Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Age impairs mnemonic discrimination of objects more than scenes: A web-based, large-scale approach across the lifespan

Jeremie Güsten [a,b,*], *Gabriel Ziegler* [a,b], *Emrah Düzel* [a,b,d] and *David Berron* [a,c]

[a] *German Center for Neurodegenerative Diseases, Magdeburg, Germany*
[b] *Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke University, Magdeburg, Germany*
[c] *Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Lund, Sweden*
[d] *Institute of Cognitive Neuroscience, University College London, London, United Kingdom*

ARTICLE INFO

Article history:
Received 28 June 2020
Reviewed 29 July 2020
Revised 13 October 2020
Accepted 5 December 2020
Action editor Asaf Gilboa
Published online 19 January 2021

Keywords:
Mnemonic discrimination
Aging
Memory

ABSTRACT

Recent findings suggest that the effect of aging on recognition memory is modality-dependent, affecting memory for objects and scenes differently. However, the lifespan trajectory of memory decline in these domains remains unclear. A major challenge for assessing domain-specific trajectories is the need to utilize different types of stimuli for each domain (objects and scenes). We tested the large sample required to cover much of the adult lifespan using a large stimulus range via web-based assessments. 1554 participants (18–77 years) performed an online mnemonic discrimination task, tested on a pool of 2708 stimuli (Berron et al., 2018). Using corrected hit-rate (*Pr*) as a measure of performance, we show age-related decline in mnemonic discrimination in both domains, notably with a stronger decline in object memory, driven by a linear increase in the false recognition rate with advancing age. These data are the first to identify a linear age-related decline in mnemonic discrimination and a stronger, linear trajectory of decline in the object domain. Our data can inform basic and clinical memory research on the effects of aging on memory and help advancing the implementation of digital cognitive research tools.

1. Introduction

The study of memory and aging has taken on particular significance due to the increased longevity resulting from an improved ability to extend life in the face of major diseases. Neurodegenerative diseases such as Alzheimer's disease (AD) have become a major cause of death and increased costs for society, sparking a quest to find early and easily accessible markers of memory decline and age-related pathology. When trying to detect significant changes in memory in elderly people, mnemonic discrimination is a promising candidate as it is known to be particularly sensitive to age-related cognitive

* *Corresponding author*. Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke University, Magdeburg, Germany.
E-mail address: jeremie.guesten@med.ovgu.de (J. Güsten).
https://doi.org/10.1016/j.cortex.2020.12.017

decline (Holden, Toner, Pirogovsky, Kirwan, & Gilbert, 2013). In mnemonic discrimination tasks, older adults are typically more prone to falsely recognize similar stimuli as old, with no difference in the correct identification of old stimuli, which is often explained with a bias towards pattern completion in older age (Stark, Stevenson, Wu, Rutledge, & Stark, 2015; Vieweg, Stangl, Howard, & Wolbers, 2015; Yassa et al., 2011). Moreover, memory decline might vary depending on information content. The Posterior Medial Anterior Temporal (PMAT) framework (Ranganath & Ritchey, 2012; Ritchey, Libby, & Ranganath, 2015) defines two information processing and memory pathways in the brain, which receive input from two separate visual streams: A posterior-medial system (PM), including parahippocampal (PHC) and retrosplenial (RSC) cortex, involved in the processing of and memory for spatial layouts, context and scenes; and an anterior-temporal system (AT), including the perirhinal cortex (PRC), predominantly involved in the processing and memory of objects and content. Importantly, the transentorhinal region, which entails part of the PRC and entorhinal cortex (ERC), and subserves object and content information processing, is among the earliest regions affected by tau accumulation (Braak & Braak, 1991). In a recent paper using tau positron emission tomography (PET), Maass et al. (2019) showed that the anterior-temporal system was predominantly affected by tau accumulation, which in turn was related to age-related deficits in mnemonic discrimination of objects. Similarly, cerebrospinal fluid (CSF) levels of phosphorylated tau have been shown to be related to object mnemonic discrimination in cognitively unimpaired older adults (Berron et al., 2019). Taken together, the pattern of tau deposition in aging would suggest an earlier decline in object memory as compared to spatial or scene memory. Recently, several studies have focused on the investigation of age-related decline as a function of stimulus domain. In rodents, Johnson et al. (2017) reported greater impairment in mnemonic discrimination of objects than spatial configurations in aged rats. Other groups have developed tasks to directly assess age-related decline in mnemonic discrimination of objects and spatial configurations in humans. Reagh et al. (2018) used a task in which stimuli either changed in terms of their identity or location. The authors reported a weaker ability for correct rejections of objects than spatial configurations in the elderly. In line with these findings, Stark and Stark (2017) observed generally worse discrimination of similar objects as well as scenes in the elderly group. Since age-related decline in object and scene discrimination was tested in independent models, the question of preferential decline for objects could not be answered directly, however. In fact, the question of domain-specific decline is far from being resolved and certainly warrants close examination, as age-related decrease has been reported for spatial memory in a host of other studies (Borella, Meneghetti, Ronconi, & Beni, 2014; Newman & Kaszniak, 2000; Rosenbaum, Winocur, Binns, & Moscovitch, 2012). Using the present object-scene task, Berron et al. (2018) found no age interaction for the mnemonic discrimination ability of objects versus scenes. However, they observed an imbalance in domain-specific blood-oxygen-level-dependent (BOLD) activity (scene minus object activity) in the PRC of older adults. It should be noted that the object-scene task (Berron et al., 2018) differs from other mnemonic discrimination tasks such as the ones mentioned above (Reagh et al., 2018; Stark & Stark, 2017). Importantly, stimulus presentation and test phases are alternating in blocks of 2 stimuli, whereas the learning and test phase are separated by several minutes in the former tasks.

