Dentate Gyrus Integrity is Necessary for Behavioral Pattern Separation but not Statistical Learning

[1]Helena Shizhe Wang, [2,3]R. Shayna Rosenbaum, [2,3]Stevenson Baker, [2,3]Claire Lauzon, [1,4]Laura J. Batterink*, [1,3,4]Stefan Köhler*

[1]The Brain and Mind Institute, University of Western Ontario

[2]Department of Psychology, York University

[3]Rotman Research Institute, Baycrest

[4]Department of Psychology, University of Western Ontario

* indicates equal contribution

Corresponding Author:

Stefan Köhler

Department of Psychology,

University of Western Ontario

stefank@uwo.ca

Abstract

Pattern separation, the creation of distinct representations of similar inputs, and statistical learning, the rapid extraction of regularities across multiple inputs, have both been linked to hippocampal processing. It has been proposed that there may be functional differentiation within the hippocampus, such that the trisynaptic pathway (entorhinal cortex > dentate gyrus > CA3 > CA1) supports pattern separation, whereas the monosynaptic pathway (entorhinal cortex > CA1) supports statistical learning. To test this hypothesis, we investigated the behavioral expression of these two processes in BL, an individual with highly selective bilateral lesions in the dentate gyrus that presumably disrupts the trisynaptic pathway. We tested pattern separation with two novel auditory versions of the continuous Mnemonic Similarity Task, requiring the discrimination of similar environmental sounds and trisyllabic words. For statistical learning, participants were exposed to a continuous speech stream made up of repeating trisyllabic words. They were then tested implicitly through a reaction time-based task and explicitly through a rating task and a forced-choice recognition task. BL showed significant deficits in pattern separation on the Mnemonic Similarity Tasks and on the explicit rating measure of statistical learning. In contrast, BL showed intact statistical learning on the implicit measure and the familiarity-based forced-choice recognition measure. Together, these results suggest that dentate gyrus integrity is critical for high-precision discrimination of similar inputs, but not the implicit expression of statistical regularities in behaviour. Our findings offer unique new support for the view that pattern separation and statistical learning rely on distinct neural mechanisms.

Introduction

The ability to distinctly remember events with overlapping features is an integral part of episodic memory functioning. Pattern separation refers to the formation of discrete neural representations of similar inputs and is thought to support the ability to distinguish between similar events in memory (Marr et al., 1971). A separate, yet equally important feature of memory is the extraction of commonalities shared across events, which supports the ability to predict future events. The process of extracting such regularities in the environment over time is referred to as statistical learning (Saffran, Aslin, et al., 1996). At the functional level, pattern separation and statistical learning serve different goals and may rely on computations that cannot be supported by the same neural circuitry. The current study aims to investigate whether the two processes are dissociable in their neural mechanisms by examining the impact of a highly selective lesion in a structure that has previously been linked to both of them, namely, the hippocampus.

The hippocampus has long been known to play an important role in episodic memory (Scoville & Milner, 1957; Squire & Zola-Morgan, 1991; Vargha-Khadem et al., 1997). Within the hippocampus, computational models have posited that the sparse coding of granule cells in the dentate gyrus subregion form distinct representations of each episodic instance, which are then passed onto the downstream CA3 through the mossy fiber pathway (Marr et al., 1971; McClelland et al., 1995; Norman & O'Reilly, 2003). Recordings from granule cells in rodent models have shown heightened activation levels during spatial pattern separation (Leutgeb et al., 2007). Evidence of pattern separation in humans comes from studies that employed functional magnetic resonance imaging (MRI) adaptation paradigms in combination with presentation of pictures of objects of varying similarity (Bakker et al., 2008; Lacy et al., 2011). In these studies, the activation level of a hippocampal region that comprised dentate gyrus and CA3 elicited by visual "lures" (items similar to previously encountered items) was found to more closely resemble that elicited by new items than by repeated items. This result has been taken to suggest that dentate gyrus/CA3 is sensitive to small changes in input signals (Bakker et al., 2008; Lacy et al., 2011), supporting a potential role in pattern separation. Ultrahigh-resolution functional neuroimaging has also revealed distinguishable activation patterns for similar scenes in the dentate gyrus, and not in other hippocampal subfields when the stimuli were matched on the novelty-familiarity dimension (Berron et al., 2016). There is also evidence from patient studies, indicating that individuals with amnesic mild cognitive impairments and reduced volume in dentate gyrus/CA3 show deficits in their ability to discriminate previously seen objects from similar lures (Yassa et al., 2010). Collectively, these results support the role of dentate gyrus in pattern separation. It must be noted, however, that some of the imaging protocols employed did not have sufficient resolution to allow for clear differentiation of dentate gyrus from neighboring CA3.

Further support for the critical role of the dentate gyrus in pattern separation is provided by the case report of BL, an amnesic individual with highly selective hippocampal lesions affecting the dentate gyrus bilaterally (Baker et al., 2016; Kwan et al., 2015). BL's pattern separation abilities were assessed on a behavioral task consisting of incidental encoding of a series of visual objects, followed by a test requiring the classification of old objects, lures, and novel foils as "old", "similar", or "new" (Mnemonic Similarity Task; MST; Stark et al., 2013). BL was able to correctly identify most of the targets and foils, but showed a heightened tendency to endorse

lures as previously seen, indicating normal recognition abilities along with a specific deficit in hippocampally-dependent pattern separation (Baker et al., 2016).

While many different sources of evidence have implicated the hippocampus in computations of pattern separation that allow for the representation of highly similar events, recent findings suggest it is also involved in rapidly generalizing and extracting statistical regularities across multiple events (Covington et al., 2018; Ellis et al., 2021; Henin et al., 2021; Schapiro et al., 2014; Turk-Browne et al., 2009). Statistical learning is typically studied by presenting participants with a continuous stream of repeating triplets (e.g., ABCDEFABCGHI…), such that items within a triplet (e.g., A-B) co-occur more frequently than cross-boundary items (e.g., C-D). Participants' ability to become sensitive to this hidden structure is subsequently assessed on a recognition test, in which participants discriminate between target triplets and foils (Saffran, Newport, et al., 1996). Neuroimaging studies have reported enhanced hippocampal activation during viewing of visual items in a structured order compared to a random order (Ellis et al., 2021; Schapiro et al., 2012; Turk-Browne et al., 2009). Similarly, using intracranial recording, Henin et al. (2021) found evidence for unit (pairs or triplets)-based item organization in the hippocampus during both visual and auditory statistical learning. Further evidence comes from developmental works showing that individual differences in the volume of the hippocampus predict auditory (Finn et al., 2019) and visual (Schlichting et al., 2017) statistical learning in young children and adults. Although beyond the focus of the current study, many other brain regions outside of the hippocampus have also been implicated in statistical learning, including sensory/perceptual cortical regions, the left inferior frontal gyrus, and the striatum (see Batterink et al., 2019 for review). Their involvement has been suggested to differ depending on various aspects of statistical learning, such as sensory modality, input complexity, engagement of attention, and the types of statistical representations that are acquired (Conway, 2020; Henin et al., 2021).

Neuropsychological evidence also supports the role of the hippocampus in statistical learning. Schapiro et al. (2014) tested the statistical learning abilities of the case LSJ, who suffered complete bilateral hippocampal tissue loss and broader MTL damage, across four different types of visual and auditory stimuli (shapes, scenes, syllables, and tones). LSJ showed at-chance performance in the explicit recognition of target triplets from novel foils across all four types of stimuli. Using the same paradigm, Covington et al. (2018) reported that four other individuals with either selective hippocampal damage or more extensive MTL damage also showed impaired recognition performance relative to healthy controls, although here performance was still above chance. However, a key limitation of these studies is that statistical-learning performance was tested only using the typical forced-choice recognition task, which primarily captures explicit knowledge based on stimulus familiarity. Given previous work that statistical learning produces dissociable implicit and explicit knowledge in healthy adults (Batterink et al., 2015), these previous studies leave open the possibility that sensitivity to statistical structure in behaviour that does not require explicit memory judgements may occur independently of hippocampal contributions. Such implicit markers of statistical structure may be present, for example, in implicit reaction-time measures of learning (Batterink et al., 2015).

To reconcile the involvement of the hippocampus in both pattern separation and statistical learning, Schapiro et al. (2017) have postulated that the two processes may rely on separate neural circuitry within the hippocampus. Building on the influential complementary learning

systems model that differentiates between functionally distinct hippocampal and neocortical learning mechanisms (McClelland et al., 1995), Schapiro et al. hypothesized that similar complementary division of labour exists within the hippocampus, in the form of the trisynaptic pathway and the monosynaptic pathway, respectively. In Schapiro et al.'s model, the trisynaptic pathway, which projects from the entorhinal cortex to the dentate gyrus, CA3, and CA1 in successive stages, is considered to underlie the encoding of specific individual episodic instances. By contrast, the monosynaptic pathway, projecting directly from the entorhinal cortex to CA1, is proposed to support the integration of inputs across multiple temporally separated instances. Schapiro et al. tested their theory using a neural network model that simulated these physiological properties of the hippocampus. When the model was presented with an item sequence with explicit pair boundaries (AB|CD|EF|GH|AB…), paired items were represented distinctly within the dentate gyrus and CA3. In contrast, when items were presented continuously such that item pairings could only be learned through tracking the statistical regularities over time (ABCDABEFABEF…), item pairs were represented more strongly within CA1 than in the dentate gyrus or CA3. Moreover, a model with a "lesioned" trisynaptic pathway still successfully tracked statistical regularities, suggesting that the monosynaptic pathway alone is sufficient to support statistical learning.

