Cognition 239 (2023) 105564

Contents lists available at ScienceDirect

Cognition

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

Separate but not independent: Behavioral pattern separation and statistical learning are differentially affected by aging

Helena Shizhe Wang [a], Stefan Köhler [a,b,c,1], Laura J. Batterink [a,b,1,*]

[a] Western Institute for Neuroscience, University of Western Ontario, London, Ontario, Canada
[b] Department of Psychology, University of Western Ontario, London, Ontario, Canada
[c] Rotman Research Institute, Baycrest, Toronto, Ontario, Canada

ARTICLE INFO

Keywords:
Pattern separation
Statistical learning
Cognitive aging
hippocampus

ABSTRACT

Our brains are capable of discriminating similar inputs (pattern separation) and rapidly generalizing across inputs (statistical learning). Are these two processes dissociable in behavior? Here, we asked whether cognitive aging affects them in a differential or parallel manner. Older and younger adults were tested on their ability to discriminate between similar trisyllabic words and to extract trisyllabic words embedded in a continuous speech stream. Older adults demonstrated intact statistical learning on an implicit, reaction time-based measure and an explicit, familiarity-based measure of learning. However, they performed poorly in discriminating similar items presented in isolation, both for episodically-encoded items and for statistically-learned regularities. These results indicate that pattern separation and statistical learning are dissociable and differentially affected by aging. The acquisition of implicit representations of statistical regularities operates robustly into old age, whereas pattern separation influences the expression of statistical learning with high representational fidelity and is subject to age-related decline.

1. Introduction

Our daily life consists of many repeated activities, such as eating breakfast and commuting to work. To function effectively, we need to be able to discriminate between highly overlapping but ultimately distinct events *and* to generalize across overlapping events. For example, we must remember where we parked our car this morning, rather than the day before, while also being able to generalize across multiple days in order to appreciate that one part of the lot regularly tends to have more empty spots than another. Although most commonly discussed in the context of episodic memory, discrimination and generalization also play crucial roles in the context of language acquisition. For example, children who learn their native language and adults who learn a second language need to distinguish the sound of a new word (e.g., “write”) from previously encountered words (“white”) that sound similar but have distinct meanings. By generalizing across their multiple encounters with these two words, they also learn that the sound “ing” is more likely to follow “write” than “white”.

These examples build on processes of pattern separation and statistical learning, respectively. Pattern separation refers to the creation of discrete orthogonalized representations of similar inputs. It allows us to distinguish unique events with many overlapping features in memory. On the other hand, statistical learning is the more gradual process of becoming sensitive to similarities shared across multiple events, supporting the extraction of statistical regularities across inputs over time. One critical difference between these two types of learning that has been highlighted in the literature is that pattern separation allows for learning based on single snapshots (one-shot learning), whereas statistical learning requires the extraction of repeated structure over extended time periods (Schapiro, Turk-Browne, Botvinick, & Norman, 2017).

While it is widely acknowledged that pattern separation and statistical learning are conceptually different processes, whether they operate independently remains largely unknown. A recently proposed computational model suggests that each process relies on computations that are implemented by distinct neural pathways, suggesting separation at the neural level (Schapiro et al., 2017). Specifically, the model hypothesizes that pattern separation and statistical learning are performed by

* Corresponding author at: Department of Psychology, University of Western Ontario, Western Interdisciplinary Research Building, London, Ontario N6A 3K7, Canada.
E-mail address: lbatter@uwo.ca (L.J. Batterink).

[1] Indicates equal contribution

https://doi.org/10.1016/j.cognition.2023.105564
Received 25 August 2022; Received in revised form 23 June 2023; Accepted 11 July 2023
Available online 18 July 2023

separate pathways within the hippocampus, the trisynaptic pathway and the monosynaptic pathway, respectively. The trisynaptic pathway projects from the entorhinal cortex to hippocampal subregion CA1 through the dentate gyrus and CA3, two subregions that have been consistently implicated in pattern separation (e.g., Berron et al., 2016; Leutgeb, Leutgeb, Moser, & Moser, 2007). In contrast, the monosynaptic pathway projects directly from the entorhinal cortex to CA1 and is considered to support statistical learning. A neural network model that simulated these hippocampal subregions demonstrated pattern separated representations in the dentate gyrus and CA3 and representations that reflected statistical learning in CA1, supporting the hypothesis that the trisynaptic and monosynaptic pathways support pattern separation and statistical learning, respectively (Schapiro et al., 2017).

Behavioral research that speaks to the proposed independence of these computational processes is limited at present, as research on pattern separation and statistical learning has largely proceeded independently, and has made use of different tasks and stimulus materials (e.g., Frost, Armstrong, & Christiansen, 2019; Saffran, Aslin, & Newport, 1996; Yassa & Stark, 2011). However, we recently developed a set of tasks for direct comparison with speech-based stimuli. In neuropsychological research, we showed that an individual with circumscribed bilateral dentate gyrus lesions (affecting the trisynaptic hippocampal pathway) demonstrated a deficit in pattern separation alongside intact expression of statistical learning, suggesting a dissociation between these two processes (Wang et al., 2023). Still, it remains unclear whether these two processes can also be dissociated behaviorally in individuals without brain lesions. Of particular theoretical importance in any such comparison is also the manner in which knowledge about statistical regularities in speech sounds is probed. In our previous neuropsychological study, we found evidence for preservation of statistical learning only when task performance could be supported by the implicit expression of knowledge (through reaction times) or by a coarse, gist-based familiarity signal (through a forced-choice recognition-memory task). In contrast, the explicit retrieval of high-resolution statistical representations was found to be impaired. Against this background, dependencies between pattern separation and statistical learning may emerge specifically under conditions that share high-precision explicit retrieval demands.

The goal of the current study was to further test whether pattern separation and statistical learning are dissociable processes within the brain. To this end, we focused on age as a naturally-occurring between-subjects manipulation to compare the behavioral expressions of these two processes in healthy individuals, using highly similar stimulus materials across tasks and probing knowledge of statistical regularities in multiple test formats. Based on the proposed computational model (Schapiro et al., 2017), we hypothesized that healthy aging would differentially impact pattern separation and statistical learning, with greater age-related effects for pattern separation compared to statistical learning as expressed under conditions that do not require explicit access to high-precision knowledge. The rationale for this hypothesis is outlined further below, as we next review what is currently known about aging effects within each of these domains.

2. Effects of aging on pattern separation and statistical learning

Pattern separation has often been studied in humans using the Mnemonic Similarity Task (MST; Stark, Yassa, Lacy, & Stark, 2013; Stark, Stevenson, Wu, Rutledge, & Stark, 2015). A typical MST begins with incidental encoding of a list of visual objects, followed by a recognition phase in which participants classify studied items, perceptually similar lures, and novel foils as “old”, “similar”, or “new”. The ability to successfully differentiate old items from similar lures has been consistently associated with the dentate gyrus and CA3 (Baker et al., 2016; Bakker, Kirwan, Miller, & Stark, 2008; Lacy, Yassa, Stark, Muftuler, & Stark, 2011; Yassa et al., 2010), supporting the use of the MST as a reliable behavioral index of pattern separation. A number of studies using the visual MST have found that, relative to younger adults, older adults frequently report similar lures as previously seen, indicative of a deficit in pattern separation (Kirwan & Stark, 2007; Yassa, Lacy et al., 2011). This age-related deficit has been documented across multiple task formats, including versions of the MST that have separate study and test phases or a continuous task design and versions with three-choice (“old”, “similar”, “new”) or two-choice response formats (“old”, “new”; Stark et al., 2015). In addition, age effects have been reported across various types of visual stimuli such as objects and scenes (Stark & Stark, 2017), emotional stimuli (Leal & Yassa, 2014), and written words (Ly, Murray, & Yassa, 2013).

While pattern separation is typically studied in the visual modality, statistical learning was first investigated in the context of speech segmentation (Saffran et al., 1996), and continues to be most frequently studied using auditory linguistic stimuli (although statistical learning studies in the visual domain are also becoming increasingly common, e.g., Kirkham, Slemmer, & Johnson, 2002; Turk-Browne, Jungé, & Scholl, 2005). In a typical linguistic statistical learning task in adults, participants are initially exposed to a continuous speech stream consisting of repeating trisyllabic “words” (e.g., “babupupatubitutibu…”; Saffran et al., 1996). Participants are then asked to explicitly discriminate between “words” and foils, composed of recombined syllables from the stream. Above-chance performance on this forced-choice recognition test has been the standard way to indicate the occurrence of statistical learning. Critically, recent studies have provided evidence that the knowledge of the learned regularities acquired through statistical learning can be expressed both implicitly and explicitly (Arciuli, 2017; Batterink, Reber, Neville, & Paller, 2015; Bertels, Franco, & Destrebecqz, 2012), such that the typical, explicit forced-choice recognition task may not adequately capture all aspects of statistical learning. To address this issue, many studies have also employed implicit measures of the knowledge gained through statistical learning that do not require explicit, intentional retrieval of the learned statistical regularities. One such measure is the target detection task, which measures participants' reaction time as they respond to a target syllable in a continuous speech stream (Batterink et al., 2015; Kim, Seitz, Feenstra, & Shams, 2009; Siegelman, Bogaerts, & Frost, 2017). Faster reaction time to the more predictable syllables (i.e., the second or third syllable in a triplet) is taken as evidence of statistical learning.