Previous research on mnemonic discrimination exhibits two major shortcomings regarding the generalizability of domain-specific effects: the limited size of stimulus sets as well as the age groups under study. Regarding the former, the small size of stimulus sets used in previous studies might limit the generalizability of the observed effects. For the latter, age-related domain-specific memory decline has so far mostly been studied in a categorical fashion, contrasting young and old participants, even though the underlying decline is a fundamentally continuous process (Nyberg, Lövden, Riklund, Lindenberger, & Bäckman, 2012). One open issue with testing various item sets is that it is expected to require significantly larger sample sizes than can typically be recruited in a lab-based setting. This problem is even more apparent when studying the entire adult age-range rather than confined age groups. A promising and increasingly used remedy to the problems of small samples is web-based research. In the social sciences, recent years have already seen a surge of recruitment via crowdsourcing platforms such as Amazon Mechanical Turk (Mturk), with a rise of yearly publications from 50 to 500 between 2011 and 2015 (Chandler & Shapiro, 2016). Other than the opportunity to collect data from large samples from a pool of roughly 7000 active workers, Mturk also offers access to more demographically diverse samples than typical undergraduate populations (Buhrmester, Kwang, & Gosling, 2011; Casler, Bickel, & Hackett, 2013). Furthermore, Mturk data quality in terms of participants' attention to the task seems to be equal or partially even better than in typical studies using undergraduates (Chandler & Shapiro, 2016; Hauser & Schwarz, 2016). However, the benefits of large pools of online data do not come without potential shortcomings, which will be discussed below.

In the present study, we investigated the phenomenon of age-related, domain-specific mnemonic discrimination decline in a web-based large-scale sample across a very wide age-range. We addressed methodological limitations of lab-based studies that use small item sets by presenting stimulus subsets out of a very large item pool. The present study aims to replicate and expand on lab-based findings regarding age-related domain-specific memory decline in a web-based sample that largely differs in size and diversity from previous studies. Using a version of the object-scene task (Berron et al., 2018) and treating age as a continuous variable, we investigated domain-specific age-related change. Based on the existing literature, we hypothesized that accuracy for object discrimination would decline in a steeper fashion than scene discrimination across age. We also tested the modalityindependent hypothesis that the ability to reject similar lures, specifically, declines in aging.

2. Methods

2.1. Participants

Participants were recruited through the online crowdsourcing platform Amazon Mechanical Turk (Mturk) (https://www.mturk.com/). The experiment advert was hosted via Psiturk (https://www.psiturk.org/), a platform that facilitates performing behavioral experiments on Mturk. Each participant was asked to provide demographic information such as age, handedness and gender prior to the experiment. Before performing the task, participants were instructed to pay attention to the task and informed that insufficient engagement with the task could lead to the exclusion from the study. After excluding participants based on criteria reported below, a sample of 1554 adults (18–77 years, $M = 37.19$, $SD = 11.61$, 61% females) was analyzed. Sample size was previously determined based on the minimum amount of evaluations that we wished to collect for each stimulus (pair). In accordance with general pay on the platform, a remuneration of 2.50$ was paid upon successful participation. The study was approved by the ethics committee of the Otto-von-Guericke University, Magdeburg. All subjects gave informed consent for their participation by explicitly ticking an "I agree" box on the consent form, in accordance with ethics and data security guidelines of the Otto-von-Guericke University.

2.2. Materials and stimuli

The stimuli consisted of computer-generated (3ds Max, Autodesk Inc., San Rafael, USA) and isoluminant images depicting everyday indoor objects or scenes (1486 and 1222 stimuli respectively). Most images were paired into two very similar versions that only differed in specific features, for the objects, or the spatial configuration or shape of boundaries, for the scenes. For a subset of stimuli, object (17.8%) and scene stimuli (4.4%) only consisted of one version, without a similar counterpart. In order to have a better control of balanced stimulus exposure, we randomly divided the entirety of stimuli into 10 distinct sets. This process was repeated 3 times, resulting in a total of 30 sets. The sets were not all equal in length, given that the amount of pairs and single stimuli for objects and scenes differed and we aimed to make balanced use of all available stimuli. In order to keep the repeat/lure-ratio and the object/scene-ratio within a close range across sets, and further making sure that all stimuli were presented equally often, we had to set up an elaborate partition of sets. An item set would thereby define the set of scene and object items, which themselves consisted of either a pair of similar stimuli or a single stimulus. During test, one of four random sequences was used to assign the role of lure or repeat to every stimulus pair, while single stimuli were always shown as repeats. The balanced combination of test sets with randomized sequences was achieved via the inbuilt Psiturk functionality, which keeps track of and counterbalances database entries. Each participant was tested on 72–76 objects and 58–60 scenes, the exact amount depending on the stimulus set that was assigned. The focus in the present study was to test participants on a maximum amount of different stimuli. While this meant not having equally difficult sets for scenes and objects, we believe that the diversity of stimuli employed strengthens the generalizability of the results. That is, demonstrating that a domain-effect emerges across a large pool of different stimuli reduces the risk of obtaining stimulus-specific effects unrelated to their domain, and increases the chances of replicability.

2.3. Task and procedures

The task, which had a total duration of up to 30 min, was adapted from Berron et al. (2018) and consisted of a 2-back design (Fig. 1). In trials of 4 stimuli, participants were shown pictures of either objects or scenes. The first 2 stimuli (presentation phase) of a trial were always new images, whereas the following 2 (test phase) could be either an exact repetition (repeat) or a very similar version of the previous ones (lure). Stimuli were presented for 3 sec each, separated by a blank page of 1 sec. The end of a trial was marked by a 1 sec fixation cross, followed by .5 sec blank page. Subjects had to respond to each test stimulus with old/new judgments pressing the left or right arrow key. In order to avoid widely differing answer strategies, participants were asked to respond both as correctly and as fast as possible. Object and scene trials contained, respectively, only object or scene stimuli. Two optional breaks of 60 sec were included for participants to rest. There were 4 possible trial types: two test types ("First – Repeat"/"First – Lure"); and two presentation orders, where either of two picture versions could be shown as a "First". The trial types were then: Version 1 – Repeat, Version 1 – Lure, Version 2 – Repeat, Version 2 – Lure. In order to balance the probability for a given stimulus to be presented as a "repeat" or "lure", 4 randomized sequences were created for the paired stimuli, to which participants were pseudo-randomly assigned. Using only paired stimuli would therefore have led to an equal distribution of all trial types. But owing to additional unpaired stimuli, further "First-Repeat" trials were included, the amount of total trials therefore varying as a function of the stimulus set. In order to test stimuli in both directions, an additional set of "mirror" sequences were added that reversed the above sequences in terms of presentation order. In order to avoid confounds of stimulus order, paired and unpaired stimuli were pseudo-randomly assigned to trials. In addition, trial order was shuffled, creating a pseudo-random sequence of object and scene trials of the four different trial types. Although we used the object-scene task from Berron et al. (2018), there were the following differences: The item pool used in the present study was considerably larger. Related to this, the task length could vary as a function of the item set, whereas item set and duration were fix in Berron et al. (2018). We instructed participants not to respond