A few empirical studies provide some preliminary support for the proposed division of labour between the trisynaptic pathway and the monosynaptic pathway. In infants, the correlation between statistical learning and hippocampal volume has been found to be stronger in the anterior hippocampus (Ellis et al., 2021)—which contains a greater proportion of the CA1 subfield (Canada et al., 2021; Malykhin et al., 2010)—than the posterior hippocampus. Sherman et al. (2020) hypothesized that, due to the CA1 region being shared by the two pathways, episodic encoding and statistical learning are in direct competition with each other. Supporting this hypothesis, a neuroimaging study reported that when faced with competing goals of encoding the current, predictive (A) stimulus and predicting the upcoming (B) stimulus, the hippocampus represented the current stimulus more weakly than the future stimulus, which also coincided with lower recognition performance for predictive stimuli (Sherman & Turk-Browne, 2020). The finding that prediction of future events impairs the episodic encoding of the current event is in line with the view that the two processes are in direct competition due to the overlap between their proposed neural mechanisms. However, to our knowledge, no study in human participants has directly tested the involvement of the trisynaptic pathway and the monosynaptic pathway in pattern separation and statistical learning, respectively.

We had a unique opportunity to assess whether there is a division of labor for pattern separation and statistical learning within the hippocampus by testing the unique case BL. BL has a rare, highly selective lesion of bilateral dentate gyrus, along with documented highly specific behavioral impairments in discriminating similar objects in memory (Baker et al., 2016). Based on the model of hippocampal circuit organization proposed by Schapiro et al. (2017), we specifically tested the hypothesis that a selective dentate gyrus lesion would impair pattern separation but leave statistical learning intact. To match learning materials across these two domains, highly similar spoken syllables were used as stimuli across both pattern separation and statistical learning tasks. Following the rationale put forward in the computational modelling work conducted by Schapiro et al. (2017), participants' ability to encode and distinguish individually chunked items was taken as measure of pattern separation, and their ability to extract items from a continuous syllable sequence was taken as a measure of statistical learning.

To capture both implicit and explicit knowledge accrued as a result of statistical learning, we used three separate statistical learning tasks: 1) a rating task in which subjects rated the familiarity of target triplets (words) and two types of foil triplets (partwords and nonwords), each presented in isolation, 2) a forced-choice recognition task in which participants had to select target triplets that were pitted directly against nonwords, and 3) a reaction time-based task that indirectly measured learning of statistical regularities through reaction time facilitation. Moreover, to ensure that the behavioural evidence for auditory pattern separation in BL was not specific to linguistic stimuli, we also administered a second pattern separation task using common environmental sounds. We predicted that BL would demonstrate a deficit in behavioural markers of pattern separation on both tasks. Critically, we further predicted that implicit expression of statistical learning would be preserved in BL, but that this intact performance would possibly go hand in hand with impairments on statistical learning tasks that require explicit stimulus discrimination, as previously observed in patients with less selective hippocampal lesions (Covington et al., 2018; Schapiro et al., 2014).

Methods

Participants

BL

BL was 60 years old at the time of testing and has 13 years of education. At age 24, BL suffered anoxic brain injury as a result of an electrical injury and cardiac arrest. As reported by Baker et al. (2016), high-resolution 3T MRI scans revealed highly selective bilateral ischemic lesions in the hippocampus that were primarily restricted to the dentate gyrus and a portion of CA3 (Figure 1), with signal abnormalities detected in the dentate gyrus bilaterally but not in CA1-2 or subiculum. Comparison of the volume of hippocampal subfields between BL and 119 age-matched controls revealed that BL's dentate gyrus volume is approximately 50% smaller along the entire anterior-posterior axis relative to controls (Baker et al., 2016). CA1, in contrast, is not reduced in size; in fact, it is numerically larger (8%) than in controls (Baker et al., 2016). Whole-brain imaging (Baker et al., 2016) also revealed small volume reductions in the left superior-posterior parietal cortex and right precuneus. BL's recent neurocognitive performance was reported in Mitchnick et al. (2022). BL scored 24/30 on the Montreal Cognitive Assessment,

which is slightly below the cut-off point for identifying mild cognitive impairment (26/30; Nasreddine et al., 2005).
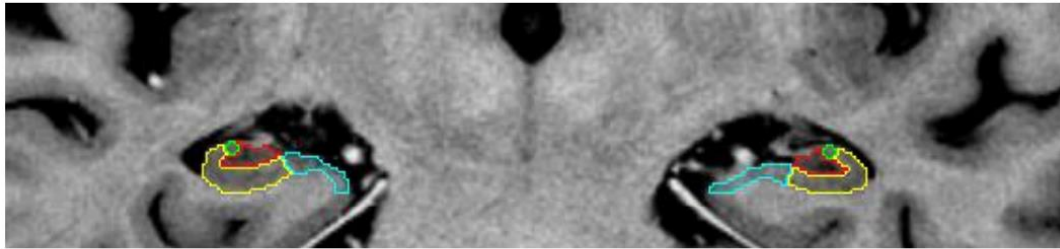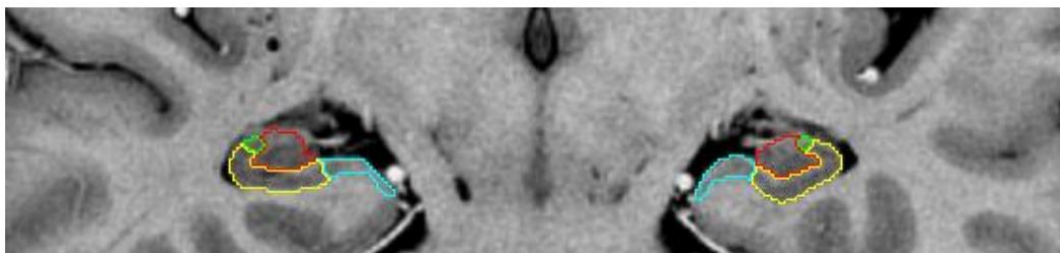

Figure 1. Hippocampal segmentation of BL (A) and controls (B). Red = Dentate gyrus & CA3, Green = CA1-2 transition, Yellow = CA1, Blue = Subiculum. (Adapted from the supplementary material of "The Human Dentate Gyrus Plays a Necessary Role in Discriminating New Memories" by S. Baker et al., 2016, *Current Biology*, 26(19), 2629-2634. with permission.)

Controls

Control participants matched to BL in terms of age, education level, and linguistic background were recruited both from the local community and through a crowdsourcing platform for online studies, Prolific.co. Community participants (n = 12, six females) had an average age of 60.5 (range = 57 – 65) and completed an average of 15 years of education (range = 10 – 18). All participants were monolingual English speakers, had normal vision and hearing, and had no history of neurological or psychiatric disorders. Testing of the community participants took place at the University of Western Ontario over two separate days.

Additional participants were recruited from Prolific to achieve a sample size of control participants similar to that of a previous study that also compared BL's MST performance to controls (Baker et al., 2016, who report a sample of *n* = 20). Prolific participants were filtered based on their response to Prolific's screening questionnaire. Participants were required to be between 55 and 65 years old, to identify as male, to be native English speakers, to have completed 12 to 14 years of education, to have no hearing difficulties, and to have no history of neurological or psychiatric disorders. Recruitment on Prolific was conducted separately for Word MST, Sound MST, and statistical learning tasks, and each participant completed just a single task. A total of 11 participants were recruited for Word MST, 11 for Sound MST, and 14 for statistical learning (nonoverlapping samples). Informed consent was obtained from all participants in compliance with the Research Ethics Board of the University of Western Ontario,

York University, and Baycrest Health Sciences. All participants were compensated for their time.

Stimuli

Word Mnemonic Similarity Task

The stimuli for this task were generated from a set of 25 unique trisyllabic nonsense words (e.g., *gopula*), 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. Another five of the original set were each used to create additional five "Lures" by recombining the three syllables in five different ways (e.g., *gopula: golapu, lagopu, lapugo, pugola, pulago*; referred to as a "lure family;" see the bottom of Figure 2A). The words in each of these lure families are highly similar to the original word and to one another; correctly differentiating them requires one to remember not only the three syllables but also the order in which they were combined (Forest et al., 2022; Park et al., 2018). The remaining 15 First presentation items served as foils and were never repeated. Thus, in total, the task consisted of 25 First presentation trials, 25 Lure trials, and 50 Repeat trials, for a total of 100 trials (Figure 2A). During the task, the items were ordered such that each Repeat or Lure item was separated by an average of 6 intervening items (range: 2 – 12 items). The first trial associated with an "Old" response occurred approximately 10 items into the task.

All word items were created by randomly pairing three different consonants with three different vowels (i.e., CVCVCV), with the 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 approximately 1s long and the perceived loudness was normalized across all sound files.

