's use of kinesthetic cues and visible consequences. *Journal of Experimental Child Psychology, 44*(3), 335–347. https://doi.org/10.1016/0022-0965(87)90038-5

Foley, M. A., Harris, J. F., & Hermann, S. (1994). Developmental comparisons of the ability to discriminate between memories for symbolic play enactments. *Developmental Psychology, 30*(2), 206–217. https://doi.org/10.1037/0012-1649.30.2.206

Foley, M. A., & Johnson, M. K. (1985). Confusions between memories for performed and imagined actions: A developmental comparison. *Child Development, 56*, 1145–1155.

Foley, M. A., & Ratner, H. H. (1998). Children's recoding in memory for collaboration: A way of learning from others. *Cognitive Development, 13*(1), 91–108. https://doi.org/10.1016/S0885-2014(98)90022-3

Foley, M. A., Santini, C., & Sopasakis, M. (1989). Discriminating between memories: Evidence for children's spontaneous elaborations. *Journal of Experimental Child Psychology, 48*(1), 146–169. https://doi.org/10.1016/0022-0965(89)90045-3

Foley, M. A., Wilder, A., McCall, R., & Vorst, R. V. (1993). The consequences for recall of children's ability to generate interactive imagery in the absence of external supports. *Journal of Experimental Child Psychology, 56*, 173–200.

Gabrieli, J. D. E., Fleischman, D. A., Keane, M. M., Reminger, S. L., & Morrell, F. (1995). Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science, 6*(2), 76–82.

Ghetti, S., Qin, J., & Goodman, G. S. (2002). False memories in children and adults: Age, distinctiveness, and subjective experience. *Developmental Psychology, 38*(5), 705–718. https://doi.org/10.1037/0012-1649.38.5.705

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia, 53*, 104–114. https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Glendøs, M., & Berliner, P. (2017). Forty years of research concerning children and youth in Greenland: A mapping review. *International Journal of Circumpolar Health, 76*(1323526), 1–13. https://doi.org/10.1080/22423982.2017.1323526

Gómez, C. M., Barriga-Paulino, C. I., Rodríguez-Martínez, E. I., Rojas-Benjumea, M.Á., Arjona, A., & Gómez-González, J. (2018). The neurophysiology of working memory development: From childhood to adolescence and young adulthood. *Reviews in the Neurosciences, 29*(3), 261–282. https://doi.org/10.1515/revneuro-2017-0073

Gómez, R., & Edgin, J. (2016). The extended trajectory of hippocampal development: Implications for early memory development and disorder. *Developmental Cognitive Neuroscience, 18*, 57–69. https://doi.org/10.1016/S2214-109X(16)30265-0.Cost-effectiveness

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(3), 501–518. https://doi.org/10.1037/0278-7393.11.3.501

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal, 26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Green, M. F., Penn, D. L., Bentall, R., Carpenter, W. T., Gaebel, W., Gur, R. C., & Heinssen, R. (2008). Social cognition in schizophrenia: An NIMH workshop on definitions, assessment, and research opportunities. *Schizophrenia Bulletin, 34*(6), 1211–1220. https://doi.org/10.1093/schbul/sbm145

Greenbaum, J. L., & Graf, P. (1989). Preschool period development of implicit and explicit remembering. *Bulletin of the Psychonomic Society, 27*(5), 417–420. https://doi.org/10.3758/BF03334643

Gulya, M., Rossi-George, A., Hartshorn, K., Vieira, A., Rovee-Collier, C., Johnson, M. K., & Chalfonte, B. L. (2002). The development of explicit memory for basic perceptual features. *Journal of Experimental Child Psychology, 81*(3), 276–297. https://doi.org/10.1006/jecp.2001.2654

Guttentag, R., & Dunn, J. (2003). Judgments of remembering: The revelation effect in children and adults. *Journal of Experimental Child Psychology, 86*(2), 153–167. https://doi.org/10.1016/S0022-0965(03)00135-8

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. https://doi.org/10.1002/hipo.23146

Hayne, H., & Imuta, K. (2011). Episodic memory in 3- and 4-year-old children. *Developmental Psychobiology, 53*(3), 317–322. https://doi.org/10.1002/dev.20527