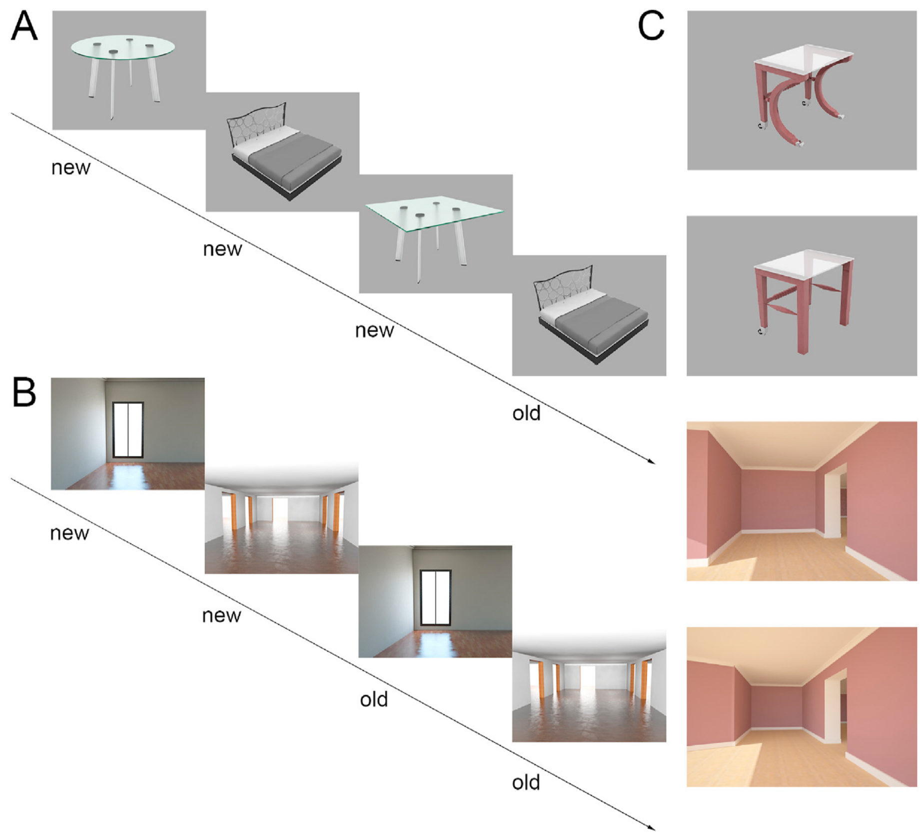

Fig. 1 – Task sequence and stimuli. Sequences used during the object and scene paradigm. For each trial, 2 object (A) or scene (B) stimuli were either identically repeated (correct response: old) or presented again in a very similar but not identical version (correct response: new). Lure and repetition stimuli only differed in shape or geometry (C).

during a presentation phase, while they had been instructed to respond "new" also in this phase in the original task. This allowed participants to unequivocally indicate presentation phases. We also shortened the pauses between stimulus presentations, from 1.63 sec on average to 1 sec, and between trials, from 2.43 sec on average to 1.5 sec. Pauses between stimuli were not jittered, as such an optimization with regard to fMRI-analysis was not necessary in the present task. Finally, we indicated within-trial pauses (blank page) differently from between-trial pauses (fixation cross followed by blank page), in order to help orientation during the task. After finishing the task, participants had to fill out a short questionnaire asking about the perceived difficulty of the task, the correct display of all stimuli as well as the overall time the experiment took them. Prior to the experiment, participants were given detailed instructions and had to complete a short training session in order to get familiarized with the task. Furthermore, a simple arrow task asking to discriminate between a set of left and right arrows was administered, in order to assess baseline reaction time (RT) and check for any obvious visual impairments (Stark & Squire, 2001).

2.4. *Data preprocessing and exclusion*

While recent work has shown that data collected via crowdsourcing platforms such as Mturk can provide easy access to high quality data, it has been emphasized that rigorous (data-based) exclusion criteria need to be applied (Downs, Holbrook, Sheng, & Cranor, 2010; Thomas & Clifford, 2017). Inclusion/exclusion criteria were established following data collection and based on previous literature as well as performance measures considered to be reasonable by the authors (see below). First, following previous work on memory recognition, we excluded all individual trials that had RTs below 200 msec (Berry, Shanks, Speekenbrink, & Henson, 2012). We collected a total of 2352 datasets, of which 1554 were finally analyzed, due to exclusions for different reasons: in a first step, 60 data sets had to be discarded because either subject information was missing, or performance could not be computed due to a lack of response data. Furthermore, the very simple arrow task was used as a criterion, with more than one active mistake leading to the exclusion. Due to technical problem, arrow task data from 122 participants were not available. In general, missing responses were not considered and did not lead to exclusion. Moreover, we excluded data based on performance measures, taking the median of the absolute deviations from the data's median (MAD) to measure variance. The exclusion threshold was set at the standard value of 3 MAD, meaning that participants exceeding this value on any of the given measures were discarded, and calculated with the R package univOutl (D'Orazio, 2019). The following measures were used: percent of wrong answers during the presentation phase ("new" responses were accepted), the ratio of "old" versus "new" responses, task performance in corrected hits, percent of total responses. The above criterion was applied concurrently to all measures (failing on multiple criteria was possible), and the number of participants not meeting them was distributed as follows: arrow task (121), "old/new"-ratio (75), wrong presentation responses (432), corrected hits (20), total responses (339). Finally, two participants were excluded for pressing only one button.