Environmental Sound Mnemonic Similarity Task
The experimental stimuli consisted of 88 unique common environmental sounds collected from Internet sources (FindSounds.com; Free SFX.co.uk; ZapSplat.com), which included sounds produced by animals, humans, and man-made objects. Task stimuli were selected on the basis of a pilot study, in which 25 online participants between ages 18 and 30 listened to 100 sounds and generated descriptive labels for each sound. The sounds were then ranked in terms of concept agreement (i.e., the proportion of participants who assigned the same label to a sound) and the 67 sounds with the highest agreement were used in either the actual experiment (64 sounds) or the practice (3 sounds). Of the 64 experimental sounds, 40 sounds were duplicated to create 40 "Repeat" pairs. The remaining 24 sounds were each paired with an additional sound from the same semantic category collected from the same Internet sources, creating 24 "Lure" pairs (e.g., two different "ringing bell" sounds). In total, the task consisted of 64 First presentation trials (40 first presentations of Repeat sounds + 24 first presentations of Lure sounds), 24 Lure trials (24 second presentations of Lure sounds), and 40 Repeat trials (40 second presentations of Repeat sounds; Figure 2B). All stimuli were manipulated using Audacity software to be monophonic, have a duration of 1 – 2.5 seconds, and have approximately the same perceived loudness. Items were ordered such that Repeats and Lures were separated by an average of 13 intervening items (range: 5 – 24) from their counterparts. The first trial associated with an "Old" response occurred approximately 10 items into the task.
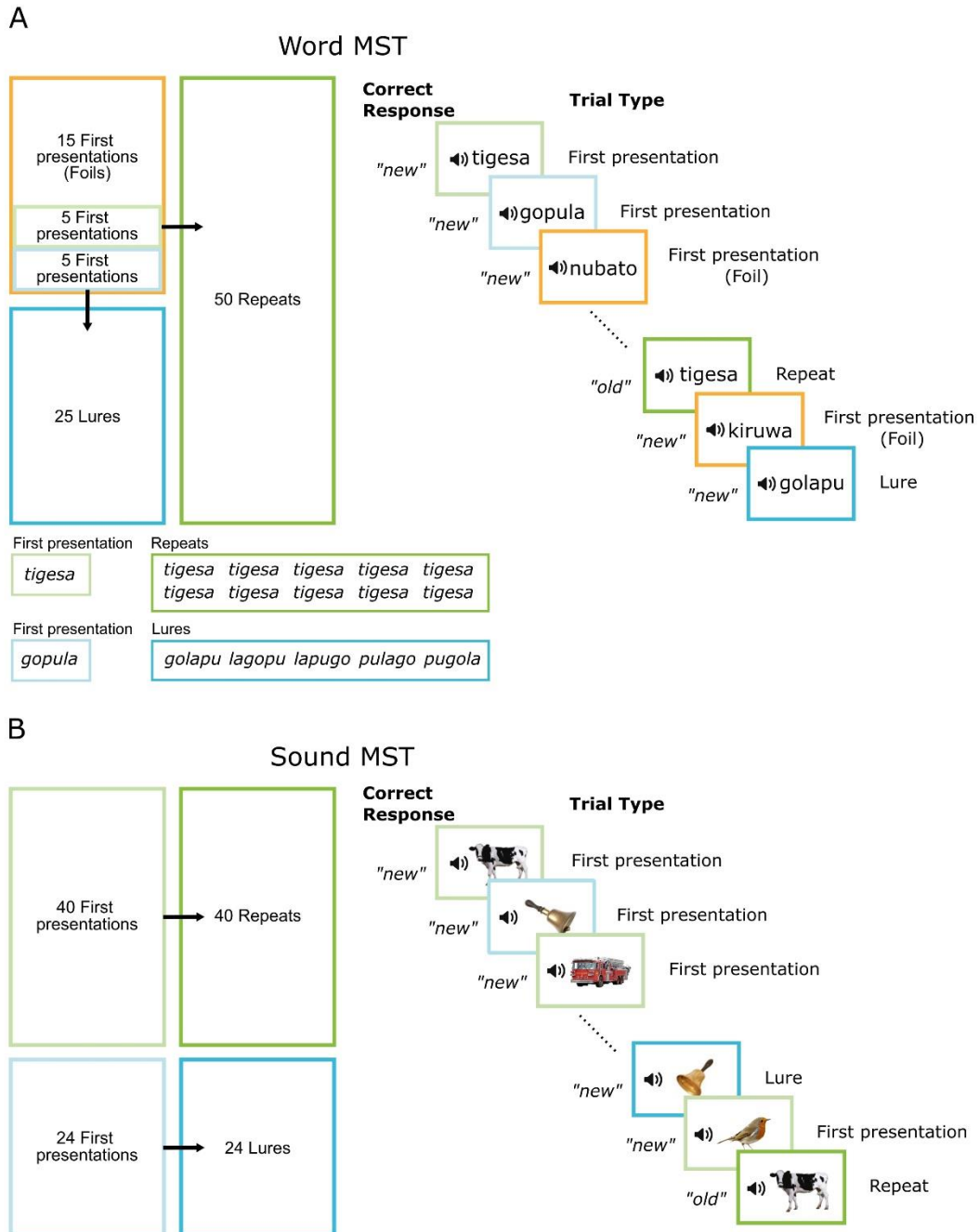
Figure 2. (A) Diagram of the Word Mnemonic Similarity Task. Of the initial 25 unique word items (orange), five were used to create 50 Repeat trials (green) and another five were used to create 25 Lure trials (blue). The items were ordered in a manner such that there was an average of 6 intervening items between a Repeat/Lure item and the subsequent Repeat/Lure item. Participants responded "Old" to repeated items and "New" to items that were presented for the first time or items that sounded similar to but different from a previous item. (B) The stimulus set for the Environmental Sound Mnemonic Similarity Task consisted of 40 "Repeat pairs" and 24 "Lure pairs".

Statistical Learning Task

The stimuli for the statistical learning tasks consisted of 12 unique syllables, taken from Batterink & Paller (2019), which were combined to form four trisyllabic nonsense words (*tafuko*, *regeme*, *rupuni*, *fetisu*; Figure 3). The sound file for each syllable was 300 ms in duration.

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-minute-long continuous stream.

Target Detection Task. A total of 36 speech streams were created. Each stream consisted of all four words in the language presented four times each (48 total syllables), concatenated together in the same manner as in the Exposure Phase but with the constraint that the word containing the target syllable never appeared as the first or the last word in the stream. This yielded 4 targets per stream, and a total of 48 targets per triplet position across the entire task. Each speech stream was 18.24 seconds long.

In order to examine possible learning effects during the target detection task itself, the 36 streams were 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 a way such that the syllable position of the target (word-initial, word-middle, word-final) were evenly distributed across the block (e.g., initial, middle, final, middle, initial, final…).

Explicit Statistical Learning Tasks. The stimuli for the Rating Task consisted of 12 trisyllabic items: four Words from the Exposure Phase, four 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 that never occurred adjacent to each other during the exposure phase (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. The same four Words and four Nonwords were used as the stimuli for the Two-Alternative Forced-Choice (2AFC) Recognition Task (Partwords were not included).
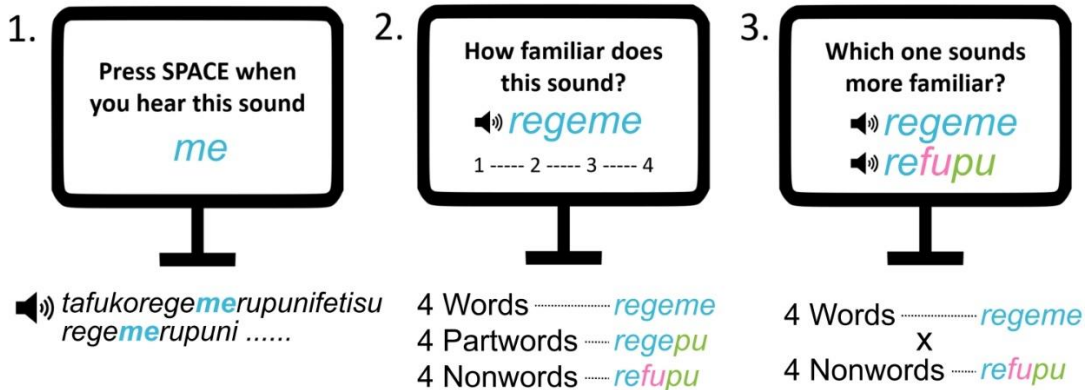
Exposure phase

Test phase

Figure 3. Diagram of the statistical learning paradigm. Following a 6.8 min exposure phase, participants completed 1) the Target Detection task, 2) the Rating task, and 3) the 2AFC Recognition task. The four Partwords used in the Familiarity Rating task were created by combining two syllables from the same word and one syllable from another word. The Nonwords used in the Familiarity Rating and the 2AFC Recognition tasks consisted of syllables from three different words.

Procedure

Word and Sound MSTs

Both Word and Sound MSTs were modeled after the continuous version of the Mnemonic Similarity Task (Stark et al., 2015; Figure 2), which requires participants to make continuous "Old" and "New" recognition judgments to a list of items. Each trial began with a 1.5-second pause followed by the auditory presentation of a word or a sound item along with a prompt on the screen ("New or Old?"). Participants were instructed to label an item as "Old" if the item had been presented before or "New" if the item had never been presented before. They were also specifically instructed that some of the items may sound similar to one another, but that similar-sounding items should also be labelled as "New" if they had not been previously presented. Before the task started, participants underwent five practice trials during which they were given feedback on whether their responses were correct or incorrect.

Statistical Learning Task

Exposure Phase. Participants listened to a 6.84-minute speech stream, which was divided into three 2.28-minute 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 seconds). 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. All participants performed well on this cover task, with no participant missing more than one pause.

Target Detection Task. After the Exposure Phase, participants completed the Target Detection task, designed to indirectly assess participants' 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 make a keypress each time they detected the target syllable. Both speed and accuracy were emphasized.

Prior to starting the task, participants completed two practice trials. A different speech stream made up of three trisyllabic words were 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.

Rating Task. This task was designed to assess participants' explicit knowledge of the words. 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 random order.

2AFC 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 for 1 second before each word was presented, and participants responded with the number associated with the more familiar sounding word. The same four Nonwords used in the Familiarity Rating task were paired exhaustively 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.

Perceptual Similarity Rating Task

BL and the community participants were administered an additional perceptual similarity rating task with the stimuli used for the sound MST to ensure that any performance deficits were not due to difficulties with discriminating the sounds at the perceptual level. The task consisted of a total of 83 pairs of sounds which included 24 lure pairs, 24 "different" pairs in which a sound was paired with another sound (not its lure counterpart) and 35 "identical" pairs in which the same sound was repeated twice. All of the 88 unique sounds were used at least once in this task. Each pair was presented successively with a 1 second interstimulus interval followed by a prompt on the screen asking participants to rate the similarity of the two sounds on a scale of 1 to 4. Participants were instructed to use 1 when the sounds were completely different, 4 when they were identical, and 2 or 3 when they were similar.