Holliday, R. E., Reyna, V. F., & Brainerd, C. J. (2008). Recall of details never experienced: Effects of age, repetition, and semantic cues. *Cognitive Development, 23*(1), 67–78. https://doi.org/10.1016/j.cogdev.2007.05.002

Holliday, R. E., & Weekes, B. S. (2006). Dissociated developmental trajectories for semantic and phonological false memories. *Memory, 14*(5), 624–636. https://doi.org/10.1080/09658210600736525

Holten, D. (2006). Hierarchical edge bundles: Visualization of adjacencyrelations in hierarchical data. *IEEE Transactions on Visualization and Computer Graphics, 12*(5), 741–748.

Hömberg, V., Bickmann, U., & Müller, K. (1993). Ontogeny is different for explicit and implicit memory in humans. *Neuroscience Letters, 150*(2), 187–190. https://doi.org/10.1016/0304-3940(93)90532-P

Horner, A. J., & Burgess, N. (2014). Pattern completion in multielement event engrams. *Current Biology, 24*(9), 988–992. https://doi.org/10.1016/j.cub.2014.03.012

Howe, M. L. (2006). Developmentally invariant dissociations in children's true and false memories: Not all relatedness is created equal. *Child Development, 77*(4), 1112–1123. https://doi.org/10.1111/j.1467-8624.2006.00922.x

Howe, M. L. (2008). Visual distinctiveness and the development of children's false memories. *Child Development, 79*(1), 65–79. https://doi.org/10.1111/j.1467-8624.2007.01111.x

Howe, M. L. (2013). Memory development: Implications for adults recalling childhood experiences in the courtroom. *Nature Reviews Neuroscience, 14*(12), 869–876. https://doi.org/10.1038/nrn3627

Howe, M. L., & Courage, M. L. (1993). On resolving the enigma of infantile amnesia. *Psychological Bulletin, 113*(2), 305–326. https://doi.org/10.1037/0033-2909.113.2.305

Howe, M. L., & Wilkinson, S. (2011). Using story contexts to bias children's true and false memories. *Journal of Experimental Child Psychology, 108*(1), 77–95. https://doi.org/10.1016/j.jecp.2010.06.009

Howe, M. L., Wimmer, M. C., & Blease, K. (2009). The role of associative strength in children's false memory illusions. *Memory, 17*(1), 8–16. https://doi.org/10.1080/09658210802438474

Hudson, J. A. (1990). Constructive processing in children's event memory. *Developmental Psychology, 26*(2), 180–187. https://doi.org/10.1037/0012-1649.26.2.180

Hudson, J. A., & Mayhew, E. M. Y. (2009). The development of memory in infancy and childhood. In M. L. Courage, & N. Cowan (Eds.), *The development of memory for recurring events* (pp. 69–91). Psychology Press.

Hudson, J., Fivush, R., & Kuebli, J. (1992). Scripts and episodes: The development of event memory. *Applied Cognitive Psychology, 6*, 483–505. https://doi.org/10.1159/000407966

Hudson, J., & Nelson, K. (1983). Effects of script structure on children's story recall. *Developmental Psychology, 19*(4), 625–635. https://doi.org/10.1037/0012-1649.19.4.625

Hudson, J., & Nelson, K. (1986). Repeated encounters of a similar kind: Effects of familiarity on children's autobiographic memory. *Cognitive Development, 1*(3), 253–271. https://doi.org/10.1016/S0885-2014(86)80004-1

Iyer, E. S., Weinberg, A., & Bagot, R. C. (2021). Ambiguity and conflict: Dissecting uncertainty in decision-making. *Behavioral Neuroscience*. https://doi.org/10.1037/bne0000489

James, E., Ong, G., Henderson, L.-M., & Horner, A. J. (2021). The formation and retrieval of holistic event memories across development. *Journal of Cognition*. https://doi.org/https://doi.org/10.5334/joc.149.