Statistical learning has been demonstrated to occur robustly from infancy to young adulthood (e.g., Raviv & Arnon, 2017; Saffran et al., 1996; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009). However, compared to pattern separation, studies on statistical learning in the healthy aging population are scarce, and the existing literature provides more mixed findings. For example, older adults performed similarly to younger adults on a reaction time-based measure of statistical learning in the visual modality (Campbell, Zimerman, Healey, Lee, & Hasher, 2012; Ong & Chan, 2019) and on forced-choice recognition tasks for trisyllabic words (Shaqiri, Danckert, Burnett, & Anderson, 2018). In contrast, other studies with forced-choice recognition tasks have observed poorer statistical learning in older adults in both auditory and visual modalities (Palmer, Hutson, & Mattys, 2018; Schevenels, Altvater-Mackensen, Zink, De Smedt, & Vandermosten, 2021). Similarly, using a novel continuous statistical learning paradigm, Herff, Zhen, Yu, and Agres (2020) observed differences in learning trajectories between older and younger adults. As a possible explanation for the observed age differences, Palmer et al. (2018) suggested that statistical learning, while relying on an implicit component that operates automatically, can be further boosted by an explicit component, which relies on working memory and is subject to age-related decline. An additional, not mutually exclusive possibility is that age differences arise due to other task demands related to how learning is measured, rather than in the ability to extract statistical regularities itself. Explicit, direct measures of learning inherently rely on peripheral processes such as explicit memory retrieval (e.g., Christiansen, 2019), which are themselves susceptible to aging (Korkki, Richter, Jeyarathnarajah, & Simons, 2020) and may influence the assessment of the impact of aging on statistical learning. Given that previous aging studies of statistical learning have mostly relied on explicit measures of learning, it remains a possibility that implicit expression of statistical learning is preserved.

Whether aging affects other types of memory that involve incidental learning of regularities over time has been inconsistent. For example, motor sequence learning, measured as a decrease in reaction time over time to a structured sequence of key presses, has been reported to be age-resistant (Gaillard, Destrebecqz, Michiels, & Cleeremans, 2009; Schevenels et al., 2021), but age-related deficits emerge when sequence structure becomes more complex (Howard & Howard, 2013; Janacsek, Fiser, & Nemeth, 2012) or when behavioral expression of learning is probed with a confidence rating-based recognition task (Gaillard et al., 2009). Similarly, in artificial grammar learning, older adults have demonstrated preserved abilities to differentiate grammatical from ungrammatical sentences in some studies (Neger, Rietveld, & Janse, 2015) but not in others (Lukács & Kemény, 2015; Neger, Rietveld, & Janse, 2014). Importantly, using multiple measures of learning, one study observed age-related declines in motor sequence learning on tasks requiring explicit judgments, along with intact performance on implicit measures (Gaillard et al., 2009). This latter finding suggests that implicit behavioral expression of statistical learning may be relatively resistant to age-related decline. It also highlights the importance of using both implicit and explicit measures to understand the effect of aging on statistical learning.

3. The current study

In the current study, we investigated pattern separation and statistical learning in older and younger adults, to test whether they are distinct computational processes with dissociable effects on behavior. Our primary hypothesis was that pattern separation and the implicit expression of statistical learning are dissociable. A related, secondary hypothesis was that participants' ability to explicitly retrieve precise representations of their statistical knowledge would be related to their pattern separation abilities. To test these dissociations and associations, we focused on effects of aging.

One challenge involved in directly comparing these two processes is that they have typically been studied using entirely different paradigms and stimulus materials. While statistical learning tasks often use auditory linguistic stimuli, pattern separation is usually assessed with visual stimuli, such as common objects, scenes, and abstract pictures. To address this challenge, we employed a novel auditory-linguistic version of the widely employed Mnemonic Similarity Task that measures participants' ability to discriminate spoken nonsense trisyllabic words, which ensured that the learning materials for the pattern separation and statistical learning tasks were highly similar. Furthermore, we used three separate measures of statistical learning to fully capture both implicit and explicit expressions of knowledge accrued during learning: (1) an explicit rating task, in which participants rated their familiarity with the original "words" and two types of foil items (highly similar "partwords" and less similar "nonwords") presented in isolation; (2) the explicit standard forced-choice recognition task, which requires discrimination of the learned items from nonword foils presented in direct opposition; (3) the implicit target detection task, which requires participants to make speeded responses to embedded syllables within continuous speech streams. While the rating task and recognition task both require participants to make conscious reference to previously learned information, the target detection task does not require conscious or intentional retrieval but probes statistical learning indirectly based on response times. Notably, these three tasks also differ in their reliance on pattern separation. While performance on the rating task relies on the ability to maintain highly precise representations of words in long-term memory, the 2AFC task can be solved by comparing gist-based familiarity for target words and novel nonwords (Bastin & Van der Linden, 2003; Holdstock et al., 2002), in the absence of highly precise knowledge for the words. Finally, the target detection task is sensitive to participants' ability to predict syllables within a word, and facilitation on this task also does not require a high precision memory trace.

With these task characteristics in mind, we predicted that older adults would show specific deficits in distinguishing old versus highly similar items on both the MST and on the explicit rating measure of statistical learning, and that performance on these two tasks would be correlated across participants. In contrast, we expected that performance on the implicit, reaction time-based statistical learning task and on the forced-choice recognition measure would be preserved across age, given their relative independence on pattern separation mechanisms.

4. Methods

4.1. Participants

A total of 94 younger adults and 137 older adults were recruited anonymously via Prolific.co, a crowdsourcing platform for recruiting and managing participants for online studies. Participants between ages 18 and 30 were recruited as younger adults and participants between ages 60 and 89 were recruited as older adults. All participants were required to be native English speakers (though not necessarily monolingual), to have no hearing difficulties, and to have no history of neurological or psychiatric disorders (all determined via self-report). Informed consent was obtained from all participants in compliance with the Research Ethics Board of the University of Western Ontario.

Participants were excluded from the final analysis for failing to complete the study ($n = 28$), for failing the attention check during the statistical learning task ($n = 3$), for failing the headphone check ($n = 5$), or for scoring below the normal hearing threshold on the hearing assessment ($n = 73$; 14 younger adults, 59 older adults). Participants whose performance score on either the MST or the statistical learning tasks was above or below the overall mean of their respective age by more than two standard deviations were considered outliers and removed from the final sample. Four younger adults and five older adults were excluded based on this criterion. After exclusions, the final sample consisted of 59 younger adults between ages 18 and 30 (38 females and 21 males, average age = 21.9, average years of education = 14.3) and 61 older adults between ages 60 and 86 (41 females and 20 males, average age = 64.7, average years of education = 15.0). The two groups did not significantly differ in the years of education completed ($t(108) = 1.64$, $p = .10$). We aimed to obtain a sample size of approximately 60 participants per group after exclusions, which represents a larger sample size than most previous aging studies in the fields of statistical learning and pattern separation (e.g., Campbell et al., 2012: $n = 24$ per group; Ong & Chan, 2019: $n = 20$ per group; Palmer et al., 2018: $n = 24$ per group; Stark & Stark, 2017: $n = 26$–28 per group).

4.2. Stimuli

4.2.1. Word mnemonic similarity task

The stimuli for this task consisted of a set of 25 unique trisyllabic nonsense words (e.g., *golapu*), created from 75 unique syllables. These 25 words were presented in the task as "First presentation" items. Of this set, five words were each repeated 10 additional times to create "Repeat" items and another five words were used to create an additional 25 similar "Lure" words by recombining the syllables in each word in five different ways (e.g., *gopula*: *golapu*, *lagopu*, *lapugo*, *pugola*, *pulago*). The remaining 15 First presentation items served as foils and were never repeated. Thus, in total, there were 25 First presentation trials, 25 Lure trials, and 50 Repeat trials, resulting in a total of 100 trials (Fig. 1).