2.5. *Statistical analysis*

As performance measures, hit rates (HR), false alarm rates (FAR) and corrected hit rates (hit rates − false alarm rates) were calculated for the object and scene conditions separately. The corrected hit-rate sensitivity measure is also known as *Pr* and provides an unbiased measure of old–new discrimination, with higher values corresponding to more accurate recognition memory (Snodgrass & Corwin, 1988). In addition, we evaluated response times (RT) for each answer type.

Given the nested nature of the design, with repeated measures for each subject, we used a linear mixed effects model (LME) to assess effects of age, stimulus domain and the age × stimulus domain interaction, in addition to covariates. The LME included random intercepts for each participant, accounting for variance due to individual performance differences between subjects. The analysis was performed using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). One potential confound to account for was that participants were tested on different object and scene sets, which could vary in difficulty. Even though stimulus sets, randomized sequence and presentation order were all pseudo-randomly assigned to participants, they could by chance have been unevenly distributed across age. We therefore included these design features and all their higher order interactions as further regressors into the model. Furthermore, we added the potential demographic covariates gender and handedness. In addition, we accounted for differences in technical equipment such as browser and operating system used (e.g., Mac or Windows). This same model was tested on all outcome measures mentioned above. All effects were tested via *F*-statistics from a type III Anova using Satterthwaite's method.

Finally, an empirical analysis of effect power was performed for the effect of age and the age × domain interaction. To do so, 1000 random subsamples were drawn for each level of N = 150 up to N = 1550 in incremental steps of 50. To each of these subsamples an LME was fitted. For every level of N, power was then approximated as the proportion of models for which $p < .05$.

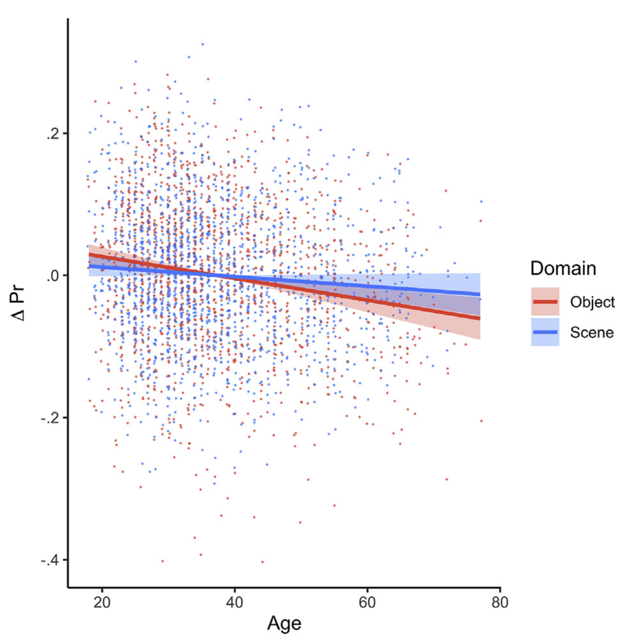

Fig. 2 – Age × Domain interaction trend in person ability *Pr*. While negative age effect is significant for both domains, the effect is more robust for objects (red) than scenes (blue). Plot shows prediction line, partial residuals and confidence band of linear mixed-effects model.

4. Results

4.1. *Mnemonic discrimination performance declines with age*

We observed a significant negative main effect of age on *Pr*, indicating that performance worsened as a function of higher age ($F(1, 1304) = 11.08$, $p = .0009$). *Pr* also differed as a function of domain. That is, on average *Pr* was higher for objects (M = .34, SD = .18) compared to scenes (M = .17, SD = .15), leading to a significant main effect of domain ($F(1552) = 172.15$, $p < .0001$).

4.2. *Stronger age-related decline for objects than scenes*

We found a significant age × domain interaction ($F(1, 1552) = 5.38$, $p = .0205$; Fig. 2), driven by a steeper decrease in object performance compared to scene performance. A subsequent simple slope analysis using the R package reghelper (Hughes, 2020) revealed that the negative effect of age was significant for objects ($t(2141.782) = -4.0451$, $p < .0001$), while merely a negative trend was found for the scenes ($t(2141.782) = -1.6989$, $p = .0895$).

4.3. *Performance decline is driven by change in false alarms but not hits*

Next, we investigated what was driving the age × domain interaction in performance. More precisely, we were interested to know whether the HRs, the FARs or a combination of both caused the decline in accuracy. We therefore analyzed these measures separately. Keeping the same model, we first used hit rates as our outcome measure which did not yield a significant effect of age ($F(1, 1304) = 1.48$, $p = .2237$), but a main effect of domain ($F(1, 1552) = 13.18$, $p = .0003$), driven by higher hit rates in response to object stimuli (M = .83, SD = .12) as compared to scene stimuli (M = .80, SD = .14). Importantly, the

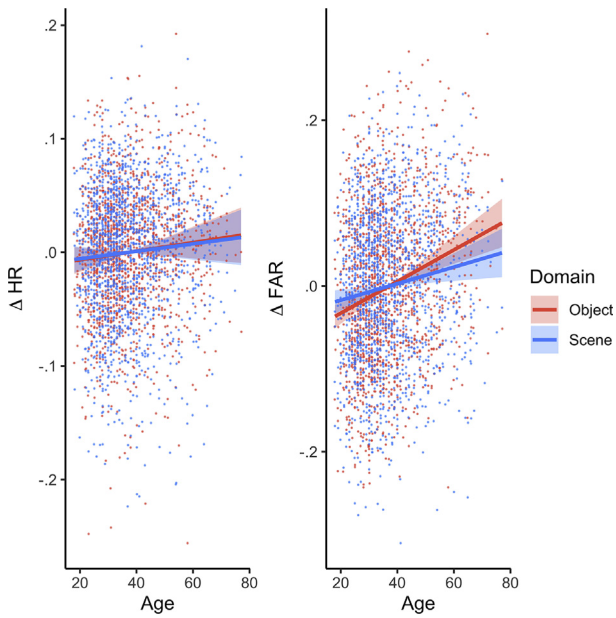

Fig. 3 – Age × Domain interaction in hit rates (HR) and false alarm rates (FAR). False alarms increase with age, but there is no age effect on HR. Positive age effect on FAR is stronger for objects (red) than scenes (blue). Plot shows prediction line, partial residuals and confidence band of linear mixed-effects model.

age × domain interaction was not significant ($F(1, 1552) = .04$, $p = .8371$).