James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence, 5*(1), 1–13. https://doi.org/10.1186/s13750-016-0059-6

Jarrold, C., Hall, D., Harvey, C. E., Tam, H., Towse, J. N., & Zarandi, A. L. (2015). What can we learn about immediate memory from the development of children's free recall? *Quarterly Journal of Experimental Psychology, 68*(9), 1871–1894. https://doi.org/10.1080/17470218.2014.995110

Johnson, E. G., Prabhakar, J., Mooney, L. N., & Ghetti, S. (2020). Neuroimaging the sleeping brain: Insight on memory functioning in infants and toddlers. *Infant Behavior and Development, 58*. https://doi.org/10.1016/j.infbeh.2020.101427

Johnson, M., Hashtroudi, S., & Lindsay, S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3–28.

Johnson, M. K., Raye, C. L., Hasher, L., & Chromiak, W. (1979). Are there developmental differences in reality-monitoring? *Journal of Experimental Child Psychology, 27*, 120–128. https://doi.org/10.1016/0022-0965(79)90064-X

Jung, Y., Walther, D. B., & Finn, A. S. (2020). Children automatically abstract categorical regularities during statistical learning. *Developmental Science, November*, 1–12. https://doi.org/10.1111/desc.13072

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*(34), 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*(8), 676–686. https://doi.org/10.1016/j.tics.2018.05.004

Keresztes, A., Raffington, L., Bender, A. R., Bögl, K., Heim, C., & Shing, Y. L. (2020). Hair cortisol concentrations are associated with hippocampal subregional volumes in children. *Scientific Reports, 1–12*. https://doi.org/10.1038/s41598-020-61131-x

Keresztes, A., Raffington, L., Bender, A. R., Bögl, K., Heim, C., & Shing, Y. L. (2022). Longitudinal developmental trajectories do not follow cross-sectional age associations in hippocampal subfield and memory development. *Developmental Cognitive Neuroscience, 54*. https://doi.org/10.1016/j.dcn.2022.101085

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

Koppenol-Gonzalez, G. V., Bouwmeester, S., & Vermunt, J. K. (2014). Short-term memory development: Differences in serial position curves between age groups and latent classes. *Journal of Experimental Child Psychology, 126*, 138–151. https://doi.org/10.1016/j.jecp.2014.04.002

Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. L. (2011). Distinct pattern separation related transfer functions in human CA3 / dentate and CA1 revealed using high- resolution fMRI and variable mnemonic similarity. *Learning & Memory*, *18*(Ward 2001), 15–18. https://doi.org/10.1101/lm.1971111.

Lambert, F. R., Lavenex, P., & Lavenex, P. B. (2015). Improvement of allocentric spatial memory resolution in children from 2 to 4 years of age. *International Journal of Behavioral Development, 39*(4), 318–331. https://doi.org/10.1177/0165025415584808

Landau, B., Smith, L., & Jones, S. (1998). Object perception and object naming in early development. *Trends in Cognitive Sciences, 2*(1), 19–24.

Lavenex, P., & Banta Lavenex, P. (2013). Building hippocampal circuits to learn and remember: Insights into the development of human memory. *Behavioural Brain Research, 254*, 8–21. https://doi.org/10.1016/j.bbr.2013.02.007

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). https://doi.org/10.1007/978-3-319-50406-3.

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*, Article 100741. https://doi.org/10.1016/j.dcn.2019.100741

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. https://doi.org/10.1111/cdev.12447

Light, R. J., & Smith, P. V. (1971). Accumulating evidence : Procedures for resolving contradictions among different research studies. *Harvard Educational Review, 41* (429–471).

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*(1), 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*(3), 297–318. https://doi.org/10.1016/0022-0965(91)90065-Z

Liu, K. Y., Gould, R. L., Coulson, M. C., Ward, E. V., & Howard, R. J. (2016). Tests of pattern separation and pattern completion in humans-A systematic review. *Hippocampus, 26*(6), 705–717. https://doi.org/10.1002/hipo.22561

Lloyd, M. E., Doydum, A. O., & Newcombe, N. S. (2009). Memory binding in early childhood: Evidence for a retrieval deficit. *Child Development, 80*(5), 1321–1328. https://doi.org/10.1111/j.1467-8624.2009.01353.x

Lloyd, M. E., & Miller, J. K. (2014). Implicit Memory. In P. J. Bauer, & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory* (pp. 336–359). Wiley Blackwell.

Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences, 5*(10), 443–450. https://doi.org/10.1016/S1364-6613(00)01752-6

Markham, R. (1991). Development of reality monitoring for performed and imagined actions. *Perceptual and Motor Skills, 72*, 1347–1354.