Visual MST

In addition to the two auditory MSTs, BL was also tested on the original, visual version of the MST used in Baker et al. (2016) to examine whether his pattern separation for visual objects had

changed since his last testing. The original MST (Stark et al., 2013) was administered in a third separate session that took place four weeks after the testing of the two auditory MSTs using the same protocol as in Baker et al. (2016). To avoid possible practice effects from viewing set "C" in 2016, BL was tested using set "D" in 2021. During the initial study phase, BL viewed 128 images of everyday objects (e.g., toothbrush) while judging whether they were indoor or outdoor items. Later, in a separate test phase, he was instructed to discriminate between previously studied target items (e.g., toothbrush), new items that did not appear in the study phase (e.g., spoon), and lures that were visually and conceptually similar to the studied items (e.g., toothbrush in a different color), by identifying them as "old", "new", or "similar". There were 64 of each stimulus type.

General testing set-up
BL and the community controls were tested in-person over two sessions separated by a 1-week period. The first session consisted of the two MSTs and the perceptual similarity rating task. The second session consisted of the statistical learning tasks. In addition, the controls were also administered a pure-tone audiometry and Montreal Cognitive Assessment in the second session to assess their hearing and neurocognitive status. Participants were asked to verbally provide responses for all the tasks except for the target detection task. The experiment was carried out in a quiet room with an experimenter who provided task instructions at the beginning of each task and recorded subjects' verbal responses. The tasks were administered on a ThinkPad laptop with two portable speakers on either side of the laptop.

The online controls completed the tasks on their own personal computers and provided responses using the keyboard. They were instructed to use headphones to listen to the auditory stimuli. Each task began with a volume adjustment task during which subjects listened to music and adjusted their sound volume to a comfortable level. To ensure that subjects were using working headphones as instructed, a headphone check task adapted from Woods et al. (2017) was administered.

For the entire sample, all tasks were created using PsychoPy 2020.2.10 and were hosted and administered online via Pavlovia.

Data analysis
To test for potential differences between the controls tested in person versus online, t-tests were conducted to compare both control groups in the core analyses; as reported throughout results, no significant group differences were found on any measure. Therefore, data from both control groups were collapsed for comparisons between BL and control participants.

MST
Auditory MSTs. The signal detection theory was used to index sensitivity to Lure and First presentation items while controlling for potential response biases (Leal et al., 2014; Stark et al., 2015; Yassa et al., 2011). Subjects' ability to discriminate similar lures (Lure Discrimination score) was computed by subtracting the probability of responding "New" to any given item from the correct rejection rate for Lure items (p("New"|Lure) – p("New"|Repeat)). Recognition score was computed by the probability of responding "New" to any given item from the correct rejection rate for First presentation items (p("New"|First presentation) – p("New"|Repeat)). BL and controls' performance was compared using a modified t-test designed for comparing a single subject to a small sample (< 50) of controls (Crawford & Howell, 1998). Effect sizes were

estimated as BL's score expressed as z-scores of controls' scores (Z_{cc}), following the method for effect size estimation described in Crawford et al. (2010). Four separate one-tailed t-tests were conducted comparing BL and controls' Lure Discrimination and Recognition scores from the two MSTs.

Visual MST. BL's performance on the original visual MST was analyzed in the same way as in Baker et al. (2016). Pattern separation was measured using the Lure Discrimination Index (LDI; Stark et al., 2013), computed by subtracting the proportion of "Similar" responses to foils from the proportion of "Similar" responses to lures (p("Similar" | Lure) – p("Similar" | Foil)). His general recognition was computed by subtracting the proportion of "Old" responses to foils from the proportion of "Old" responses to repeats (p("Old" | Repeat) – p("Old" | Foil)).

Statistical Learning Task
Target Detection Task. Responses made within 0-1200ms of a target syllable onset were considered as "hits" and were used toward analyses (Batterink et al., 2015; Batterink & Paller, 2017, 2019). All other responses were considered false alarms. Average RT was computed for syllables at the initial, middle, and final triplet position. Controls' RTs were then entered into a repeated-measures ANOVA with triplet position (1-3) and stream position (4 – 45) as within-subject factors; stream position was included as a factor to rule out any potential confounds of stream position on the main triplet effect of interest (Himberger et al., 2019). BL's RTs were analyzed with an item-based ANOVA, with syllable position and stream position as between-item factors, to test for a single subject priming effect. Linear contrasts for syllable position are reported. A significant linear effect of triplet position, with faster RTs to third syllables than first syllables, was considered to indicate RT facilitation.

To compare BL and controls' RT facilitation, a "RT prediction score" was computed by subtracting the average RT to the final triplet position from the average RT to the initial triplet position and dividing the difference by the average RT to the initial triplet position (Batterink & Paller, 2019). A modified one-tailed t-test was then used to compare RT prediction scores between BL and controls (Crawford & Howell, 1998). Furthermore, we tested whether BL's observed RT facilitation was significantly greater than would be expected given the null hypothesis of no statistical learning. As part of this process, for a given iteration, BL's RTs for each of the 144 targets were scrambled across triplet positions and a RT prediction score was calculated based on this scrambled data, using the same formula as above. This process was repeated 1,000 times, generating a null distribution of BL's RT prediction scores. BL was considered to have shown a significant RT facilitation if his observed RT prediction score fell in the 95th percentile or above in this null distribution.

Rating Task. Average rating was computed for each of the three word categories (word, part-word, nonword). For controls, ratings were entered into a repeated-measures ANOVA with word category as a within-subjects factor. BL's ratings were entered into an item-based ANOVA, with word category as a between-item factor. Additionally, a "Word-Partword (W-PW) score" and a "Word-Nonword (W-NW) score" was computed for each individual by taking the difference between the average ratings for Words and Partwords and between the average ratings for Words and Nonwords, respectively. BL's W-PW and W-NW scores were each compared to controls using the modified one-tailed t-test (Crawford & Howell, 1998).

2AFC Recognition Task. A one-sample t-test was conducted to test if the performance of controls was significantly different from chance level (50%). A modified one-tailed t-test (Crawford & Howell, 1998) was conducted to compare the average performance accuracy between BL and controls.

<div align="center">

Results

</div>

MST

On each of the four measures (Lure Discrimination and Recognition scores for both Word MST and Sound MST), BL was the lowest performing individual among all participants (Figure 5).

Word MST
BL demonstrated high accuracy for Repeat items (72%) but numerically lower accuracy for both Lure (36%) and First presentation (44%) items, revealing an overall bias toward responding "Old" (Figure 4A). BL's Lure Discrimination score was marginally lower than that of controls ($t(22) = -1.57$, $p = .065$, $Z_{cc} = -1.60$) and his Recognition score was significantly lower than that of controls ($t(22) = -2.28$, $p = .016$, $Z_{cc} = -2.33$; Figure 5A). There was no significant difference between the two control groups (in-person sample versus online sample) on either Lure Discrimination ($t(21) = 1.11$, $p = .28$) or Recognition ($t(21) = 1.40$, $p = .17$).

In the Word MST, participants were exposed to multiple lures from the same lure family – words that consisted of the same three syllables but combined in different orders (e.g., *gopula* vs. *golapu*). Motivated by the finding in (Forest et al., 2022) that statistical learning can result in order-independent representations of triplet grouping, we explored whether participants formed increasingly strong generalized lure family representations that would become independent of order as they encountered more lures from the same family; such generalized representations would be reflected in increased difficulties in separating lures from the same family in memory as the task progressed. To test this, we fitted a logistic regression model to the controls' old/new responses on the lure trials with the number of prior occurrences of a lure within a given family (1-5), overall trial number, and lure family as predictor variables. Consistent with this idea, Wald test (df = 4) indicated that the number of prior lure occurrences within a lure family significantly increased the likelihood of responding "Old" to a lure ($p = .025$; Table 1).

In contrast, BL's data did not show evidence of increased likelihood of responding "Old" to a lure as more lures from a family were encountered (Table 1). On the contrary, BL's likelihood of responding "Old" was the lowest after five prior lure occurrences within a lure family. However, BL's low accuracy on the lure trials (he erroneously labelled 16/25 lures as old) might have made it challenging to detect any pattern in his data.

Although not present in BL, the finding in controls that mnemonic discrimination became increasingly poorer as more lures from a family were encountered supports the interpretation that general, order-independent representations of triplet membership accrued and strengthened gradually over the course of the task, potentially supported by statistical learning mechanisms.