To ensure that each trisyllabic word item was as distinct as possible, the three syllables within each word were created by randomly pairing three different consonants with three different vowels, with the

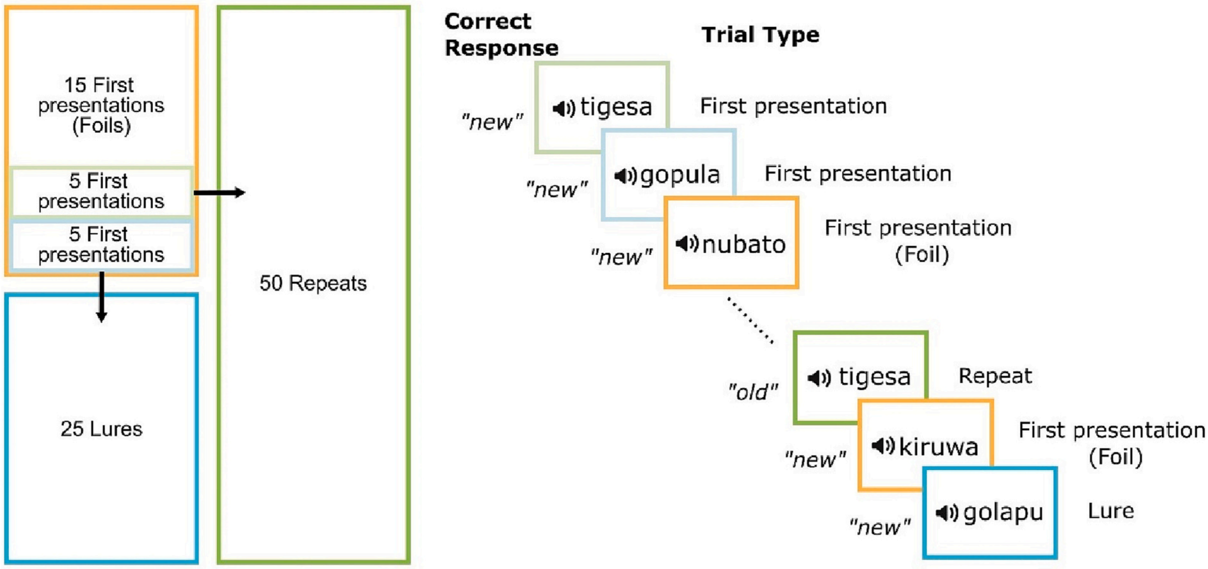

Fig. 1. Diagram of Word Mnemonic Similarity Task. Of the initial 25 unique word items (orange), five were used to create 50 Repeat trials (green) and another five were used to create 25 Lure trials (blue). Participants were instructed to respond "Old" to repeated items and "New" to items that were presented for the first time even if similar in sound to previously encountered items. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

constraint that the syllables used in Lure and Repeat items were never used in another word.

To create the auditory stimuli for this task, Microsoft Word's "Read Aloud" function was used to produce individual syllables which were recorded and combined into words with Audacity. Each word was 1 s in duration (range: 0.95 s – 1.16 s) and the perceived loudness was normalized using Audacity to 12 LUFS (Loudness Unit relative to Full Scale) across all sound files.

4.2.2. Statistical learning

The stimuli for the statistical learning tasks consisted of 12 unique syllables, taken from Batterink and Paller (2019), which were combined to form four trisyllabic nonsense words (*tafuko, regeme, rupuni, fetisu; see Fig. 2*). Of the 12 syllables, six (*ge, pu, fu, ti, su, fe*) also appeared in the words used in the MST, although a different voice was used to generate the syllables for this task. To control for potential syllable-specific effects, the syllables in a given word were each assigned to the first, second, and third position across three counterbalance conditions (e.g., *tafuko, fukota, kotafu*). Participants were randomly assigned to one of the three conditions at the beginning of the task. The sound file for each syllable was 300 ms in duration.

4.2.2.1. Stimuli for exposure phase. To create the continuous speech streams used in the initial Exposure Phase, each word was concatenated in pseudorandom order with the constraint that the same word never appeared consecutively, at a rate of 380 ms per syllable. Each word was repeated a total of 90 times, resulting in a 6.84 min long continuous stream.

4.2.2.2. Stimuli for the target detection task. A total of 36 speech streams were created. Each stream consisted of a total of 48 syllables (16 words),

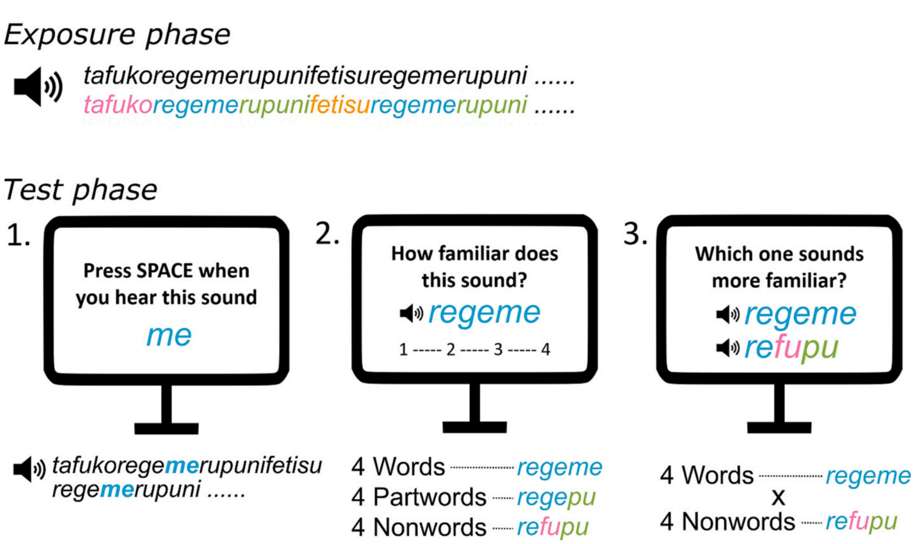

Fig. 2. Diagram of statistical learning tasks. Following a 6 min exposure phase, participants underwent 1) Target Detection task, 2) Rating task, and 3) 2AFC Recognition task. Partwords used in Rating task were created by combining two syllables from one word and one syllable from another word. Nonwords used in Rating and 2AFC Recognition tasks consisted of syllables from three different words.

concatenated together in the same manner as in the Exposure Phase and presented at the same rate. Each speech stream was approximately 18 s long.

4.2.2.3. Stimuli for the explicit SL tasks. The stimuli for the Rating Task consisted of 12 trisyllabic items: four Words from the Exposure Phase, four similar Partwords that contained two syllables from the same word and one syllable from another word (1. rege + ko, 2. feti + me, 3. ta + puni, 4. ru + tisu), and four Nonwords that contained syllables from three different words (1. pu + ge + ti, 2. ni + su + ta, 3. fu + ru + me, 4. ko + re + fe). None of the Partwords appeared across word boundaries during the exposure phase. Both Partwords and Nonwords were used to probe participants' recognition of the specific syllable combinations that occurred in the Exposure Phase. The same four Words and four Nonwords were used as the stimuli for the 2AFC Recognition Task.

4.3. Experimental tasks

4.3.1. Word MST

This task required participants to make "Old" and "New" recognition judgments to items presented in a continuous stream without a separate study phase (see Figure 1). It was modeled after the continuous version of the Mnemonic Similarity Task (Stark et al., 2015; Yassa et al., 2010). It should be noted that while the overt encoding on this continuous version differs from the typical MST with a separate incidental encoding phase, this difference does not affect the nature of the task. Each trial began with a 1.5 s pause followed by the auditory presentation of a word item along with a prompt on the screen ("New or Old?") and ended when participants made a keyboard response. Participants were instructed to label a word as "Old" if the word had been presented before or "New" if the word had never been presented before. They were also specifically instructed that some of the words would sound similar to one another, but that similar-sounding words should also be labelled as "New" if they had not been previously presented in exactly that form. Before the task started, participants underwent five practice trials during which they were given feedback on whether their responses were correct or incorrect. Participants were explicitly informed that the items used for the practice trials would not appear in the main task. During the main task, the items were presented in such a manner that each Repeat or Lure item was separated by an average of six intervening items (range = 2–12 items; Fig. 1). To control for potential item order effects, six counterbalance conditions were created and each participant was randomly assigned to one of the six conditions.

4.3.2. SL

4.3.2.1. Exposure phase. Participants listened to a 6.84-min speech stream, which was divided into three 2.28-min blocks. At the end of each block, participants were asked to guess the total number of unique syllables used in the speech stream and were then given an optional break (maximum 30 s). While listening, participants also performed a cover task in which they responded to pauses within the speech with a keypress. A total of 18 short pauses were inserted into the speech stream and the number of hits to the pauses were used to confirm that participants were continuously listening to the stream. The timing of the pauses was pseudo-random with the constraint that they always occurred after the second syllable in a word, so as not to indicate word boundaries. Those who missed more than one pause were excluded from the final sample ($n = 3$).

4.3.2.2. Target detection task. After the Exposure Phase, participants completed the target detection task, designed to assess participants' implicit knowledge of the statistics of the speech stream. This task requires participants to make speeded responses to target syllables embedded in short segments of the continuous speech stream. At the beginning of each trial, a written form of the target syllable (e.g., "ta") was displayed on the screen while the auditory syllable was presented twice. The written form of the syllable then remained on the screen while the short speech stream was presented. Participants were required to make a keypress each time they detected the target syllable. Both speed and accuracy were emphasized. Each stream consisted of all four words in the language presented four times each, concatenated in the same manner as the original exposure speech stream but with the constraint that the word containing the target syllable never appeared as the first or the last word in the stream.

In order to examine possible learning effects during the target detection task itself, the task was subdivided into three blocks of 12 streams, with each of the 12 syllables serving as the target syllable once per block. Within each block, the 12 syllables were ordered in such a way that the syllable positions of the targets (word-initial, word-middle, word-final) were evenly distributed across the block (e.g., *initial, middle, final, middle, initial, final*…). To monitor participants' attention and engagement during the target detection task, a short-answer question that was irrelevant to the task (e.g., "What is your favorite color?") was inserted at the end of each block and participants were required to type in their answers.

Prior to starting the task, participants completed three practice trials. A different speech stream made up of three trisyllabic words was used for the practice. The words in the practice speech stream contained none of the 12 syllables in the exposure speech stream and were generated using a different speech synthesizer voice. At the end of each practice trial, participants were given their average response time and their total number of hits.