Mastrogiuseppe, M., Bertelsen, N., Bedeschi, M. F., & Lee, S. A. (2019). The spatiotemporal organization of episodic memory and its disruption in a neurodevelopmental disorder. *Scientific Reports, 9*(1), 1–12. https://doi.org/10.1038/s41598-019-53823-w

Mccartney, K. A., & Nelson, K. (1981). Children's use of scripts in story recall. *Discourse Processes, 4*, 59–70. https://doi.org/10.1080/01638538109544506

McClain, M. B., Callan, G. L., Harris, B., Floyd, R. G., Haverkamp, C. R., Golson, M. E., & Benallie, K. J. (2021). Methods for addressing publication bias in school psychology journals: A descriptive review of meta-analyses from 1980 to 2019. *Journal of School Psychology, 84*, 74–94. https://doi.org/10.1016/j.jsp.2020.11.002

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*(3), 213–219. https://doi.org/10.1007/978-3-642-11202-7_20

McGeown, S. P., Gray, E. A., Robinson, J. L., & Dewhurst, S. A. (2014). What factors underlie children's susceptibility to semantic and phonological false memories? Investigating the roles of language skills and auditory short-term memory. *Cognition, 131*(3), 323–329. https://doi.org/10.1016/j.cognition.2014.02.005

McGrady, M. E., & Hommel, K. A. (2016). Targeting health behaviors to reduce health care costs in pediatric psychology: Descriptive review and recommendations. *Journal of Pediatric Psychology, 41*(8), 835–848. https://doi.org/10.1093/jpepsy/jsv083

Metzger, R. L., Warren, A. R., Shelton, J. T., Price, J., Reed, A. W., & Williams, D. (2008). Do Children "DRM" Like Adults? False Memory Production in Children. *Developmental Psychology, 44*(1), 169–181. https://doi.org/10.1037/0012-1649.44.1.169

Moreau, D., & Wiebels, K. (2022). Psychological constructs as local optima. *Nature Reviews Psychology*. https://doi.org/10.1038/s44159-022-00042-2

Morris, G., Baker-Ward, L., & Bauer, P. J. (2010). What remains of that day: The survival of children's autobiographical memories across time. *Applied Cognitive Psychology, 24*, 527–544. https://doi.org/10.1002/acp

Mullally, S. L., & Maguire, E. A. (2014). Learning to remember: The early ontogeny of episodic memory. *Developmental Cognitive Neuroscience, 9*, 12–29. https://doi.org/10.1016/j.dcn.2013.12.006

Murphy, K., McKone, E., & Slee, J. (2003). Dissociations between implicit and explicit memory in children: The role of strategic processing and the knowledge base. *Journal of Experimental Child Psychology, 84*(2), 124–165. https://doi.org/10.1016/S0022-0965(03)00002-X

Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour, 3*(3), 221–229. https://doi.org/10.1038/s41562-018-0522-1

Nelson, C. A. (1998). The nature of early memory. *Preventive Medicine, 27*(2), 172–179. https://doi.org/10.1006/pmed.1998.0272

Nelson, K. (2007). Development of extended memory. *Journal of Physiology - Paris, 101*(4–6), 223–229. https://doi.org/10.1016/j.jphysparis.2007.11.004

Nelson, K., & Gruendel, J. (1981). Generalized event representations: Basic building blcoks of cognitive development. In A. Brown, & M. Lamb (Eds.), *Advances in Developmental Psychology*. New Jersey: Hilllsdale.

Ngo, C. T., Horner, A. J., Newcombe, N. S., & Olson, I. R. (2019). Development of Holistic Episodic Recollection. *Psychological Science, 30*(12), 1696–1706. https://doi.org/10.1177/0956797619879441

Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2018). The ontogeny of relational memory and pattern separation. *Developmental Science, 21*(2). 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*(5), 1569–1578. https://doi.org/10.1111/cdev.13297

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*, 1–8. https://doi.org/10.1016/j.cub.2021.03.088

Ngo, C. T., Lin, Y.i., 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. https://doi.org/10.1037/xge0000583

Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2017). The ontogeny of relational memory and pattern separation. *Developmental Science, 21*(2). https://doi.org/10.1111/desc/12556

Norman, K. A., & O'Reilly, R. C. (2003). Modeling Hippocampal and Neocortical Contributions to Recognition Memory: A Complementary-Learning-Systems Approach. *Psychological Review, 110*(4), 611–646. https://doi.org/10.1037/0033-295X.110.4.611

Ofen, N. (2012). The development of neural correlates for memory formation. *Neuroscience and Biobehavioral Reviews, 36*(7), 1708–1717. https://doi.org/10.1016/j.neubiorev.2012.02.016

Ofen, N., Tang, L., Yu, Q., & Johnson, E. L. (2019). Memory and the developing brain: From description to explanation with innovation in methods. *Developmental Cognitive Neuroscience, 36*(December 2018), Article 100613. https://doi.org/10.1016/j.dcn.2018.12.011

Ornstein, P. A. (1995). Children's long-term retention of salient personal experiences. *Journal of Traumatic Stress, 8*(4), 581–605. https://doi.org/10.1007/BF02102890

Paré, G., & Kitsiou, S. (2016). Methods for literature reviews. In F. Lau & C. Kuziemsky (Eds.), *Handbook of eHealth evaluation: An evidence-based approach* (pp. 157–179). https://doi.org/10.1353/book.68987.

Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information and Management, 52*(2), 183–199. https://doi.org/10.1016/j.im.2014.08.008

Pathman, T., Coughlin, C., & Ghetti, S. (2018). Space and time in episodic memory: Effects of linearity and directionality on memory for spatial location and temporal order in children and adults. *PLoS ONE, 13*(11), 1–18. https://doi.org/10.1371/journal.pone.0206999

Pathman, T., & Ghetti, S. (2014). The Eyes Know Time: A Novel Paradigm to Reveal the Development of Temporal Memory. *Child Development, 85*(2), 792–807. https://doi.org/10.1111/cdev.12152

Pathman, T., & Ghetti, S. (2016). More to it than meets the eye: How eye movements can elucidate the development of episodic memory. *Memory, 24*(6), 721–736. https://doi.org/10.1080/09658211.2016.1155870

Pathman, T., Larkina, M., Burch, M., & Bauer, P. J. (2013). Young children's memory for the times of personal past events. *Journal of Cognition and Development, 14*(1), 120–140. https://doi.org/10.2217/nnm.12.167.Gene

Pawar, A. V., & Spence, S. A. (2003). Defining thought broadcast: Semi-structured literature review. *British Journal of Psychiatry, 183*, 287–291. https://doi.org/10.1192/bjp.183.4.287

Peters, B. C. M., & Wood, W. (2017). Autism and equine-assisted Interventions: A systematic mapping review. *Journal of Autism and Developmental Disorders, 47*(10), 3220–3242. https://doi.org/10.1007/s10803-017-3219-9

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. *12th International Conference in Evaluation Assessment in Software Engineer*, 1.

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

Picard, L., Cousin, S., Guillery-Girard, B., Eustache, F., & Piolino, P. (2012). How Do the Different Components of Episodic Memory Develop? Role of Executive Functions and Short-Term Feature-Binding Abilities. *Child Development, 83*(3), 1037–1050. https://doi.org/10.1111/j.1467-8624.2012.01736.x

Powell, M. B., Roberts, K. P., Ceci, S. J., & Hembrooke, H. (1999). The effects of repeated experience on children's suggestibility. *Developmental Psychologysychology, 35*(6), 1462–1477. https://doi.org/10.1037/0012-1649.35.6.1462

Price, H. L., & Connolly, D. A. (2013). Suggestibility effects persist after one year in children who experienced a single or repeated event. *Journal of Applied Research in Memory and Cognition, 2*(2), 89–94. https://doi.org/10.1016/j.jarmac.2013.03.001

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*(1–2), 25–45. https://doi.org/10.1080/02643294.2019.1667316

Ramsaran, A. I., Schlichting, M. L., & Frankland, P. W. (2019). The ontogeny of memory persistence and specificity. *Developmental Cognitive Neuroscience, 36*, Article 100591. https://doi.org/10.1016/j.dcn.2018.09.002

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*(1), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