We next looked at FARs. In contrast to the HRs, we found a positive main effect of age ($t(1, 1304) = 19.37$, $p < .0001$), revealing an increase in the FAR with age. Furthermore, as was the case for the hit rates, we found a main effect of domain ($F(1, 1552) = 140.03$, $p = <.0001$). As we expected from previous experience with the task, FARs were generally lower for objects (M = .50, SD = .17) than scenes (M = .63, SD = .16), that is, participants had more problems in rejecting lure scene stimuli, again revealing the overall greater difficulty of scene stimuli.

Interestingly, we found a significant age × domain interaction ($F(1, 1552) = 6.94$, $p = .0085$), driven by a greater increase in FARs across age for object stimuli: the simple slope analysis revealed a steeper and more robust increase in false alarms for object performance ($t(2038.911) = 5.1246$, $p < .0001$) than scene performance ($t(2038.911) = 2.608$, $p = .0091$). To summarize, the age × domain interaction on overall performance was driven by changes in the FARs across age, namely a stronger increase in false alarms for objects (Fig. 3).

Furthermore, we tested whether the greater difficulty of scenes might have led to floor effects causing the interaction. Given that average accuracy for scenes and objects was well above chance level, we believed this to be unlikely, however. Nevertheless, in order to account for possible floor effects, we reran the analyses excluding participants with negative corrected hit rates (chance level). The resulting pattern of effects still remained unchanged. Neither average false alarm nor hit rates were near ceiling, both in general, or for either domain in particular. We therefore believe that the effect is mainly driven by a specific decline in object processing.

4.4. *RT changes with age depend on stimulus domain and answer type*

We observed a linear effect of age on RT ($F(1, 1303.3) = 16.24$, $p = .0001$), driven by an increase in RT (in ms) with age (estimate = 2.94).

We also found a significant domain effect ($F(1, 10725.6) = 20.97$, $p < .0001$), caused by lower RTs to objects (M = 1471.56, SD = 339.16) than scenes (M = 1503.80, SD = 347.24). Furthermore, there was a significant effect of condition ($F(3,10727.1) = 47.33$, $p < .0001$). That is, RTs were lowest for hits (M = 1383.70, SD = 311.91), similar for false alarms (M = 1502.21, SD = 340.97) and correct rejections (M = 1509.94, SD = 309.25), and highest for misses (M = 1557.31, SD = 384.77).

Looking at interactions, we found a significant age × domain interaction ($F(1, 10725.9) = 3.87$, $p = .0492$). The effect was driven by greater increase in RTs to objects (estimate = 3.03) compared to scenes (estimate = 2.41).

We also observed an age × condition effect ($F(3, 10726.1) = 2.97$, $p = .0306$), driven by a stronger increase in RTs for FA (estimate = 3.30) and H (estimate = 2.94), compared to CR (estimate = 2.63) and M (estimate = 2.01) with age. However, when performing pairwise tukey-correct post-hoc contrasts with the emmeans package (Lenth, 2020), only the slope-difference between FA and M remained significant ($t(10725) = 2.866$, $p = .0216$).

4.5. *Covariates of no interest*

Of the remaining factors, only the experimental control factors stimulus set ($F(29, 1304) = 1.87$, $p = .0035$) and stimulus presentation order ($F(1, 1304) = 42.03$, $p = <.0001$) exhibited main effects on performance. Post-hoc analysis of the effect of presentation order revealed a significant presentation order × domain interaction ($F(1, 1551) = 95.65$, $p < .0001$). Pairwise contrasts showed that while presenting the original item version at test resulted in worse performance for objects ($t(2195) = 10.622$, $p = <.0001$), this order did not affect scene performance ($t(2195) = .455$, $p = .6495$). Importantly however, including the interaction in the model did not change the remaining pattern of results. Regarding the higher order effects, we obtained a significant stimulus set × presentation order interaction ($F(29, 1304) = 1.53$, $p = .0358$), meaning that the effect of presentation order depends on the set of item pairs being tested. Also, the stimulus set × randomized sequence interaction was significant ($F(87, 1304) = 1.35$, $p = .0200$), showing that the randomized sequence would affect performance within each set by defining which stimulus was shown as lure or repeat. None of the remaining experimental control effects was significant. As a side note, the above effects were more related to the false alarm rates, where the main effects were significant (stimulus set: $F(29, 1304) = 1.69$, $p = .0125$; presentation order: $F(1, 1304) = 40.33$, $p = <.0001$) and the interactions trended (stimulus set × presentation order: $F(29, 1304) = 1.36$, $p = .0971$; stimulus set × randomized sequence: $F(87, 1304) = 1.27$, $p = .0513$), while no notable effect on hit rates was observed. Moreover, no effect on *Pr* was found for either of the remaining variables: gender, handedness, browser or the operating system used.

Regarding RTs, only the operating system had a significant effect ($F(4, 1302) = 2.53$, $p = .0388$), with Android (M = 1855.43, SD = 206.39) and MacOS (M = 1511.99, SD = 327.54) users giving slower responses than Windows (M = 1484.36, SD = 345.71), Linux (M = 1442.15, SD = 335.13) and ChromeOS (M = 1426.35, SD = 370.87) users. It should be mentioned that the apparent outlier Android only consisted of 16 users.

4.6. *Power analysis*

We report values of N for which power was >.8 (see Supplementary Fig. 2). Regarding the main effect of age, a power of .835 was obtained for N = 1000. As for the age × domain effect, we obtained a power of .875 for N = 1450. Generally, we see that given the current design, data and model structure, a considerable N is necessary to reliably detect the observed effects.