4.3.2.3. Rating task. This task was designed to assess participants' recognition of the words encountered during the Exposure Phase. On each trial, participants listened to a trisyllabic item and rated their familiarity with the item on a scale of 1–4 (1 = least familiar). The task consisted of 12 trials (four Words, four Partwords, and four Nonwords), presented in a random order.

4.3.2.4. Two-alternative forced-choice recognition task. This task served as an additional measure of participants' explicit knowledge. On each trial, participants listened to a Word and Nonword pair, and selected the one that sounded more familiar to them. The numbers "1" and "2" appeared on the screen for1s before each item was presented, and participants responded with the number associated with the more familiar sounding item. The same four Nonwords used in the Rating task were paired in all possible combinations with the four Words, resulting in a total of 16 trials. The Word appeared as the first item in half of the trials, and trials were presented in random order.

All participants completed the statistical learning tasks in the following order: Exposure phase, Target Detection task, Rating task, and 2AFC task (see Figure 2). This order was chosen so that the participants could complete the Target Detection task prior to being exposed to the isolated triplets in the Rating and the 2AFC tasks, which could increase their explicit knowledge of the triplets. In addition, the Rating task was administered before the 2AFC because the 2AFC task provided additional exposures to Words and Nonwords but not Partwords. All tasks were created using PsychoPy 2020.2.10 and were hosted and presented online via Pavlovia.

4.4. Procedure

The Word MST and the SL tasks were administered in two separate sessions, separated by approximately 24 h, in order to minimize task interference and fatigue caused by completing a prolonged online study. The Word MST took approximately 15 min, and the SL tasks took approximately 35 min to complete. The order of the two tasks was counterbalanced across participants by randomly assigning them to one of the tasks in the first session. Twenty-four hours after completing the first session, participants were sent an invitation link through Prolific to sign up for the task that they were not assigned to in the first session. Participants were compensated through Prolific after completion of each session. Information on this two-session study structure was given to participants at the beginning of the first session. Those who completed the first session but did not return for the second session were removed from the final sample ($n = 28$).

All tasks were performed online with the participants' own laptops or personal computers. To minimize distractions during the study, participants were instructed to complete the tasks in a quiet listening environment and to use headphones or earphones for the entire duration of the session. Each session began with a volume adjustment task during which participants listened to music and adjusted their sound volume to a comfortable level. Participants then completed a headphone check task (Woods, Siegel, Traer, & McDermott, 2017). Individuals who failed on more than one trial on the headphone check task were removed from the final sample ($n = 5$).

To ensure participants displayed normal hearing, this study also included an online hearing assessment. After completing the first session, participants were automatically directed to a Qualtrics survey containing demographic questions as well as instructions and a link to a hearing check website, SHOEBOX Online (shoeboxonline.com). SHOEBOX Online is a web-based hearing assessment that has been validated for sensitivity to hearing losses against in-person hearing assessment (Reed et al., 2022). It consists of a questionnaire about one's hearing experience and a pure tone audiometry test and uses the results to categorize ears into three categories, "Normal hearing", "Loss", and "Significant loss". Those whose hearing thresholds at 1000, 2000, and 4000 Hz were equal to or lower than 25 dB were considered to have normal hearing. Participants were instructed to return to the Qualtrics survey after completing the hearing assessment and select the category assigned to each of their ears from a drop-down menu. Of the initial sample of 231 people, 3 (1 younger adults, 2 older adults) reported "Significant loss" for both ears, 70 (13 younger adults, 57 older adults) reported "Loss" for at least one ear, and 142 (70 younger adults, 72 older adults) reported "Normal hearing" for both ears. Only those who completed the test and reported normal hearing in both ears were included in the final sample.

4.5. Data analyses

4.5.1. Word MST

Participants' response data were entered into a 2 × 3 mixed effects ANOVA with age group (young, old) as the between-subject factor and item condition (Repeat, Lure, First presentation) as the within-subject factor. Furthermore, d' estimates derived from signal detection theory (z(Hit rate) – z(False Alarm rate)) were used to index memory discrimination while correcting for response bias (Stark et al., 2015; Leal, Tighe, & Yassa, 2014; Yassa et al., 2011). All response rates were transformed to z-scores for d' calculations. Lure Discrimination (p("New"|Lure) – p("New"|Repeat)) was computed as the probability of correctly responding "New" to Lure items minus the tendency to falsely label a repeated item as "New", and was used as the behavioral measure for pattern separation. Recognition (p("New"|First presentation) – p("New"|Repeat)) was computed as the probability of correctly responding "New" to First presentation items minus the tendency to falsely label a repeated item as "New", and was used as the behavioral measure for item recognition. Response bias was indexed with c' which is defined as criterion c (z(Hit rate) + z(False Alarm rate)/−2) divided by discrimination (d'). A positive c' indicates a general tendency for responding "Old" and a negative c' indicates a general tendency for responding "New". C' was computed separately for Lure trials and First presentation trials.

4.5.2. SL

4.5.2.1. Target detection task. Responses made within 0–1200 ms of a target syllable onset were considered as "hits" and their RTs were used for analyses. All other responses were considered false alarms. RTs were entered into a mixed effects model with age group, syllable position (word-initial, word-middle, word-final), the position of the target within each stream (4–45), and stream number (1–36) as fixed effects and participants as a random effect. The position of the target within a stream and stream number were included as covariates to explore any learning that might have occurred within each stream (target position) and across streams (stream number) due to the additional exposure to the speech stream. Planned contrasts were used to test whether RTs decreased linearly as a function of syllable position. In addition, the effect of age and syllable position on RT was also assessed by a Bayesian mixed ANOVA.

To quantify SL performance using a single metric while controlling for individual baseline response times, an "RT prediction score" was computed by subtracting the average RT for the final syllable position from the average RT for the initial syllable position and dividing it by the average RT for the initial syllable position (Batterink & Paller, 2019). An independent samples t-test was used to compare RT prediction scores between younger and older adults. An age-related difference in statistical learning would be revealed by a significant difference in RT prediction score between the two groups.

4.5.2.2. Rating task. Average familiarity rating was computed for each word category (Word, Partword, Nonword) and was entered into a 2 × 3 mixed effects ANOVA with age group as the between-subjects factor and word category as the within-subjects factor. Additionally, a "Word-Partword (W-PW) score" was computed by taking the difference between the average ratings for Words and Partwords and a "Word-Nonword (W-NW) score" was computed by taking the difference between the average ratings for Words and Nonwords. We also computed a "Partword-Nonword (PW-NW) score" by taking the difference between the average ratings for Partwords and Nonwords to examine whether there was a difference in response patterns to the two types of foil words between younger and older adults.

4.5.2.3. 2AFC recognition task. An independent samples t-test was conducted to compare the average accuracy in younger and older adults across all trials. Since we hypothesized that there would be no difference in accuracy between the two age groups, a Bayesian analysis was also conducted to determine whether there was evidence supporting the lack of an age effect on accuracy.

4.5.2.4. Correlational analysis. To examine the relationship between performance on different measures of statistical learning, we conducted exploratory correlational analyses between RT score, W-PW score, W-NW score, and accuracy on the 2AFC task while correcting for multiple comparisons using the Bonferroni procedure.

4.5.3. Mediation analysis

To examine whether the age effects on discrimination of similar words in the Word MST and the Rating task were related, a mediation analysis was conducted with W-PW score as the dependent variable, age as the independent variable, and Lure Discrimination as a possible mediator (Preacher & Hayes, 2008). Bootstrapping with 1000 replications was used to calculate bias-corrected standard errors of the indirect paths. The same procedure was used to investigate the relationship between age and W-NW score with Lure Discrimination as a possible mediator.

5. Results

All variables were confirmed to meet the normality and homoscedasticity assumptions of ANOVA with the Shapiro test and the Bartlett test.

5.1. Word MST

Accuracy rates differed as a function of item condition ($F(2, 354) = 151.37$, $p < .001$, $\eta^2 = 0.46$), with accuracy being the highest for First presentation items and lowest for Lure items (Fig. 3a). On average, participants erroneously responded "Old" to 40.2% (OA: 42.5%, YA: 37.8%) of Lures and 13.1% (OA: 16.5%, YA: 9.6%) of First presentation items. This is a pattern that is consistent with that reported in other studies that employed the MST in the visual modality (e.g., Baker et al., 2016; Stark et al., 2015). Across all conditions, younger adults performed more accurately than older adults ($F(1, 354) = 5.79$, $p = .017$, $\eta^2 = 0.02$) but no age by condition interaction was identified ($F(2, 354) = 0.20$, $p = .82$).

Critically, when signal-detection analyses were conducted using *d'*, there was a significant age effect on Lure Discrimination, such that younger adults showed greater sensitivity in discriminating similar Lure from Repeat items than older adults ($t(118) = -2.52$, $p = .013$, $d = -0.46$). No age difference was found in Recognition discrimination scores between First presentation and Repeat items ($t(118) = -1.05$, $p = .30$; Fig. 3b). However, when *d'* scores were entered into a mixed effects ANOVA with age and memory measure (Lure Discrimination, Recognition) each as between- and within-subjects factors, there was not a significant interaction between age and memory measure ($F(1, 118) = 0.13$, $p = .72$, $\eta^2 < 0.001$); instead older adults showed numerically lower scores on both measures as indicated by a marginal main effect of age ($F(1, 118) = 3.45$, $p = .066$, $\eta^2 = 0.03$). It should be noted, however, that these ANOVA results need to be interpreted with caution as both Recognition and Lure Discrimination depend on participants' false alarm rate on the Repeat trials and are thus not entirely independent.