Richmond, J. L., & Pan, R. (2013). Thinking about the future early in life: The role of relational memory. *Journal of Experimental Child Psychology, 114*(4), 510–521. https://doi.org/10.1016/j.jecp.2012.11.002

Roberts, K. P. (2002). Children's ability to distinguish between memories from multiple sources: Implications for the quality and accuracy of eyewitness statements. *Developmental Review, 22*(3), 403–435. https://doi.org/10.1016/S0273-2297(02)00005-9

Roberts, K. P., & Blades, M. (1995). Children's discrimination of memories for actual and pretend actions in a hiding task. *British Journal of Developmental Psychology, 13*(4), 321–333. https://doi.org/10.1111/j.2044-835x.1995.tb00683.x

Roberts, K. P., & Blades, M. (1998). The effects of interacting in repeated events on children's eyewitness memory and source monitoring. *Applied Cognitive Psychology, 12*(5), 489–503. https://doi.org/10.1002/(sici)1099-0720(199810)12:5<489::aid-acp535>3.0.co;2-#

Roberts, K. P., Brubacher, S. P., Drohan-Jennings, D., Glisic, U., Powell, M. B., & Friedman, W. J. (2015). Developmental Differences in the Ability to Provide Temporal Information About Repeated Events. *Applied Cognitive Psychology, 29*(3), 407–417. https://doi.org/10.1002/acp.3118

Roberts, K. P., Evans, A. D., & Duncanson, S. (2016). Binding an event to its source at encoding improves children's source monitoring. *Developmental Psychology, 52*(12), 2191–2201. https://doi.org/10.1037/dev0000213

Roberts, K. P., & Powell, M. B. (2006). The consistency of false suggestions moderates children's reports of a single instance of a repeated event: Predicting increases and decreases in suggestibility. *Journal of Experimental Child Psychology, 94*(1), 68–89. https://doi.org/10.1016/j.jecp.2005.12.003

Roberts, M., Ashinoff, B. K., Castellanos, F. X., & Carrasco, M. (2018). When attention in intact in adults with ADHD. *Psychonomic Bulletin & Review, 25*(4), 1423–1434. https://doi.org/10.3758/s13423-017-1407-4.When

Robinson, C. W., Best, C. A., Deng, W., & Sloutsky, V. M. (2012). The role of words in cognitive tasks: What, when, and how? *Frontiers in Psychology, 3*(95), 1–8. https://doi.org/10.3389/fpsyg.2012.00095

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Rollins, L., & Cloude, E. B. (2018). Development of mnemonic discrimination during childhood. *Learning and Memory, 25*(6), 294–297. https://doi.org/10.1101/lm.047142.117

Rolls, E. T. (2016). Pattern separation, completion, and categorisation in the hippocampus and neocortex. *Neurobiology of Learning and Memory, 129*. https://doi.org/10.1016/j.nlm.2015.07.008

Rovee-Collier, C., & Cuevas, K. (2009). Multiple memory systems are unnecessary to account for infant memory development: An ecological model. *Developmental Psychology, 45*(1), 160–174. https://doi.org/10.1037/a0014538.Multiple

Roze des Ordons, A. L., Sinuff, T., Stelfox, H. T., Kondejewski, J., & Sinclair, S. (2018). Spiritual distress within inpatient settings—A scoping review of patients' and families' experiences. *Journal of Pain and Symptom Management, 56*(1), 122–145. https://doi.org/10.1016/j.jpainsymman.2018.03.009

Sadoughi, F., Behmanesh, A., & Sayfouri, N. (2020). Internet of things in medicine: A systematic mapping study. *Journal of Biomedical Informatics, 103*(February), Article 103383. https://doi.org/10.1016/j.jbi.2020.103383

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Sander, M. C., Werkle-Bergner, M., Gerjets, P., Shing, Y. L., & Lindenberger, U. (2012). The two-component model of memory development, and its potential implications for educational settings. *Developmental Cognitive Neuroscience, 2*(Suppl 1), S67–S77. https://doi.org/10.1016/j.dcn.2011.11.005

Schacter, D. L., Chiu, P. C. Y., & Ochsner, K. N. (1993). Implicit memory: A selective review. *Annual Review of Neuroscience, 16*, 159–182. https://doi.org/10.1016/B978-012264841-0/50024-X

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1711). https://doi.org/10.1098/rstb.2016.0049

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*(1), 37–51.