Table 1. Proportion of "Old" responses per each lure-family occurrence for BL and controls

Prior Lure Occurrence	1	2	3	4	5

Controls	0.24	0.40	0.40	0.45	0.52
BL	0.80	0.60	0.60	0.80	0.40

Sound MST

Similar to performance on the Word MST, BL again demonstrated a response bias to responding "Old", with numerically higher accuracy for Repeat items (85%) than Lure (25%) and First presentation items (47%; Figure 4B). On accuracy measures that control for response bias, BL's performance was significantly poorer than controls for both Lure Discrimination ($t(22) = -1.96$, $p = .031$, $Z_{cc} = -2.00$) and Recognition ($t(22) = -3.10$, $p = .003$, $Z_{cc} = -3.16$; Figure 5B). Again, the two control groups did not differ in Lure Discrimination ($t(21) = 0.14$, $p = .89$) or Recognition ($t(21) = -0.05$, $p = .96$).
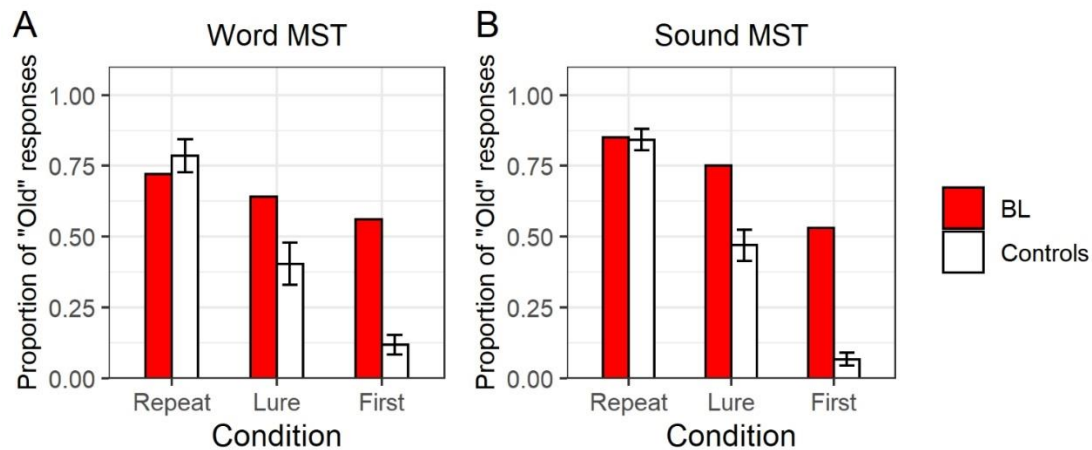

Figure 4. BL and controls' average accuracy rate per each word condition on the Word MST (A) and the Environmental Sound MST (B). Error bars represent 95% confidence intervals.
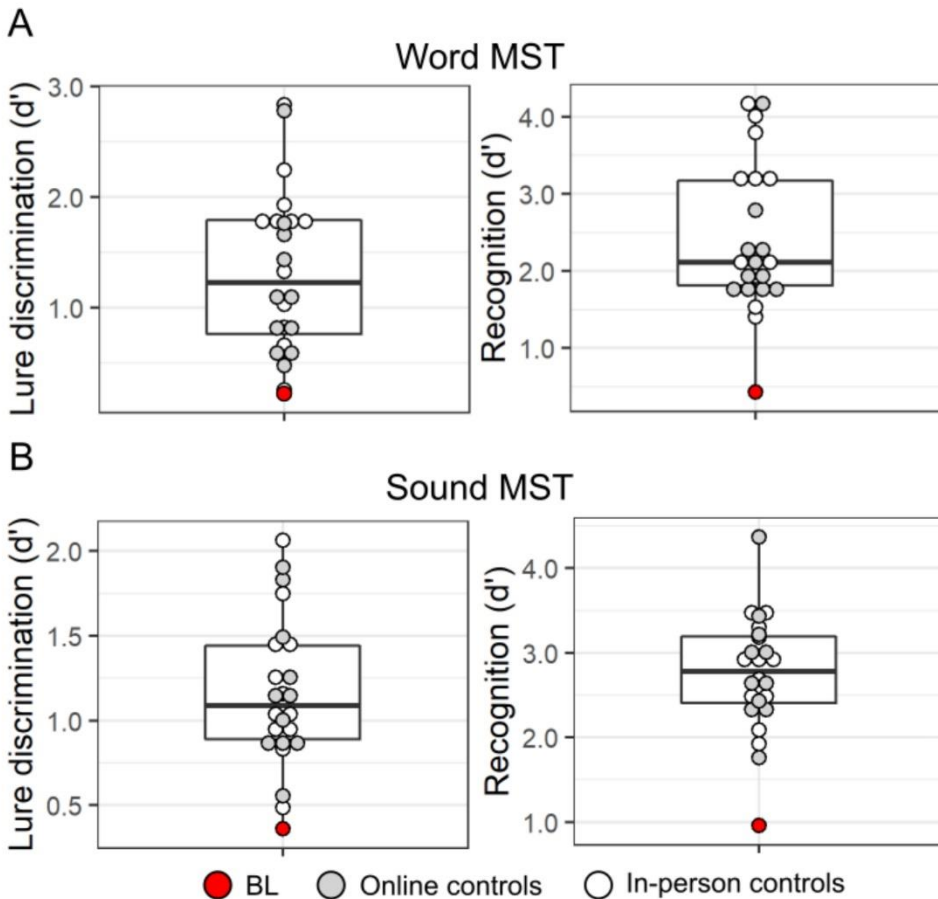
Figure 5. BL and controls' Lure Discrimination and Recognition scores on the Word MST and Sound MST. BL was the lowest performer on each of the four measures.

Visual MST

Since BL was the only subject who completed the original visual MST, BL's performance in 2021 (age 60) was compared to his performance in 2015 (age 54), reported in Baker et al. (2016). When given three response options, BL still demonstrated a bias toward responding "Old" on lure trials, similar to his performance on the auditory MSTs (Figure 6A). However, unlike his previous performance, BL also exhibited a tendency toward responding "Similar" to foil items. This high rate of "Similar" responses on foil trials resulted in a low accuracy (56%) for foils and an LDI of -0.13 that is lower than his LDI in Baker et al. (2016; 0.01; Figure 6B). BL's Recognition score, which is not dependent on the probability of "Similar" responses, showed a slight improvement compared to his earlier performance (Baker et al., 2016).
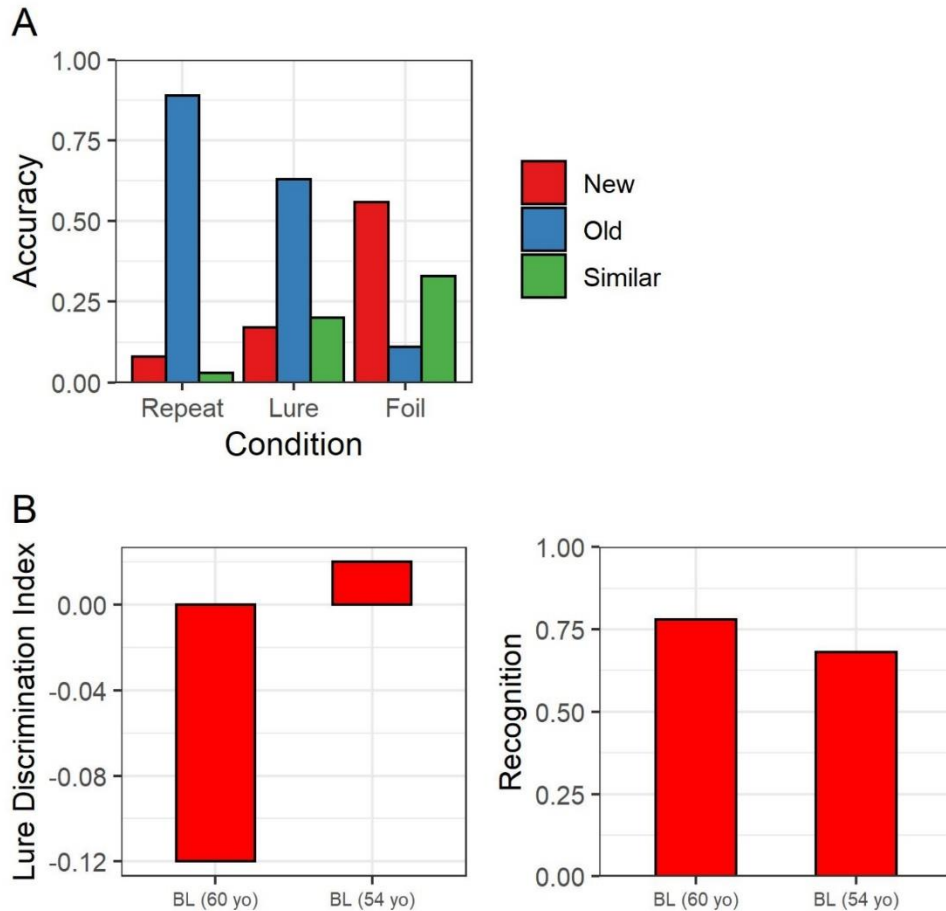
Figure 6. BL's performance in 2021 on the original (visual) version of the MST (Baker et al., 2016). (A) In addition to poor performance on the Lure trials, BL also showed a high tendency to respond "Similar" to Foil trials. (B) After correcting for response bias, BL's LDI was found to be lower in 2021 (60 years old) than in Baker et al. (2016; 54 years old). In contrast, there was a slight increase in BL's Recognition from Baker et al. (2016) to 2021.

Statistical Learning Task

BL was severely impaired on the Rating Task, but performed within normal limits on the 2AFC Recognition Task and the Target Detection Task.

Target Detection Task

BL performed well on the task, with an average hit rate of 84.7% and average false alarm rate of 14.3%. The hit rate and false alarm rate of controls were 86.3% and 10.4%, respectively (Figure 8). While BL's overall RT (average = 635.8 ms) was slower than controls (average = 527.5 ms), he showed a robust statistical learning effect that was comparable to control participants, with both BL and control participants' RTs becoming progressively faster for later-occurring syllables (BL: $F(2, 118) = 3.22$, $p = .043$, $\eta^2 = 0.052$; Linear effect: $F(1, 118) = 6.38$, $p = .013$, $\eta^2 = 0.052$; Controls: $F(2, 3215) = 28.3$, $p < .001$, $\eta^2 = 0.017$; Linear effect: $F(1, 3215) = 32.62$, $p < .001$, $\eta^2 = 0.010$; Figure 7A). For controls, there was also a significant main effect of stream position of the target on RTs such that RTs became progressively slower toward the end of each trial ($F(1$,

3215) = 18.22, $p < .001$, $\eta^2 = 0.006$). Stream position did not significantly predict BL's RTs ($F(1, 118) = 2.08$, $p = .15$). BL's RT prediction score was significantly greater than the score based on the random RT distribution generated from BL's own data (99th percentile), also indicative of significant statistical learning. BL's RT prediction score did not differ from that of controls ($t(25) = 0.91$, $p = .81$). The two control groups' RT prediction scores also did not differ from each other ($t(24) = 0.45$, $p = .65$).