Both older and younger adults showed a tendency to respond "Old" to Lure trials (C': OA = 0.35, YA = 0.30) and "New" to First presentation trials (C': OA = −0.10, YA = −0.095). There was no significant difference between the two groups in *c'* for Lure trials ($t(118) = 0.62$, $p = .53$) or First presentation trials ($t(118) = -0.27$, $p = .78$), suggesting that the observed age differences are at the level of discrimination rather than response criteria.

5.2. Statistical learning

5.2.1. Target detection task

The average hit rate in the Target Detection task was 88.1% (YA = 87.6%, OA = 88.6%). Across both groups, RTs showed the expected decrease as a function of syllable position ($F(2, 15{,}100) = 88.11$, $p < .001$, $\eta^2 = 0.32$; linear effect: $F(1, 15{,}101) = 173.13$, $p < .001$). Interestingly, there was no interaction between age group and syllable position ($F(2, 15{,}100) = 0.048$, $p = .95$), indicating that the effect of syllable position on RTs was not modulated by age group (Fig. 4a). This lack of an age-position interaction was also supported by a follow-up Bayesian mixed ANOVA indicating extreme evidence in favor of an absence of an interaction effect ($BF_{10} = 614.5$). There was also no age difference in the RT prediction scores ($t(113) = 0.28$, $p = .78$).

Additionally, RTs also increased significantly both as a function of target position ($F(1, 15{,}096) = 69.73$, $p < .001$) and stream number ($F(1, 15{,}096) = 51.18$, $p < .001$) such that RTs became slower toward later occurring targets within a stream as well as for later streams within the task. Both target position and stream number also significantly interacted with syllable position ($F(2, 15{,}096) = 7.99$, $p < .001$; $F(2, 15{,}096) = 10.73$, $p < .001$), such that the facilitation effect was stronger for later occurring targets and later trials. Age did not moderate the effect of target position ($F(1, 15{,}096) = 0.88$, $p = .35$) or stream ($F(1, 15{,}096) = 0.55$, $p = .46$), indicating that RTs increased for later targets and streams for both younger and older adults. The effects of block and stream position on statistical learning are consistent with findings in previous studies (Batterink et al., 2015; Batterink & Paller, 2017), and could be due to a nonspecific effect of fatigue counteracted by the facilitation effect of statistical learning on latter syllables, or the greater cost incurred by the anticipation of the upcoming syllables on the initial syllable (Sherman & Turk-Browne, 2020; Turk-Browne, Scholl, Johnson, & Chun, 2010).

5.2.2. Rating task

Across both groups, familiarity ratings varied significantly by word category ($F(2, 236) = 66.77$, $p < .001$, $\eta^2 = 0.84$), such that ratings were highest for previously encountered Words, followed by Partwords, and then Nonwords (Fig. 4b). This pattern is consistent with the results reported in prior statistical learning studies that employed the Rating task (Batterink & Paller, 2017, 2019). There was no main effect of age group ($F(1, 118) = 2.32$, $p = .13$). There was a marginally significant interaction between word type and age group ($F(2, 236) = 2.44$, $p = .089$).

We then tested for an age effect on W-PW scores and W-NW scores individually to probe for patterns reflected in the marginally significant interaction. There was a significant age effect in the W-PW scores ($F(1, 118) = 4.94$, $p = .028$, $\eta^2 = 0.04$), such that older adults showed reduced sensitivity to familiarity between Words and Partwords than younger adults. No age difference was found in the W-NW scores ($F(1, 118) = 2.16$, $p = .14$; Fig. 4c). When both W-PW and W-NW scores were analyzed with a mixed effects ANOVA with age and score type (W-PW, W-NW) each as between- and within-subject factors, there was a significant main effect of age with older adults scoring lower than younger adults ($F(1, 118) = 4.27$, $p = .041$, $\eta^2 = 0.03$), but there was not an interaction between age and score type ($F(1, 118) = 0.32$, $p = .57$). The two groups also did not differ in the PW-NW scores ($t(118) = 0.57$, $p =$

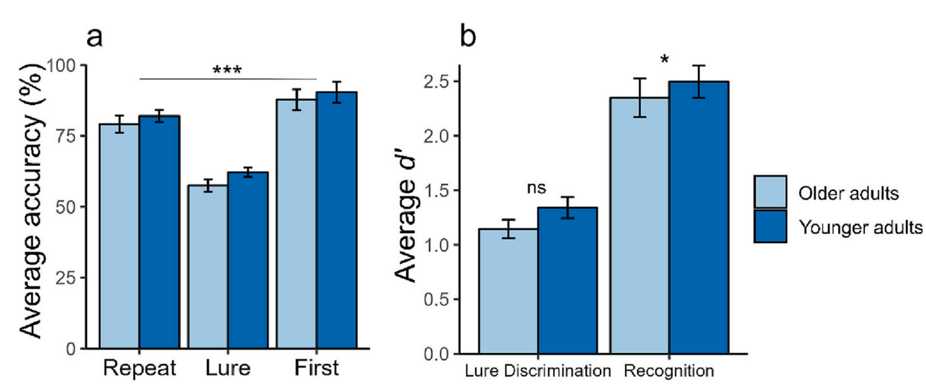

Fig. 3. Average accuracy for each item type (a) and average Lure Discrimination and Recognition scores (b) on the MST. Error bars represent 95% confidence intervals. * = $p < .05$, *** = $p < .001$.

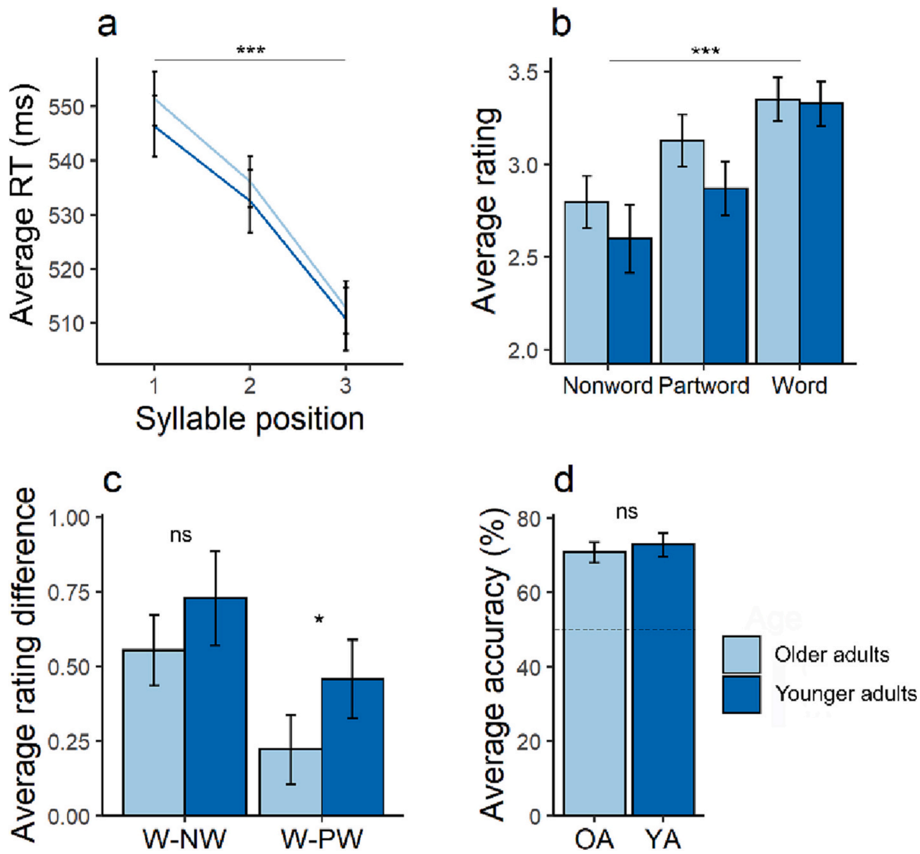

Fig. 4. Results of the statistical learning tasks. (a) Average reaction time as a function of syllable position on the Target Detection task. (b) Average familiarity rating for each condition on the Rating task. (c) Average rating difference between Words and Nonwords (W-NW) and between Words and Partwords (W-PW). (d) Average accuracy on the 2AFC Recognition task. Error bars represent 95% confidence intervals. * = $p < .05$, *** = $p < .001$.

.57), suggesting that the lower W-PW scores for older adults were not due to an increased tendency for older adults to rate words as familiar.

5.2.3. 2AFC recognition task

The average performance accuracy was 72.8% for younger adults and 70.7% for older adults. Both scores were significantly above chance level (OA: $t(60) = 12.26, p < .001$; YA: $t(58) = 11.67, p < .001$; Fig. 4d). No difference was found between the two groups in performance accuracy ($t(118) = -0.81, p = .42$). Bayesian analysis also provided substantial evidence in favor of the absence of an age effect ($BF_{10} = 3.83$).