Schneider, W., & Ornstein, P. A. (2019). Determinants of memory development in childhood and adolescence. *International Journal of Psychology, 54*(3), 307–315. https://doi.org/10.1002/ijop.12503

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews, 34*(7), 1080–1091. https://doi.org/10.1016/j.neubiorev.2009.11.002

Shufaniya, A., & Arnon, I. (2018). Statistical learning is not age-invariant during childhood: Performance improves with age across modality. *Cognitive Science, 42*(8), 3100–3115. https://doi.org/10.1111/cogs.12692

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*(1), 89–100. https://doi.org/10.1037/0278-7393.32.1.89

Sommer, V. R., Mount, L., Weigelt, S., Werkle-Bergner, M., & Sander, M. C. (2021). Memory specificity is linked to repetition effects in event-related potentials across the lifespan. *Developmental Cognitive Neuroscience, 48*(January), Article 100926. https://doi.org/10.1016/j.dcn.2021.100926

Stark, S. M., & Stark, C. E. L. (2017). Age-related deficits in the mnemonic similarity task for objects and scenes. *Behavioural Brain Research, 333*(June), 109–117. https://doi.org/10.1016/j.bbr.2017.06.049

Stark, S. M., Yassa, M. A., 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*(12). https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Symmank, C., Mai, R., Hoffmann, S., Stok, F. M., Renner, B., Lien, N., & Rohm, H. (2017). Predictors of food decision making: A systematic interdisciplinary mapping (SIM) review. *Appetite, 110*, 25–35. https://doi.org/10.1016/j.appet.2016.11.023

Templier, M., & Paré, G. (2015). A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems, 37*, 112–137. https://doi.org/10.17705/1cais.03706.

Thomas, K. M., & Nelson, C. A. (2001). Serial reaction time learning in preschool- and school-age children. *Journal of Experimental Child Psychology, 79*(4), 364–387. https://doi.org/10.1006/jecp.2000.2613

Thompson, R. F., & Kim, J. J. (1996). Memory systems in the brain and localization of a memory. *Proceedings of the National Academy of Sciences of the United States of America, 93*(24), 13438–13444. https://doi.org/10.1073/pnas.93.24.13438.

Toner, C. K., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2012). Visual object pattern separation deficits vary in cognitively normal older adults. *Learning & Memory, 42*, 338–342. https://doi.org/10.1101/lm.1315109.subsequent

Tulving, E. (1972). *Episodic and semantic memory. In Organization of memory*. Oxford, England: Academic Press.

Underwood, B. J. (1969). *Attributes of memory. Psycholgoical Review, 76*(6), 559–573. https://doi.org/10.4324/9780203467084

Unger, L., Fisher, A. V., Nugent, R., Ventura, S. L., & MacLellan, C. J. (2016). Developmental changes in semantic knowledge organization. *Journal of Experimental Child Psychology, 146*, 202–222. https://doi.org/10.1016/j.jecp.2016.01.005

Unger, L., Savic, O., & Sloutsky, V. M. (2020). Statistical regularities shape semantic organization throughout development. *Cognition, 198*, Article 104190. https://doi.org/10.1016/j.cognition.2020.104190

Wagner, G., Prester, J., Roche, M. P., Schryen, G., Benlian, A., Paré, G., & Templier, M. (2021). Which factors affect the scientific impact of review papers in IS research? A scientometric study. *Information and Management, 58*(3). https://doi.org/10.1016/j.im.2021.103427

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems Quarterly*, *26*(2), xiii–xxiii. https://doi.org/10.1016/j.freeradbiomed.2005.02.032.

Wu, Y., & Jobson, L. (2019). Maternal reminiscing and child autobiographical memory elaboration: A meta-analytic review. *Developmental Psychology, 55*(12), 2505–2521. https://doi.org/10.1037/dev0000821

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences, 34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Zeithamova, D., & Bowman, C. R. (2020). Generalization and the hippocampus: More than one story? *Neurobiology of Learning and Memory, 175*, Article 107317. https://doi.org/10.1016/j.nlm.2020.107317