Interestingly, controls' hit rates also increased toward the later triplet positions ($F(2, 908) = 3.53$, $p = .030$), such that the hit rate for the last position was significantly higher than the hit rate for the middle position ($t(908) = -2.58$, $p = .03$). BL's hit rates also increased numerically as a function of syllable position, but this pattern was not statistically significant ($F(2, 33) = 0.55$, $p = .58$; Figure 8).

Rating Task
BL's average rating for Words (3.0) was lower than his average rating for Partwords (3.75) and Nonwords (3.75), providing no evidence for explicit knowledge of statistical regularities acquired during learning. This pattern was opposite to controls' ratings, which showed the expected profile (Batterink & Paller, 2017, 2019) of being highest for Words, followed by Partwords, and lowest for Nonwords ($F(2, 50) = 13.01$, $p < .001$, $\eta^2 = 0.34$; Figure 7B). BL's W-PW and W-NW scores both had a negative value and were both significantly lower than controls' scores (W-PW: $t(25) = -2.23$, $p = .017$, $Z_{cc} = -2.27$; W-NW: $t(25) = -2.36$, $p = .013$, $Z_{cc} = -2.41$). The two control groups did not differ in either W-PW ($t(24) = 1.46$, $p = .16$) or W-NW scores ($t(24) = 0.75$, $p = .46$).

2AFC Recognition Task
Controls performed with an average accuracy of 67.8%, which was significantly greater than chance-level performance ($t(25) = 7.59$, $p < .001$, $d = 1.49$). BL's average accuracy on the recognition task was 62.5% and did not significantly differ from controls ($t(24) = -0.42$, $p = .34$; Figure 7C). There was no difference in accuracy between the two control groups ($t(24) = -0.58$, $p = .57$).
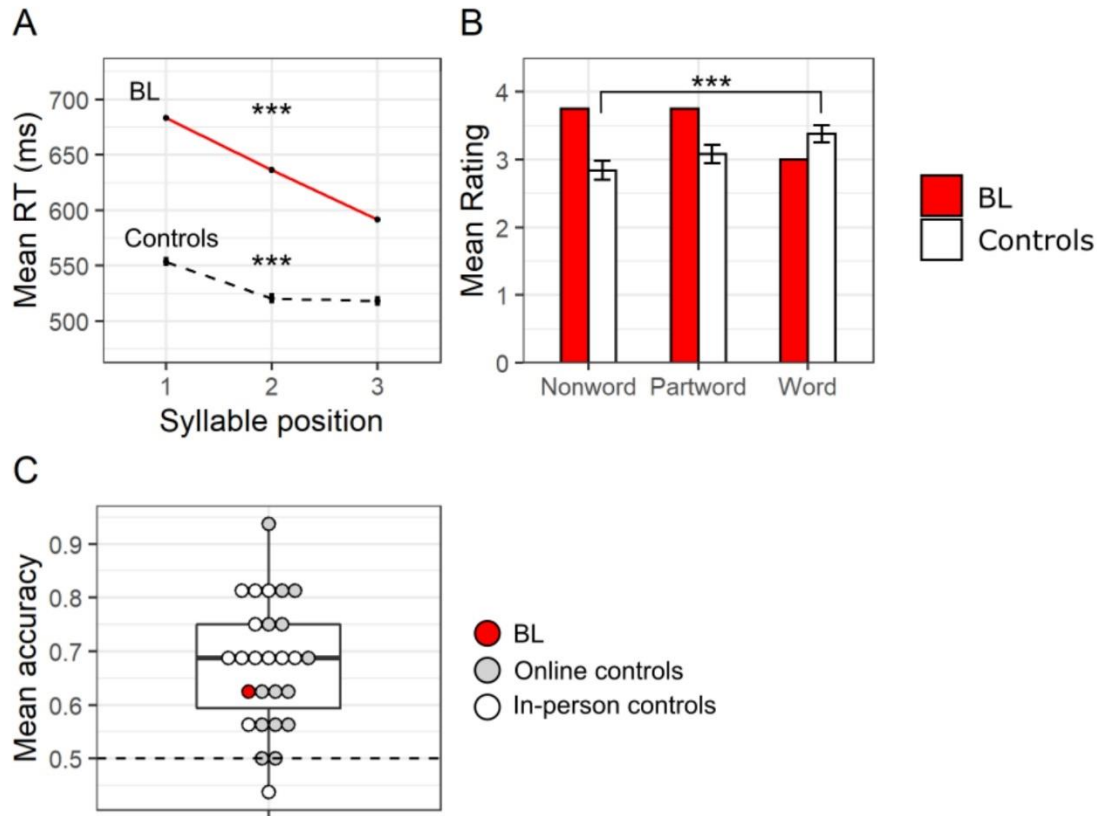
Figure 7. Results of the statistical learning tasks. (A) Average reaction time on the Target Detection task plotted as a function of syllable position. (B) Average familiarity rating to each of the three word categories. (C) BL and controls' average performance accuracy on the 2AFC task. The dashed line denotes at-chance (50%) performance. Error bars represent 95% confidence intervals. *** = p < .001.
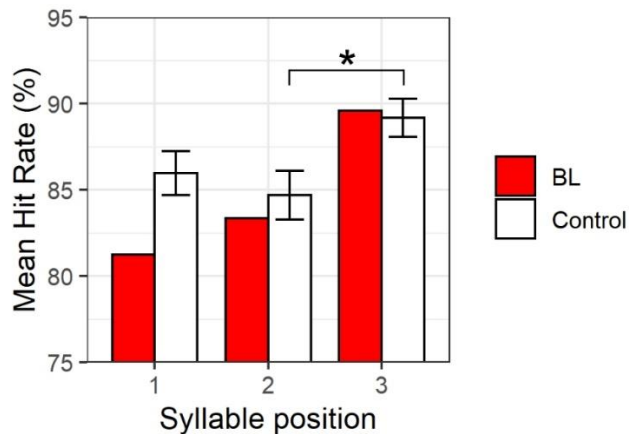

Figure 8. BL and controls' average hit rate per syllable position on the target detection task

Perceptual Similarity Rating Task
Interestingly, on the similarity rating task, BL rated significantly more lure pairs as "exactly the same" (7/24 pairs, 29%) than controls (average = 0.75/24 pairs, 3%; $t(11) = 6.93$, $p < .001$).

To examine whether BL's poor perceptual discrimination of similar sounds (relative to controls) may account for his deficits in pattern separation on our memory task, we excluded all lures that BL failed to discriminate on the perceptual task from the Sound MST analysis. BL's mnemonic discrimination accuracy did not improve even when considering only lures that he successfully discriminated at the perceptual level (25% correct → 23.5% correct), suggesting a deficit in mnemonic discrimination over and above his perceptual impairments.

In summary, BL performed significantly worse than controls on both the Word and the Sound MST as well as on the rating measure of statistical learning, but showed comparable performance to controls on the reaction time-based and the 2AFC measures of statistical learning (Table 2).

Table 2. Summary of BL's performance relative to controls

Pattern Separation				Statistical Learning		
Word MST (Discrimination)	Word MST (Recognition)	Sound MST (Discrimination)	Sound MST (Recognition)	Target Detection	Rating	2AFC Recognition
Impaired	Impaired	Impaired	Impaired	Within normal limits	Impaired	Within normal limits

Discussion

The current study investigated pattern separation and statistical learning in BL, an individual with a highly selective hippocampal lesion in the dentate gyrus (Baker et al., 2016; Kwan et al., 2015). We tested the hypothesis that the dentate gyrus—a key structure of the trisynaptic pathway— supports pattern separation but is not necessary for statistical learning. By and large, our results supported this hypothesis. When exposed to individually presented trisyllabic words, BL showed difficulties in recognizing previously heard words and differentiating previously heard words from similar lures, indicating a deficit in both pattern separation and general recognition. A similar pattern of deficits emerged when BL's pattern separation abilities were assessed with environmental sound items. However, despite his impaired ability to distinctly encode individual episodic events, BL showed successful learning of trisyllabic words embedded within a continuous syllable sequence on the target detection task and the 2AFC task, demonstrating a preserved ability to track statistical regularities across multiple inputs over time. In contrast, BL showed a severe impairment in recognizing and differentiating the learned words on arguably the most explicit measure of statistical learning – the rating task. This task requires explicit high-resolution retrieval of the learned words, and thus may in part rely on pattern separation, making it sensitive to trisynaptic pathway disruption.

The critical role of the dentate gyrus in pattern separation has been supported by a large number of studies (e.g., Berron et al., 2016; Lacy et al., 2011; Yassa et al., 2010), including a recent study on BL himself that revealed his circumscribed deficit in discriminating similar visual objects in memory judgments on the widely used MST developed by Stark and colleagues (Baker et al., 2016; Stark et al., 2013). The current finding that BL was additionally impaired in discriminating two types of auditory stimuli further supports the necessity of the dentate gyrus in pattern separation, and extends its involvement across multiple sensory domains (see also Baker,

2022, for additional evidence concerning auditory pattern separation in BL). The importance of the hippocampus in auditory episodic memory has been suggested in previous work that has revealed impaired auditory recognition memory in patients with hippocampal damage (e.g., Squire et al., 2001). Neuroimaging evidence indicates that different types of visual stimuli (faces, scenes) are represented uniquely in the hippocampus in a category-agnostic manner (Huffman & Stark, 2014), suggesting that the process of pattern separation is invariant to stimulus type, and by extension sensory domain (Kent et al., 2016). However, research on pattern separation outside the visual domain is still scarce (Bjornn et al., 2021; Herman et al., 2020; Trier et al., 2016), and whether the dentate gyrus supports pattern separation in a domain-general manner across modalities remains a topic of continuing investigation.