5. Discussion

We tested a large sample of participants (n = 1554) across a very wide adult age-range (18–77 years) using a web-based mnemonic discrimination task (Berron et al., 2018). Using *Pr* as a bias-corrected discrimination ability measure (Berron et al., 2018), we found that mnemonic discrimination performance declines across age, in line with previous research in humans (Holden, Toner, Pirogovsky, Kirwan, & Gilbert, 2013; Stark, Yassa, Lacy, & Stark, 2013; Toner, Pirogovsky, Kirwan, & Gilbert, 2009). Moreover, we found that age-related linear decline of mnemonic discrimination ability was stronger for objects than scenes. Supporting previous research on domain specific decline, we found that worse performance in general (Berron et al., 2018), and a stronger decline in discrimination ability for objects (Reagh et al., 2016, 2018), were driven by an increase in false recognition. The RT results presented a similar pattern, with an age-related linear increase in RTs, which was stronger for objects than scenes.

5.1. *Domain-specific age-related differences across tasks*

To our knowledge, the present study is the first to show such a linear trajectory of decline in mnemonic discrimination along a continuous and wide age-range. This is all the more relevant as age-related changes are assumed to occur in a gradual fashion over time rather than in a stepwise manner (Nyberg, Lövden, Riklund, Lindenberger, & Bäckman, 2012). The pattern of more robust decline for mnemonic discrimination of objects than scenes in aging fits well with previous studies using a variety of methods. Reagh et al. (2018) used a different task, showing an entire block of stimuli before the test phase, extending the time between encoding and retrieval. Accordingly, they employed much more dissimilar stimuli than in our task. That is, their spatial lures consisted of objects changing their location on the screen, whereas in our scenes, the shape or geometry of spatial boundaries would change (see Fig. 1C). Finally, they classified subjects into two distinct age groups. Still, they obtained very similar results, finding no age-related decrease in hit rates and a more pronounced deficit in correct rejections for objects than scenes in the elderly. Moreover, they found no significant difference in false alarms for spatial memory. While we did find significantly increased false alarms and a trend for decreased overall accuracy for scenes with age, the effect is far less robust than for the objects, lending support to the idea that scene performance is less affected in aging. Stark and Stark (2017) reported a similar pattern of results using the MST. However, as noted earlier, the MST and the object-scene task differ in several ways. The duration of presentation and test blocks is considerably larger in the MST, while the object-scene is a 2-back task. Also, the object-scene task only uses similar lures, while the MST additionally employs totally dissimilar foils. Accordingly, participants have two answer options ("old" *vs* "new"), instead of three answer options in the MST ("old", "similar", "new"). Another important difference is that stimuli in the MST vary on more dimensions than ours, and that scenes contain varying objects, making it more difficult to separate domain-specific contributions. While Stark and Stark (2017) found an age-related decrease in the discrimination ability of object and scene lures, the effect seemed to be stronger for objects. However, they did not explicitly test for an age by domain interaction. All in all, the fact that studies using different mnemonic discrimination tasks show a similar pattern of results suggests that age does have a domain-specific impact unrelated to individual task-requirements. Importantly however, we do not claim that spatial memory is unaffected by aging. In addition to the effects on scene memory found here, a host of other studies have reported robust age-related decline of spatial memory (Newman & Kaszniak, 2000; Rosenbaum, Winocur, Binns, & Moscovitch, 2012; Borella, Meneghetti, Ronconi, & Beni, 2014). Generally, the varying findings on spatial memory decline might be due to differences in the tasks employed (Borella, Meneghetti, Ronconi, & Beni, 2014). Compared to the present task for instance, many studies have used more complex spatial stimuli such as virtual environments (Antonova et al., 2009; Etchamendy, Konishi, Pike, Marighetto, & Bohbot, 2012; Head & Isom, 2010; Rodgers, Sindone, & Moffat, 2012), whose processing might be more sensitive to age-related changes. Also, the onset of memory decrease may be domain-dependent and occur later for spatial memory. This might mirror underlying brain changes, where in AD for instance, PRC is affected earlier by atrophy than PHC. Importantly, most studies in humans have focused on domains separately (however see Berron et al., 2018; Reagh et al., 2018), whereas we directly compare them in one model. Taken together, future studies on domain-specific memory decrease need to further investigate its reliance on stimulus properties as well as its temporal progression in AD and healthy aging. In order to improve our understanding of domain-specific temporal progression, the study of longitudinal data will prove invaluable.

5.2. *Presentation order and stimulus similarity*

The fact that task performance was not purely driven by the objective difference between two images is an interesting finding. The stimulus pairs were created in a way that certain changes were made on one stimulus – which we might call the original stimulus – in order to obtain an altered version. It was taken care that changes were realistic and plausible. However, the effect demonstrates that showing the original first and the altered version at test leads to less false alarms than the reversed direction (altered version first and original at test). The effect was also stronger for objects than scenes. While only speculative, we hypothesize that the original version is generally closer to an existing mental prototype, and might therefore facilitate completion to such a prototype (pattern completion) at test, which could have induced the higher false alarm rates found for this direction. This is in line with recent results from Naspi et al. (2020), who found increased false recognitions for stimuli that were better examples of a conceptual prototype. While we did not assess this directly, we would argue that our objects contain more concrete semantic information than the scenes, with specific attributes being expected from them (e.g., a sofa is more likely to be rectangular than round), which may in turn explain the stronger presentation order effect for objects. Interestingly, Pidgeon and Morcom (2014) observed generally higher mnemonic discrimination performance for stimuli that contain familiar semantic information as compared to abstract stimuli, which may explain generally higher performance for objects in our task. In contrast, we found that scene trials were generally more difficult even though the amount of perceptual changes within a pair is larger for scenes, speaking against the importance of simple perceptual similarity. In addition, Pidgeon and Morcom (2014) observed an age-related increase in false alarms that was linked to the semantic rather than perceptual similarity of stimuli. In their view, reliance on conceptual structure may contribute to older adults' bias towards pattern completion. Such an effect might play a role in the stronger age $\times$ domain interaction for objects found here. All in all, while the present results seem related to research regarding stimulus similarity, we are hesitant to make any conclusive statements, since our stimuli were not directly assessed in that regard. We do believe, however, that further investigation into the effect of stimulus features and similarity dimensions will be worthwhile.