5.2.4. Correlational analysis

Across age groups, performance on the target detection task was significantly correlated with accuracy on the 2AFC task ($r(120) = 0.36$, $p < .001$). This correlation remained significant when the two age groups were analyzed separately (YA: $r(59) = 0.41, p < .001$; OA: $r(61) = 0.36, p < .001$). There was not a significant correlation between RT score and the W-PW ($r(120) = 0.16, p = .081$) or the W-NW scores ($r(120) = 0.17, p = .068$) on the Rating task.

5.3. Mediation analysis

The finding that both Lure Discrimination and W-PW scores showed a significant age effect prompted us to further investigate the relationship between these variables. Compared to the W-NW score and the 2AFC recognition task, the W-PW score relies more on high-precision representations of triplets from the stream and the ability to discriminate them from partially overlapping triplets, which we expected to rely on those mechanisms of pattern separation that we found to be impaired with advanced age on the MST. Thus, we tested whether individual differences in Lure Discrimination significantly predicted W-PW using a linear regression model. The model revealed that Lure Discrimination significantly predicted W-PW score ($F(1, 118) = 10.37, p = .002$) and remained significant even after age group was added as a second predictor (Beta = 0.34, $t(118) = 2.78, p = .006$). In contrast, age group did not predict W-PW scores in this latter combined regression model (Beta = 0.17, $t(118) = 1.60, p = .11$), suggesting a possible mediating role of Lure Discrimination between age group and W-PW scores that shows specificity for high-precision discrimination. We formally tested this idea with a mediation analysis which revealed a significant indirect effect of age on W-PW score that was mediated by Lure Discrimination (Beta = 0.06, $p = .024$, tested with 1000 bootstrapping samples; Fig. 5), in addition to the direct effect of age on Lure Discrimination (Beta = 0.20, $t(118) = 2.52, p = .013$) and the direct effect of Lure Discrimination on W-PW scores (Beta = 0.34, $t(118) = 2.78, p = .006$). A linear regression model also revealed a significant correlational relationship

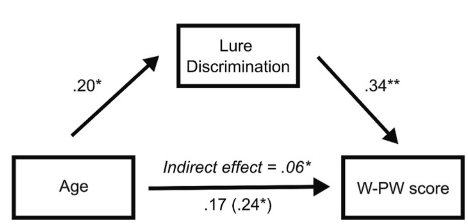

Fig. 5. Mediation model representing the effect of age on W-PW scores with Lure Discrimination as mediator. Path coefficients represent the beta coefficients from multiple regression analyses. The parentheses represent the total (direct + indirect) effect of age on W-PW score. All reported beta coefficients are unstandardized regression coefficients. * = $p < .05$, ** = $p < .01$.

between Lure Discrimination and W-NW score ($r(118) = 0.26$, $p = .0047$). However, as there was not a significant age effect on W-NW scores, no further mediation analysis was performed on W-NW scores.

6. Discussion

This study tested whether pattern separation and statistical learning are computationally separate processes by assessing the impact of healthy aging on these two processes. Using auditory linguistic versions of pattern separation and statistical learning tasks, we showed that pattern separation is impaired in aging, while performance on the reaction time-based and the forced-choice recognition measures of statistical learning remains preserved. This finding supports our hypothesis that pattern separation and statistical learning are differentially affected by aging. However, this dissociation disappeared when statistical learning was probed with an explicit rating task of item familiarity. Older adults performed worse in differentiating learned triplets from foil triplets. Critically, the age effect in high-fidelity but not low-fidelity discrimination was significantly mediated by pattern separation, suggesting that the explicit expression of statistical learning at a high-precision level depends in part on those mechanisms of pattern separation that we show are affected by age-related decline.

Aging impacted performance on the rating task but not the 2AFC recognition measure. From a perspective of explicit versus implicit retrieval demands, this result may be surprising at first glance, as both tasks require participants to make direct, explicit judgments about prior exposure of learned triplets. However, unlike the rating task, the 2AFC task can be supported by a comparison of relative familiarity between the two items on each trial, and has been shown to be more resistant to aging and age-related decline than explicit episodic recollection as probed in research on recognition memory (e.g., Bastin & Van der Linden, 2003; Holdstock et al., 2002). In contrast, the rating task requires retrieval of a high-resolution, precise memory representation of an isolated item, with such precise memory retrieval being compromised in aging (e.g., Korkki, Richter, Jeyarathnarajah, & Simons, 2020). The idea that task format is crucial for revealing aging effects is also supported by several previous statistical learning studies. Relative to younger adults, older adults are impaired in rating words, partwords and nonwords presented in isolation (Fama, Schuler, Newport, & Turkeltaub, 2022), but often show preserved performance on 2AFC recognition tasks, even when partwords are included (Ong & Chan, 2019; Shaqiri et al., 2018; Palmer et al., 2018; though note the latter study found a marginal effect of age on word versus part-word trials driven by a high cognitive load condition). However, an exception is a recent study by Schevenels et al. (2021), who reported poorer performance in older adults on a 2AFC recognition task that included only nonwords as foils.

In contrast, the implicit target detection task does not require engagement of declarative retrieval processes and memory-based decision making but relies on reaction time facilitation instead. Faster reaction time to the later syllables in a triplet reflects successful prediction of syllable occurrence that results from the acquisition of associations between consecutively encountered syllables. Such associative priming effects can occur independently of explicit awareness and recognition (Batterink et al., 2015), and there is evidence these effects are maintained into old age (Kan et al., 2011). Therefore, we conclude that older adults maintain the ability to discover and predict statistical regularities in input, but show a decline in the ability to encode, maintain and/or retrieve a high-fidelity memory representation of the learned statistical regularities. This task-specific demand could account for the observed age effect on the rating task. These findings support the recent view that statistical learning is a multi-component process (Arciuli, 2017; Conway, 2020). These findings also suggest that studies that probe statistical learning only with a task that involves high-precision memory retrieval may underestimate older adults' true statistical learning abilities.

The mediation effect of pattern separation on the word-partword score in the rating task provides direct support that the observed age effect for high-fidelity discrimination is driven by variations in pattern separation abilities, rather than statistical learning per se. Pattern separation mechanisms may contribute to the acquisition of the type of knowledge that can be expressed through the explicit rating task, potentially in parallel with the extraction of statistical regularities that are expressed implicitly. Alternatively, pattern separation mechanisms may be most critical at the time of retrieval, without playing a direct role in the learning process itself. We note that other higher-level cognitive processes such as working memory are also likely to contribute to the explicit expression of statistical learning (e.g., Palmer et al., 2018). Although the current data do not provide a mechanistic account of the interplay between these interacting processes, these initial set of findings open a window toward investigating these interconnections in aging and beyond.

The observed dissociable effects of healthy aging on explicit versus implicit measures of statistical learning show parallels with existing literature outside of the linguistic statistical learning area. A study that used a target detection task to measure visual statistical learning reported preserved learning in older adults (Campbell et al., 2012). Dissociations between explicit and implicit measures have also been found in implicit learning studies that employed multiple measurements. Implicit learning of motor sequences and novel word-pairs were expressed equally across age on indirect measures of learning, but aging effects emerged when direct, recognition-based measures were used (Gaillard et al., 2009; Light, Singh, & Capps, 1986; Schevenels et al., 2021). Our results extend these findings by showing that similar dissociations also emerge in learning of statistical regularities in syllable sequences. Together, these studies provide converging evidence that statistical learning continues to be robust into late adulthood and that any age effect is likely due to additional cognitive demands inherent to the explicit tasks.

The current findings also show parallels to the larger literature on age effects for memory about gist versus detail (see Brainerd & Reyna, 2015; Devitt & Schacter, 2016; Grilli & Sheldon, 2022 for reviews). A significant body of experimental research on episodic memory has established that gist representations of prior experiences (such as information about the superordinate categories to which previously studied items belonged; e.g., the list contained cats) tends to be well preserved among older adults. By contrast, information about specific details (e.g., the colour of the cats encountered) tends to be sensitive to aging. This dissociation has been particularly well documented in research on false memory conducted with visual recognition-memory paradigms, which has revealed that older adults frequently show false-alarm responses to novel items that are perceptually or conceptually similar to those that they previously studied (e.g., Koutstaal & Schacter, 1997). That we observed an increased rate in false alarms to perceptually similar lure items but not to dissimilar foil items on our MST converges with these findings and extends them to memory for speech material in the auditory domain.

The greater impact of aging on pattern separation relative to statistical learning, when probed implicitly, can be interpreted with reference to their proposed underlying distinct neural substrates (Schapiro et al., 2017). The decline of pattern separation in healthy aging has been associated with structural and functional changes in the dentate gyrus and CA3 hippocampal subregions. Studies using high resolution structural neuroimaging observed greater volume reduction in the dentate gyrus compared to other hippocampal subfields in older adults (Bennett & Stark, 2016; Dillon et al., 2017; Malykhin, Huang, Hrybouski, & Olsen, 2017). Findings from diffusion-weighted imaging studies demonstrated reduced perforant pathway integrity in older adults (Yassa, Mattfeld, Stark, & Stark, 2011). Functional neuroimaging studies have also revealed a correlation between dentate gyrus and CA3 activation and behavioral pattern separation performance (Reagh et al., 2018; Yassa, Lacy, et al., 2011). Through mediation analysis, Dillon et al. (2017) revealed that dentate gyrus volume reduction specifically mediates impairment in behavioral pattern separation in older adults.