BL performed comparably to controls on the target detection task and the 2AFC recognition task for word stimuli encountered in our statistical learning paradigm, but did not show evidence of learning on the rating task. Previous studies suggest that indirect and direct measures of statistical learning vary in the level of their dependency on processes of explicit memory retrieval (Batterink et al., 2015; Kiai & Melloni, 2021; Siegelman et al., 2018). The target detection task measures learning without taxing explicit retrieval and has been suggested to capture the implicit knowledge of statistical regularities acquired through new learning (Batterink et al., 2015; Kiai & Melloni, 2021). While the 2AFC task requires explicit recognition judgments, it can be completed by comparing gist-based familiarity for target words and novel nonwords, in the absence of highly precise memory for the words. Of note, our nonword foils contained multiple types of statistics that learners could potentially use to successfully discriminate them from words (Forest et al., 2022; Henin et al., 2021; Park et al., 2018), including triplet membership, ordinal position of individual syllables, and transitional probability (containing syllable pairs that had a transitional probability of 0). Previous evidence indicates that learners may represent all of these different statistical cues, with different brain regions contributing differentially to these representations (Henin et al., 2021). Further, participants can successfully distinguish target words from nonwords even if they have only developed a general association between the three syllables in a word, without knowing the specific order in which they had occurred (Forest et al., 2022; Park et al., 2018). Thus, BL may have succeeded on the current 2AFC task based solely on memory for triplet memberships, without needing to rely on more specific memories of item-to-item transitions. There is evidence to suggest that such familiarity-based recognition can also be achieved based on implicit knowledge (Voss & Paller, 2008) and that it may not require hippocampal integrity (Holdstock et al., 2002).

In contrast, performance on the rating task cannot be supported by comparing the familiarity of words and foils, but depends on the recollection of individually presented items, a process that has been suggested to functionally differ from familiarity-based recognition (Yonelinas, 2002) and that is known to depend on the integrity of the hippocampus (Bowles et al., 2010; Holdstock et al., 2002). Moreover, the rating task in the current study also included partwords, which share more overlap with target words, whereas the 2AFC task involved only nonword foils with multiple, highly distinct statistical cues, as described above. Taken together, the rating task can thus be seen as relying more heavily on the encoding and retrieval of high-precision representations of the learned triplets, and by extension is more likely to be affected by impairments in pattern separation.

Using both implicit and explicit measures of statistical learning, our study extends previous research on the involvement of the hippocampus in statistical learning (Covington et al., 2018; Schapiro et al., 2014). In these two previous studies, patients with hippocampal damage or damage in the broader MTL beyond the hippocampus showed deficits in both visual and auditory statistical learning on the 2AFC measure. By contrast, BL showed intact 2AFC recognition performance, suggesting that recognition of learned triplets may rely on hippocampal regions outside the dentate gyrus. In addition, we also found that the implicit expression of statistical learning is not affected by dentate gyrus lesion. These findings are in line with the proposal that the trisynaptic pathway is not necessary for statistical learning (Schapiro et al., 2017)—at least at the level of implicit or familiarity-based expressions of what has been learnt. However, more explicit, recollection-based expressions of statistical learning may still depend on pattern separation and the integrity of the trisynaptic pathway, as suggested by BL's poor performance on the rating task. A still-open question is whether patients with broader hippocampal damage (such as those studied by Covington, Schapiro and colleagues) would also show intact performance on the target detection task we administered, or whether their observed statistical learning deficits would also extend to performance on implicit measures. Future studies of such patients may shed light on whether the hippocampus is strictly necessary for statistical learning, or whether it merely contributes to learning in healthy populations by acquiring explicitly accessible representations of statistical regularities (Batterink et al., 2019).

The hippocampus is only part of a larger network of regions spanning both cortical and subcortical structures that have been implicated in statistical learning (Batterink et al., 2019; Frost et al., 2015). Neuroimaging studies have reported activation of a number of regions during statistical learning of auditory or visual stimuli, including the striatum, the inferior frontal gyrus, and sensory-related processing areas such as the occipital cortex and the superior temporal gyrus (Karuza et al., 2013; McNealy et al., 2006; Moser et al., 2021; Sandoval et al., 2017; Schapiro et al., 2017; Turk-Browne et al., 2009). In Henin et al. (2021), intracranial recordings during statistical learning were used to group individual syllables based on the similarity of the neural activity evoked by each syllable. The study found evidence of triplet-based syllable organization within the hippocampus and lower feature (e.g., ordinal position)-based syllable organization in other cortical regions, suggesting that different regions form complementary representations of regularities during statistical learning (Henin et al., 2021). In particular, the authors suggest that prediction-based coding may occur within cortical areas, while the role of the hippocampus may be in accurately and flexibly representing the identity of the extracted units, supporting further use of these representations in various cognitive operations. BL's profile of robust reaction-time facilitation but impaired explicit rating performance are in line with this idea. Similarly, Covington and colleagues' hippocampal-lesioned patients showed low yet above chance recognition performance, while temporal-lobe epilepsy patients performed no better than chance on the explicit 2AFC task but showed intact performance on a reaction time-based online task (Henin et al., 2021). These results suggest that statistical learning can occur without hippocampal involvement. In the current study, there is no direct evidence that the hippocampus is involved in any of the statistical learning tasks we employed. Future neuroimaging research can investigate the direct involvement of different hippocampal subregions in statistical learning to further our understanding of the exact nature of hippocampal contributions to statistical learning, including for the specific task employed in the current study.

Interestingly, when pairs of similar environmental sounds (e.g., two different types of cow moos) were presented in direct succession during the perceptual similarity rating task, BL rated more pairs as being "exactly the same" than controls. Recent studies have reported that BL shows a deficit in perceptually discriminating simultaneously presented similar faces and novel complex objects (Baker et al., 2020; Mitchnick et al., 2022), supporting the purported role of the hippocampus, and in particular the dentate gyrus, not just in mnemonic but also in perceptual discrimination (Barense et al., 2007; Erez et al., 2013; Inhoff et al., 2019; Lee et al., 2005). The findings in the current study hint that the role of the dentate gyrus in perceptual discrimination extends to the auditory domain as well. However, while BL performed poorly in differentiating similar sounds at both perceptual and mnemonic levels, our analyses suggests that his impaired pattern separation performance on our memory task cannot be fully accounted for by his difficulties in auditory perceptual discrimination. Even when considering only lures that BL successfully differentiated at the perceptual level, BL's mnemonic discrimination accuracy remained impaired.

In contrast to earlier reports with the visual MST (Baker et al., 2016), it should be noted that BL also scored significantly lower than controls on general recognition across the two pattern separation tasks. This was primarily driven by his tendency to respond "Old" to first presentation items. BL's accuracy for first presentation trials was 44% on the Word MST and 47% on the Sound MST. To determine whether this poor recognition performance was specific to the auditory stimuli used in the current study, we also administered the same visual mnemonic similarity task used in Baker et al. (2016). On this task, BL's accuracy for novel, unrelated foil trials was 56%, this time due to his tendency to respond "Similar" to novel items. As a result, BL's current Lure Discrimination Index (correct rejection rate for lure items minus the probability of responding "similar" to foils; -0.13) decreased from the value reported by Baker et al. (2016); LDI = 0.01; Figure 6B), but his Recognition Index (correct rejection rate for target items minus the probability of responding "old" to foils) did not show any decrease. The variation in response bias exhibited by BL across tasks, and across multiple assessments with the same task, poses a challenge for determining whether the changes observed reflect a true decline in his task performance. Additionally, at the time of testing for the current study, BL was 60 years old and within the age range in which age-related pattern separation deficits have been reported (e.g., Stark et al., 2015). He was also newly diagnosed with cancer. It is also possible that the process of normal aging along with changes in his health status has further exacerbated BL's deficits in pattern separation and by extension affected general recognition. Lastly, given our current design, we cannot be sure whether BL's pattern of impairment and preservations across our tests of statistical learning would extend to different stimulus types and modalities. Prior evidence suggests that statistical learning shows distinct patterns based on modality and stimulus type (Raviv & Arnon, 2017; Siegelman et al., 2018; Siegelman & Frost, 2015), and additional studies will be necessary to understand the contributions of the dentate gyrus to other types of statistical learning.

Overall, this study sheds light on the role of the dentate gyrus in pattern separation and statistical learning, providing support for a previously proposed neural computational model of these two processes (Schapiro et al., 2017). We found that, consistent with previous literature, the dentate gyrus plays an essential role in pattern separation across both visual and auditory stimuli. On the other hand, statistical learning—when probed with tasks that require implicit or low-resolution, familiarity-based expression of knowledge—can be maintained in the absence of dentate gyrus

integrity and the trisynaptic pathway of which it is a critical component. However, a measure of statistical learning that relies upon explicit, high-precision memory retrieval was also disrupted by dentate gyrus impairment and thus may rely on trisynaptic pathway-dependent memory processes. We conclude that statistical learning operates largely independently of pattern separation mechanisms, but that both processes may be called upon when high-resolution statistical memory representations are needed.

Data availability statement

All stimuli, scripts, and data can be found online at: https://github.com/helenaszwang/MST-SL-tasks-and-data.

Author contribution

Helena S. Wang; Data Curation; Formal Analysis; Investigation; Methodology; Project Administration; Visualization; Writing – Original Draft Preparation. R. Shayna Rosenbaum. Methodology; Supervision; Writing – Review & Editing. Stevenson Baker. Methodology; Resources. Claire Lauzon. Investigation. Laura J. Batterink. Conceptualization; Methodology; Resources; Supervision; Writing – Review & Editing. Stefan Köhler. Conceptualization; Methodology; Resources; Supervision; Writing – Review & Editing.

This work was jointly overseen by Laura Batterink and Stefan Köhler.