5.3. *Potential influence of sample properties*

Importantly, despite using the present task and testing for the interaction specifically, Berron et al. (2018) did not find domain-specific group differences for either hits or false alarms. Still, their functional imaging data showed diminished object-specific activity in PRC for older adults, which in turn was related to decline in object performance exclusively in that age group. Nevertheless, the fact that the same task did not show any behavioral age–domain interaction previously is somewhat surprising, and exemplifies the importance of considering other factors involved, such as the population being tested. For instance, average age in their old group was around 5 years lower (68.6 years) than in Reagh et al. (2018) (73.6 years), who also used a two-groups design. It might be that decline for one domain accelerates around that age (Rönnlund, Nyberg, Bäckman, & Nilsson, 2005; Schaie, 2005), which could explain the differing results. Our sample incorporates quite a large and continuous age-range (18–77), while our model captures the variance in age, potentially making it more sensitive to age-related change and circumventing problems associated with dichotomous age groups. A further difference is the sample size, which was considerably larger in the present study compared to Berron et al. (2018), with 1554 versus 93 participants respectively. Given the rather modest effect size of the age–domain interaction found in our data, the lack of such a finding for Berron et al. (2018) might also simply be due to a lack of statistical power.

5.4. *Challenges and opportunities of internet-based testing*

The previous section touched on a few challenges when inferring age effects from cross-sectional samples. A note of caution regarding our online-sample has to be raised, in this matter. For instance, we did not control for neuropathologies, cognitive impairment, depression, perceptual acuity or technical devices used, all of which might influence memory performance. Indeed, limitations to clinical assessment of the sample are a major challenge for internet-based research. This is all the more relevant, as there seem to be systematic differences between age subgroups of the Mturk worker population. For instance, old Mturkers seem to differ more from the general population than the young, compared to their respective age peers (Huff & Tingley, 2015). Assessing cognitive impairment in online samples does pose several problems, however. First, participants' diagnoses and clinical tests employed might vary widely. Moreover, some participants may not have been diagnosed with cognitive impairment despite being affected, others might not want to share this information. We therefore decided to stick to objective performance measures to filter out problematic data, but do acknowledge the importance of finding ways to assess cognitive impairment in future studies. As for depression, Mturkers do exhibit a higher prevalence of anxiety and depression disorders compared to the general population, which in turn have been linked to deficits in mnemonic discrimination performance (Camfield, Fontana, Wesnes, Mills, & Croft, 2017). It seems unlikely to us, however, that the age effects found in the present study were driven by depression. First, old Mturkers seem to exhibit less clinical symptoms than their young counterparts (Arditte, Çek, Shaw, & Timpano, 2016). Moreover, while younger Mturkers (<50) exhibit a higher prevalence of depression than the general US population, this effect is in fact reversed for older Mturkers (<50) (Walters, Christakis, & Wright, 2018), being less depressed on average. Thus, if depression had an effect in the present study, we would expect it to improve memory with age, contrary to the effect found here. As for visual acuity, it is known to decline with age and might have contributed to the observed age-related decline on performance. In this case, the effect would be more perceptual than mnemonic in nature. We did not directly assess visual acuity in the present study, but administered a simple arrow task, which we would expect to reveal more pronounced visual impairment. It is unclear, however, to what extent the task is able to detect subtle visual impairment, as average performance was quite high (93% correct). However, Berron et al. (2019) did apply a visual screening procedure on elderly participants (mean age = 66 years), using the object-scene task and a subset of the stimuli used here. They found no effect of visual acuity on task performance for either objects or scenes. We therefore do not expect visual acuity to have had a major effect on the present results. Finally, while Mturk allowed us to obtain data on technical equipment such as the browser and operating system used, we did not assess the exact device type (laptop, tablet, etc.) or screen size, which may affect memory performance and also vary with effects of interest such as age. It would therefore be important to collect this information and to explicitly control for those factors in future web-based studies.

While it is important to find ways of better characterizing internet-based participants, Mturk data is not by default less valid. For one thing, a whole variety of classical psychological findings have already been replicated using Mturk (Chandler & Shapiro, 2016). Furthermore, Mturk workers are more diverse than the typically studied college students and more representative of the overall population (Berinsky, Huber, & Lenz, 2012), which may thus enhance the generalizability of findings beyond confined demographics. Finally, the problem of selection bias is not unique to internet-based samples, but its recent surge in popularity seems to highlight its importance. Indeed, systematic differences between age groups are likely to exist in the classical lab-based setting as well, where the young group is typically exclusively made up of college students, and the old group is usually recruited from local elderly people that are either already in close contact with a hospital or research institute, or that show increased interest in clinical assessment. While we underline the importance of characterizing the sample under study in the future, we do believe that the statistical power that comes with large, internet-based data is a promising way to account for a host of further covariates. More generally, the ease with which large sample data can today be collected from mobile devices, in some cases combined with biological markers serving as ground truth, offers a unique opportunity to develop ever more sensitive behavioral markers for memory decline.

5.5. *Limitations*

One important goal of the present study was to assess item difficulty by using a very large number of different items, which led to stimulus difficulties not being balanced across domains. However, we argue that this does not account for the age-interaction found in our study. If the effect was due to overall higher difficulty in one domain, we would rather expect a sharper decline for the scenes in aging, given that they are generally more difficult. Still, future studies investigating domain-specific decline should aim to balance difficulty across domains in order to control for confounds unrelated to domain-specific processing.

One further shortcoming of the present study is that no brain data were available to directly inspect the relation between modality-dependence and region- or network-specific brain changes. Future studies using molecular biomarkers and MRI should investigate whether the pattern of stronger decline for object memory can indeed be directly attributed to brain changes in the anterior–temporal pathway. Furthermore, a meaningful design improvement would be the study of longitudinal cohorts: in the search for behavioral markers of age-related brain changes, it is imperative for these brain-behavior links to be reliably observed on an individual rather than merely on a population level.