In contrast, statistical learning has been hypothesized to recruit a complementary mechanism that relies on CA1-dependent hippocampal circuitry (Schapiro et al., 2017). CA1 volume has been reported to remain relatively age-resistant (Dillon et al., 2017), although findings are not entirely consistent across studies (Kurth, Cherbuin, & Luders, 2017). In addition, statistical learning is also associated with a network of brain regions outside of the hippocampus, which includes the inferior frontal gyrus, the striatum, and medial temporal cortex (Batterink & Paller, 2019; Frost, Armstrong, Siegelman, & Christiansen, 2015; Karuza et al., 2013; McNealy, Mazziotta, & Dapretto, 2006; Sandoval, Patterson, Dai, Vance, & Plante, 2017; Schapiro, Kustner, & Turk-Browne, 2012; Turk-Browne, Scholl, Chun, & Johnson, 2009). However, the interplay between these regions as a function of aging is not well understood. Studies on motor sequence learning point to a compensatory relationship between the striatum and the medial temporal lobe in aging (Rieckmann & Bäckman, 2009). Whether this relationship generalizes to statistical learning is currently unclear and deserves further investigation in future research. A direct examination of differences in the reliance on the dentate gyrus relative to the CA1 is another important avenue that could provide support for the account of the dissociation between pattern separation and statistical learning we suggest in the current study.

To ensure consistency in the stimulus materials used to assess pattern separation and statistical learning, we developed a novel, auditory linguistic version of the continuous MST (Yassa et al., 2010). To our knowledge, only a few studies have investigated pattern separation in the auditory domain (Bjornn, 2018; Herman, Baker, Cazes, Alain, & Rosenbaum, 2020; Trier, Lacy, & Marsh, 2016; Wang et al., 2023), and none has investigated effects of healthy aging on auditory pattern separation. The current findings provide evidence that the widely-documented age-related decline in pattern separation is not restricted to the visual domain but also extends into the auditory modality. Our novel paradigm may pave the way for future studies of pattern separation of auditory stimuli in aging and in other populations.

There are several limitations to this study. First, while statistical learning was measured with both implicit and explicit measures to account for possible effects of the variability in explicit memory retrieval on the expression of statistical learning, pattern separation was only measured with an explicit recognition task. As such, there remains a possibility that the observed age-related deficits in pattern separation are due to task-specific demands for explicit cognitive processing, and that the implicit expression of pattern separation would be preserved in older adults. Although implicit measures have rarely been used by previous pattern separation studies, such measures may be developed in the future by adapting methodologies commonly used in the implicit memory literature, such as priming. Such implicit pattern separation measures may further elucidate the impact of aging on pattern separation and its relationship with statistical learning. A second related limitation is that the current design does not allow us to fully disentangle the effects of precision of representation and explicit versus implicit access to these representations. The 2AFC task used in the current work involves explicit retrieval but requires only coarse knowledge. Future work could tease apart these contributions using more complex designs that orthogonalize these factors, such as by using a 2AFC task that incorporates both partwords and nonwords. Third, the current study provides limited evidence on the underlying neural mechanism of the two computational processes. Future research with neuroimaging is warranted to examine whether the observed dissociation between pattern separation and the implicit expression of statistical learning arises from the difference in their underlying neural architectures, as hypothesized by Schapiro et al. (2017). Lastly, while the mediation analysis in the current study provides an interesting exploration of the relationship between pattern separation and the explicit expression of statistical learning, its results should be interpreted with caution as it was conducted on cross-sectional data (O'Laughlin, Martin, & Ferrer, 2018; Raz & Lindenberger, 2011). Future research using a well-powered single sample with a continuous, wide age distribution may shed more light on the potential dependency between the two processes.

In conclusion, our findings provide behavioral evidence that pattern separation and statistical learning are distinct computational processes that are differentially impacted by healthy aging. This dissociation emerges most clearly when the behavioral expression of statistical learning is probed implicitly. Future research can build on the current findings and address the observed aging effect on these two processes in direct relation to changes in their neural substrates, using structural and/or functional imaging. Our results also reveal age-related dissociations within statistical learning, with older adults showing intact performance on implicit and forced-choice recognition-based measures of learning, along with poorer performance on an explicit memory task that requires access to highly precise triplet representations. This effect is at least partially mediated by task-specific demands for high precision memory that build on pattern separation and that decline with age. In contrast to explicit pattern separation mechanisms, our results suggest that the ability to acquire implicit knowledge during statistical learning is not affected by aging.

CRediT authorship contribution statement

Helena Shizhe Wang: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Stefan Köhler:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Laura J. Batterink:** Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Funding acquisition.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Natural Sciences and Engi-neering Research Council of Canada (https://dx.doi.org/10.13039/501100002790), grant numbers RGPIN-2019-05132 to L. J. B and RGPIN-2018-05770 to S. K.

References

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372*(1711), 20160058. https://doi.org/10.1098/rstb.2016.0058

Baker, S., Vieweg, P., Gao, F., Gilboa, A., Wolbers, T., Black, S. E., & Rosenbaum, R. S. (2016). The human dentate gyrus plays a necessary role in discriminating new memories. *Current Biology, 26*(19), 2629–2634. https://doi.org/10.1016/j.cub.2016.07.081

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science, 319*(5870), 1640–1642. https://doi.org/10.1126/science.1152882

Bastin, C., & Van der Linden, M. (2003). The contribution of recollection and familiarity to recognition memory: A study of the effects of test format and aging. *Neuropsychology, 17*(1), 14–24. https://doi.org/10.1037/0894-4105.17.1.14

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex, 90*, 31–45. https://doi.org/10.1016/j.cortex.2017.02.004

Batterink, L. J., & Paller, K. A. (2019). Statistical learning of speech regularities can occur outside the focus of attention. *Cortex, 115*, 56–71. https://doi.org/10.1016/j.cortex.2019.01.013

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language, 83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Bennett, I. J., & Stark, C. E. L. (2016). Mnemonic discrimination relates to perforant path integrity: An ultra-high resolution diffusion tensor imaging study. *Neurobiology of Learning and Memory, 129*, 107–112. https://doi.org/10.1016/j.nlm.2015.06.014

Berron, D., Schutze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H. J., Kumaran, D., & Duzel, E. (2016). Strong evidence for pattern separation in human dentate gyrus. *Journal of Neuroscience, 36*(29), 7569–7579. https://doi.org/10.1523/JNEUROSCI.0518-16.2016

Bertels, J., Franco, A., & Destrebecqz, A. (2012). How implicit is visual statistical learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1425–1431. https://doi.org/10.1037/a0027210

Bjornn, D. K. (2018). *Contribution of target-lure similarity and sensory modality to lure false alarms* (p. 2018).

Brainerd, C. J., & Reyna, V. F. (2015). Fuzzy-trace theory and lifespan cognitive development. *Developmental Review, 38*, 89–121. https://doi.org/10.1016/j.dr.2015.07.006

Campbell, K. L., Zimerman, S., Healey, M. K., Lee, M. M. S., & Hasher, L. (2012). Age differences in visual statistical learning. *Psychology and Aging, 27*(3), 650–656. https://doi.org/10.1037/a0026780

Christiansen, M. H. (2019). Implicit statistical learning: A tale of two literatures. *Topics in Cognitive Science, 11*(3), 468–481. https://doi.org/10.1111/tops.12332

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews, 112*, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032

Devitt, A. L., & Schacter, D. L. (2016). False memories with age: Neural and cognitive underpinnings. *Neuropsychologia, 91*, 346–359. https://doi.org/10.1016/j.neuropsychologia.2016.08.030

Dillon, S. E., Tsivos, D., Knight, M., McCann, B., Pennington, C., Shiel, A. I., … Coulthard, E. J. (2017). The impact of ageing reveals distinct roles for human dentate gyrus and CA3 in pattern separation and object recognition memory. *Scientific Reports, 7*(1). https://doi.org/10.1038/s41598-017-13853-8. Article 1.