Acknowledgements

Funding information

References

Baker, S. (2022). *Understanding the nature, extent, and brain dynamics of deficient pattern separation.* [Doctoral dissertation, York University].

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

Baker, S., Youm, A., Levy, Y., Moscovitch, M., & Rosenbaum, R. S. (2020). *Pattern Separation Contributes to Categorical Face Perception* [Preprint]. Neuroscience. https://doi.org/10.1101/2020.05.13.094276

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

Barense, M. D., Gaffan, D., & Graham, K. S. (2007). The human medial temporal lobe processes online representations of complex objects. *Neuropsychologia*, *45*(13), 2963–2974. https://doi.org/10.1016/j.neuropsychologia.2007.05.023

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., Paller, K. A., & Reber, P. J. (2019). Understanding the Neural Bases of Implicit and Statistical Learning. *Topics in Cognitive Science*, tops.12420. https://doi.org/10.1111/tops.12420

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

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

Bjornn, D. K., Van, J., & Kirwan, C. B. (2021). *The Contributions of Eye Gaze Fixations and Target-Lure Similarity to Behavioral and fMRI Indices of Pattern Separation and Pattern Completion* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.11.20.469401

Bowles, B., Crupi, C., Pigott, S., Parrent, A., Wiebe, S., Janzen, L., & Köhler, S. (2010). Double dissociation of selective recollection and familiarity impairments following two different surgical treatments for temporal-lobe epilepsy. *Neuropsychologia*, *48*(9), 2640–2647. https://doi.org/10.1016/j.neuropsychologia.2010.05.010

Canada, K. L., Hancock, G. R., & Riggins, T. (2021). Modeling longitudinal changes in hippocampal subfields and relations with memory from early- to mid-childhood. *Developmental Cognitive Neuroscience*, *48*, 100947. https://doi.org/10.1016/j.dcn.2021.100947

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

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The Necessity of the Hippocampus for Statistical Learning. *Journal of Cognitive Neuroscience*, *30*(5), 680–697. https://doi.org/10.1162/jocn_a_01228

Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect

    sizes for the case-controls design in neuropsychology: Rationale, methods,

    implementations, and proposed reporting standards. *Cognitive Neuropsychology*, *27*(3),

    245–260. https://doi.org/10.1080/02643294.2010.513967

Crawford, J. R., & Howell, D. C. (1998). Comparing an Individual's Test Score Against Norms

    Derived from Small Samples. *The Clinical Neuropsychologist*, *12*(4), 482–486.

    https://doi.org/10.1076/clin.12.4.482.7241

Ellis, C. T., Skalaban, L. J., Yates, T. S., Bejjanki, V. R., Córdova, N. I., & Turk-Browne, N. B.

    (2021). Evidence of hippocampal learning in human infants. *Current Biology*, *31*(15),

    3358-3364.e4. https://doi.org/10.1016/j.cub.2021.04.072

Erez, J., Lee, A. C. H., & Barense, M. D. (2013). It does not look odd to me: Perceptual

    impairments and eye movements in amnesic patients with medial temporal lobe damage.

    Neuropsychologia, *51*(1), 168–180.

    https://doi.org/10.1016/j.neuropsychologia.2012.11.003

Finn, A. S., Kharitonova, M., Holtby, N., & Sheridan, M. A. (2019). Prefrontal and Hippocampal

    Structure Predict Statistical Learning Ability in Early Childhood. *Journal of Cognitive*

    Neuroscience, *31*(1), 126–137. https://doi.org/10.1162/jocn_a_01342

Forest, T. A., Finn, A. S., & Schlichting, M. L. (2022). General precedes specific in memory

    representations for structured experience. *Journal of Experimental Psychology: General*,

    151(4), 837–851. https://doi.org/10.1037/xge0001104

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

Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., Doyle, W., Devinsky, O., & Melloni, L. (2021). Learning hierarchical sequence representations across human cortex and hippocampus. *Science Advances*, *7*(8), eabc4530. https://doi.org/10.1126/sciadv.abc4530

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

Himberger, K., Finn, A., & Honey, C. J. (2019). *Reconsidering the Automaticity of Visual Statistical Learning*. PsyArXiv. https://doi.org/10.31234/osf.io/r659w

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

Huffman, D. J., & Stark, C. E. L. (2014). Multivariate pattern analysis of the human medial temporal lobe revealed representationally categorical cortex and representationally agnostic hippocampus. *Hippocampus*, *24*(11), 1394–1403. https://doi.org/10.1002/hipo.22321

Inhoff, M. C., Heusser, A. C., Tambini, A., Martin, C. B., O'Neil, E. B., Köhler, S., Meager, M. R., Blackmon, K., Vazquez, B., Devinsky, O., & Davachi, L. (2019). Understanding perirhinal contributions to perception and memory: Evidence through the lens of selective perirhinal damage. *Neuropsychologia*, *124*, 9–18. https://doi.org/10.1016/j.neuropsychologia.2018.12.020

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

Kent, B. A., Hvoslef-Eide, M., Saksida, L. M., & Bussey, T. J. (2016). The representational–

hierarchical view of pattern separation: Not just hippocampus, not just space, not just

memory? *Neurobiology of Learning and Memory*, *129*, 99–106.

https://doi.org/10.1016/j.nlm.2016.01.006

Kiai, A., & Melloni, L. (2021). *What canonical online and offline measures of statistical*

learning can and cannot tell us [Preprint]. Neuroscience.

https://doi.org/10.1101/2021.04.19.440449

Kwan, D., Craver, C. F., Green, L., Myerson, J., Gao, F., Black, S. E., & Rosenbaum, R. S.

(2015). Cueing the personal future to reduce discounting in intertemporal choice: Is

episodic prospection necessary? *Hippocampus*, *25*(4), 432–443.

https://doi.org/10.1002/hipo.22431

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(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

Lee, A. C. H., Buckley, M. J., Pegman, S. J., Spiers, H., Scahill, V. L., Gaffan, D., Bussey, T. J.,

Davies, R. R., Kapur, N., Hodges, J. R., & Graham, K. S. (2005). Specialization in the

medial temporal lobe for processing of objects and scenes. *Hippocampus*, *15*(6), 782–

797. https://doi.org/10.1002/hipo.20101

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

Malykhin, N. V., Lebel, R. M., Coupland, N. J., Wilman, A. H., & Carter, R. (2010). In vivo

quantification of hippocampal subfields using 4.7 T fast spin echo imaging. *NeuroImage*,

49(2), 1224–1230. https://doi.org/10.1016/j.neuroimage.2009.09.042

Marr, D., Willshaw, D., & McNaughton, B. (1971). Simple Memory: A Theory for Archicortex.

In L. Vaina (Ed.), *From the Retina to the Neocortex: Selected Papers of David Marr* (pp.

59–128). Birkhäuser. https://doi.org/10.1007/978-1-4684-6775-8_5

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),

419–457. https://doi.org/10.1037/0033-295X.102.3.419

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

Mitchnick, K. A., Ahmad, Z., Mitchnick, S. D., Ryan, J. D., Rosenbaum, R. S., & Freud, E.

(2022). Damage to the human dentate gyrus impairs the perceptual discrimination of

complex, novel objects. *Neuropsychologia*, *172*, 108238.

https://doi.org/10.1016/j.neuropsychologia.2022.108238

Moser, J., Batterink, L., Li Hegner, Y., Schleger, F., Braun, C., Paller, K. A., & Preissl, H.

(2021). Dynamics of nonlinguistic statistical learning: From neural entrainment to the

emergence of explicit knowledge. *NeuroImage*, *240*, 118378.

https://doi.org/10.1016/j.neuroimage.2021.118378

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I.,

Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA:

A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American

Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

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

Park, S. H., Rogers, L. L., & Vickery, T. J. (2018). The roles of order, distance, and interstitial

items in temporal visual statistical learning. *Attention, Perception, & Psychophysics*,

80(6), 1409–1419. https://doi.org/10.3758/s13414-018-1556-1

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

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. (1996). Word Segmentation: The Role of

Distributional Cues. *Journal of Memory and Language*, *35*(4), 606–621.

https://doi.org/10.1006/jmla.1996.0032

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., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The Necessity of the Medial Temporal Lobe for Statistical Learning. *Journal of Cognitive Neuroscience*, *26*(8), 1736–1747. https://doi.org/10.1162/jocn_a_00578

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

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. https://doi.org/10.1162/jocn_a_01028

Scoville, W. B., & Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11–21.

Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, *32*, 15–20. https://doi.org/10.1016/j.cobeha.2020.01.015

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., Kronenfeld, O., & Frost, R. (2018). Redefining "Learning" in Statistical Learning: What Does an Online Measure Reveal About the Assimilation of Visual Regularities? *Cognitive Science*, *42*(S3), 692–727. https://doi.org/10.1111/cogs.12556

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001

Squire, L. R., Schmolck, H., & Stark, S. M. (2001). Impaired Auditory Recognition Memory in Amnesic Patients with Medial Temporal Lobe Lesions. *Learning & Memory*, *8*(5), 252–256. https://doi.org/10.1101/lm.42001

Squire, L. R., & Zola-Morgan, S. (1991). The Medial Temporal Lobe Memory System. *Science, New Series*, *253*(5026), 1380–1386.

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

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., 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

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., &

Mishkin, M. (1997). Differential Effects of Early Hippocampal Pathology on Episodic

and Semantic Memory. *Science*, *277*(5324), 376–380.

https://doi.org/10.1126/science.277.5324.376

Voss, J. L., & Paller, K. A. (2008). Brain substrates of implicit and explicit memory: The

importance of concurrently acquired neural signals of both memory types.

Neuropsychologia, *46*(13), 3021–3029.

https://doi.org/10.1016/j.neuropsychologia.2008.07.010

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., 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

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of

Research. *Journal of Memory and Language*, *46*(3), 441–517.

https://doi.org/10.1006/jmla.2002.2864