Fama, M. E., Schuler, K. D., Newport, E. L., & Turkeltaub, P. E. (2022). Effects of healthy aging and left hemisphere stroke on statistical language learning. *Language, Cognition and Neuroscience, 0*(0), 1–16. https://doi.org/10.1080/23273798.2022.2030481

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin, 145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Gaillard, V., Destrebecqz, A., Michiels, S., & Cleeremans, A. (2009). Effects of age and practice in sequence learning: A graded account of ageing, learning, and control. *European Journal of Cognitive Psychology, 21*(2–3), 255–282. https://doi.org/10.1080/09541440802257423

Grilli, M. D., & Sheldon, S. (2022). Autobiographical event memory and aging: Older adults get the gist. *Trends in Cognitive Sciences, 26*(12), 1079–1089. https://doi.org/10.1016/j.tics.2022.09.007

Herff, S. A., Zhen, S., Yu, R., & Agres, K. R. (2020). Age-dependent statistical learning trajectories reveal differences in information weighting. *Psychology and Aging, 35*(8), 1090–1104. https://doi.org/10.1037/pag0000567

Herman, D., Baker, S., Cazes, J., Alain, C., & Rosenbaum, R. S. (2020). Mismatch negativity predicts pattern separation [preprint]. *Neuroscience.*. https://doi.org/10.1101/2020.05.19.102707

Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., & Norman, K. A. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus, 12*(3), 341–351. https://doi.org/10.1002/hipo.10011

Howard, J. H., & Howard, D. V. (2013). Aging mind and brain: Is implicit learning spared in healthy aging? *Frontiers in Psychology, 4*, 817. https://doi.org/10.3389/fpsyg.2013.00817

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan: Implicit learning across human lifespan. *Developmental Science, 15*(4), 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x

Kan, I. P., Keane, M. M., Martin, E., Parks-Stamm, E. J., Lewis, L., & Verfaellie, M. (2011). Implicit memory for novel associations between pictures: Effects of stimulus unitization and aging. *Memory & Cognition, 39*(5), 778–790. https://doi.org/10.3758/s13421-011-0071-6

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language, 127*(1), 46–54. https://doi.org/10.1016/j.bandl.2012.11.007

Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters, 461*(2), 145–149. https://doi.org/10.1016/j.neulet.2009.06.030

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory, 14*(9), 625–633. https://doi.org/10.1101/lm.663507

Korkki, S. M., Richter, F. R., Jeyarathnarajah, P., & Simons, J. S. (2020). Healthy ageing reduces the precision of episodic memory retrieval. *Psychology and Aging, 35*(1), 124–142. https://doi.org/10.1037/pag0000432

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37*(4), 555–583. https://doi.org/10.1006/jmla.1997.2529

Kurth, F., Cherbuin, N., & Luders, E. (2017). The impact of aging on subregions of the hippocampal complex in healthy adults. *NeuroImage, 163*, 296–300. https://doi.org/10.1016/j.neuroimage.2017.09.016

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 (Cold Spring Harbor, N.Y.), 18*(1), 15–18. https://doi.org/10.1101/lm.1971111

Leal, S. L., Tighe, S. K., & Yassa, M. A. (2014). Asymmetric effects of emotion on mnemonic interference. *Neurobiology of Learning and Memory, 111*, 41–48. https://doi.org/10.1016/j.nlm.2014.02.013

Leal, S. L., & Yassa, M. A. (2014). Effects of aging on mnemonic discrimination of emotional information. *Behavioral Neuroscience, 128*(5), 539–547. https://doi.org/10.1037/bne0000011

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the Hippocampus. *Science, 315*(5814), 961–966. https://doi.org/10.1126/science.1135801

Light, L. L., Singh, A., & Capps, J. L. (1986). Dissociation of memory and awareness in young and older adults. *Journal of Clinical and Experimental Neuropsychology, 8*(1), 62–74. https://doi.org/10.1080/01688638608401297

Lukács, Á., & Kemény, F. (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science, 39*(2), 383–404. https://doi.org/10.1111/cogs.12143

Ly, M., Murray, E., & Yassa, M. A. (2013). Perceptual versus conceptual interference and pattern separation of verbal stimuli in young and older adults. *Hippocampus, 23*(6), 425–430. https://doi.org/10.1002/hipo.22110

Malykhin, N. V., Huang, Y., Hrybouski, S., & Olsen, F. (2017). Differential vulnerability of hippocampal subfields and anteroposterior hippocampal subregions in healthy cognitive aging. *Neurobiology of Aging, 59*, 121–134. https://doi.org/10.1016/j.neurobiolaging.2017.08.001

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *Journal of Neuroscience, 26*(29), 7629–7639. https://doi.org/10.1523/JNEUROSCI.5501-05.2006

Neger, T., Rietveld, T., & Janse, E. (2015). *Adult age effects in auditory statistical learning* (p. 2015).

Neger, T. M., Rietveld, T., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience, 8*. https://doi.org/10.3389/fnhum.2014.00628

O'Laughlin, K. D., Martin, M. J., & Ferrer, E. (2018). Cross-sectional analysis of longitudinal mediation processes. *Multivariate Behavioral Research, 53*(3), 375–402. https://doi.org/10.1080/00273171.2018.1454822

Ong, J. H., & Chan, A. H. D. (2019). *Statistical learning among young and older adults: Similar yet different?* (p. 2019).

Palmer, S. D., Hutson, J., & Mattys, S. L. (2018). Statistical learning for speech segmentation: Age-related changes and underlying mechanisms. *Psychology and Aging, 33*(7), 1035–1044. https://doi.org/10.1037/pag0000292

Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing mediation in communication research. In *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13–54). Sage Publications, Inc. https://doi.org/10.4135/9781452272054.n2.

Raviv, L., & Arnon, I. (2017). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science, 21*. https://doi.org/10.1111/desc.12593

Raz, N., & Lindenberger, U. (2011). Only time will tell: Cross-sectional studies offer no solution to the age–brain–cognition triangle: Comment on Salthouse (2011). *Psychological Bulletin, 137*, 790–795. https://doi.org/10.1037/a0024503

Reagh, Z. M., Noche, J. A., Tustison, N. J., Delisle, D., Murray, E. A., & Yassa, M. A. (2018). Functional imbalance of anterolateral entorhinal cortex and hippocampal dentate/CA3 underlies age-related object pattern separation deficits. *Neuron, 97*(5), 1187–1198.e4. https://doi.org/10.1016/j.neuron.2018.01.039

Reed, M., Freedman, M., Mark Fraser, A. E., Bromwich, M., Santiago, A. T., Gallucci, C. E., & Frank, A. (2022). Enhancing clinical visibility of hearing loss in cognitive decline. *Journal of Alzheimer's Disease, 86*(1), 413–424. https://doi.org/10.3233/JAD-215377

Rieckmann, A., & Bäckman, L. (2009). Implicit learning in aging: Extant patterns and new directions. *Neuropsychology Review, 19*(4), 490–503. https://doi.org/10.1007/s11065-009-9117-y

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*(2), 101–105. https://doi.org/10.1111/j.1467-9280.1997.tb00690.x

Sandoval, M., Patterson, D., Dai, H., Vance, C. J., & Plante, E. (2017). Neural correlates of morphology acquisition through a statistical learning paradigm. *Frontiers in Psychology, 8*, 1234. https://doi.org/10.3389/fpsyg.2017.01234

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology, 22*(17), 1622–1627. https://doi.org/10.1016/j.cub.2012.06.056

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), 20160049. https://doi.org/10.1098/rstb.2016.0049

Schevenels, K., Altvater-Mackensen, N., Zink, I., De Smedt, B., & Vandermosten, M. (2021). Aging effects and feasibility of statistical learning tasks across modalities. *Aging, Neuropsychology, and Cognition*, 1–30. https://doi.org/10.1080/13825585.2021.2007213

Shaqiri, A., Danckert, J., Burnett, L., & Anderson, B. (2018). Statistical learning impairments as a consequence of stroke. *Frontiers in Human Neuroscience, 12*. https://doi.org/10.3389/fnhum.2018.00339

Sherman, B. E., & Turk-Browne, N. B. (2020). Statistical prediction of the future impairs episodic encoding of the present. *Proceedings of the National Academy of Sciences, 117*(37), 22760–22770. https://doi.org/10.1073/pnas.2013291117

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods, 49*(2), 418–432. https://doi.org/10.3758/s13428-016-0719-z

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

Stark, S. M., Stevenson, R., Wu, C., Rutledge, S., & Stark, C. E. L. (2015). Stability of age-related deficits in the mnemonic similarity task across task variations. *Behavioral Neuroscience, 129*(3), 257–268. https://doi.org/10.1037/bne0000055

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), 2442–2449. https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience, 10*(1), 21. https://doi.org/10.1186/1471-2202-10-21

Trier, H. A., Lacy, J. W., & Marsh, B. U. (2016). Limitations of episodic memory for highly similar auditory stimuli. *Journal of Cognitive Psychology, 28*(7), 843–855. https://doi.org/10.1080/20445911.2016.1204306

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*(4), 552–564. https://doi.org/10.1037/0096-3445.134.4.552

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience, 30*(33), 11177–11187. https://doi.org/10.1523/JNEUROSCI.0858-10.2010

Wang, H. S., Rosenbaum, R. S., Baker, S., Lauzon, C., Batterink, L. J., & Köhler, S. (2023). Dentate Gyrus Integrity Is Necessary for Behavioral Pattern Separation But Not Statistical Learning. *Journal of Cognitive Neuroscience, 35*(5), 900–917. https://doi.org/10.1162/jocn_a_01981. PMID: 36877071.

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics, 79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2011). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus, 21*(9), 968–979. https://doi.org/10.1002/hipo.20808

Yassa, M. A., Mattfeld, A. T., Stark, S. M., & Stark, C. E. L. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences, 108*(21), 8873–8878. https://doi.org/10.1073/pnas.1101567108

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

Yassa, M. A., Stark, S. M., Bakker, A., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2010). High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnestic mild cognitive impairment. *NeuroImage, 51*(3), 1242–1252. https://doi.org/10.1016/j.neuroimage.2010.03.040