6. Conclusion

In conclusion, we found stronger age-related decline in the mnemonic discrimination of objects as compared to scenes. This association was mainly driven by an increase in false alarms with aging, with the effect being more pronounced for objects than scenes. We discuss several pros and cons of memory research using web-based assessment. Importantly, the present study demonstrates the feasibility of testing large online samples on a previously lab-based mnemonic discrimination task. We were able to replicate and expand on previous research of mnemonic discrimination, and furthermore control for a wide variety of covariates. While research on online-samples certainly introduces new challenges, we do believe that the availability of large, lifespan data from crowdsourcing platforms such as Mturk opens up new possibilities of exploring and building more comprehensive models of memory decline in aging. Importantly, we believe that large-scale approaches provide a great means to develop more sensitive behavioral markers for the early diagnosis of AD, given that AD pathology particularly and differentially affects functional memory systems in the medial temporal lobe.

Data availability

No part of the study procedures or analyses was preregistered prior to the research being conducted. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. Our ethical approval does not permit open sharing of participant data without prior informed consent. Anonymized data will be shared by request for the sole purpose of replicating procedures and results presented in the article. The entire analysis code as well as the task presentation code have been openly archived and are available at https://github.com/jguesten/MTURK_CORTEX. The stimulus material used in the present study belongs to neotiv GmbH via a legal agreement with the Otto-von-Guericke University, Magdeburg. Any query for access to the material must be addressed directly to neotiv GmbH.

CRediT author statement

Jeremie Güsten: Conceptualization, Data curation, Software, Formal analysis, Writing – Original draft preparation. Gabriel Ziegler: Formal analysis, Writing- Reviewing and Editing. Emrah Düzel: Funding acquisition, Writing – Reviewing and Editing. David Berron: Conceptualization, Writing- Original draft preparation.

Declaration of competing interest

David Berron and Emrah Düzel are scientific co-founders of neotiv GmbH (Magdeburg, Germany).

Acknowledgments

This work was funded by the German Research Foundation CRC 779 (TP A07). We are grateful to the Psiturk Development Team, who were very helpful in setting up the task. Moreover, we thank Hartmut Schütze and Bernd Markgraf for their technical assistance.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2020.12.017.

REFERENCES

Antonova, E., Parslow, D., Brammer, M., Dawson, G. R., Jackson, S. H. D., & Morris, R. G. (2009). Age-related neural activity during allocentric spatial memory. *Memory, 17*(2), 125–143. https://doi.org/10.1080/09658210802077348

Arditte, K. A., Çek, D., Shaw, A. M., & Timpano, K. R. (2016). The importance of assessing clinical phenomena in Mechanical Turk research. *Psychological Assessment, 28*(6), 684–691. https://doi.org/10.1037/pas0000217

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

Berron, D., et al. (2018). Age-related functional changes in domain-specific medial temporal lobe pathways. *Neurobiology of Aging, 65*, 86–97. https://doi.org/10.1016/j.neurobiolaging.2017.12.030

Berron, D., et al. (2019). Higher CSF tau levels are related to hippocampal hyperactivity and object mnemonic discrimination in older adults. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 39*(44), 8788–8797. https://doi.org/10.1523/jneurosci.1279-19.2019

Berry, C. J., Shanks, D. R., Speekenbrink, M., & Henson, R. N. A. (2012). Models of recognition, repetition priming, and fluency: Exploring a new framework. *Psychological Review, 119*(1), 40–79. https://doi.org/10.1037/a0025464

Borella, E., Meneghetti, C., Ronconi, L., & Beni, R. D. (2014). Spatial abilities across the adult life span. *Developmental Psychology, 50*(2), 384–392. https://doi.org/10.1037/a0033818

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica, 82*(4), 239–259. https://doi.org/10.1007/bf00308809

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological Science: A Journal of the Association for Psychological Science, 6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Camfield, D. A., Fontana, R., Wesnes, K. A., Mills, J., & Croft, R. (2017). Effects of aging and depression on mnemonic discrimination ability. *Aging, Neuropsychology, and Cognition, 25*(3), 464–483. https://doi.org/10.1080/13825585.2017.1325827

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior, 29*(6), 2156–2160. https://doi.org/10.1016/j.chb.2013.05.009

Chandler, J., & Shapiro, D. (2016). Conducting clinical research using crowdsourced convenience samples. *Annual Review of Clinical Psychology, 12*(1), 53–81. https://doi.org/10.1146/annurev-clinpsy-021815-093623

Etchamendy, N., Konishi, K., Pike, G. B., Marighetto, A., & Bohbot, V. D. (2012). Evidence for a virtual human analog of a rodent relational memory task: A study of aging and fMRI in young adults. *Hippocampus, 22*(4), 869–880. https://doi.org/10.1002/hipo.20948

D'Orazio, M. (2019). *univOutl: Detection of univariate outliers*. Retrieved from: https://CRAN.R-project.org/package=univOutl.

Downs, J. S., Holbrook, M. B., Sheng, S., & Cranor, L. F. (2010). *Are your participants gaming the system?: Screening Mechanical Turk workers* (pp. 2399–2402). CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

Hauser, D. J., & Schwarz, N. (2016). Attentive turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods, 48*(1), 400–407. https://doi.org/10.3758/s13428-015-0578-z

Head, D., & Isom, M. (2010). Age effects on wayfinding and route learning skills. *Behavioural Brain Research, 209*(1), 49–58. https://doi.org/10.1016/j.bbr.2010.01.012

Holden, H. M., Toner, C., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2013). Visual object pattern separation varies in older adults. *Learning & Memory (Cold Spring Harbor, N.Y.), 20*(7), 358–362. https://doi.org/10.1101/lm.030171.112

Huff, C., & Tingley, D. (2015). "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics, 2*(3). https://doi.org/10.1177/2053168015604648, 205316801560464.

Hughes, J. (2020). *Reghelper: Helper functions for regression analysis. R package version 1.0.0.* https://CRAN.R-project.org/package=reghelper.

Johnson, S. A., et al. (2017). Rodent age-related impairments in discriminating perceptually similar objects parallel those observed in humans. *Hippocampus, 27*(7), 759–776. https://doi.org/10.1002/hipo.22729

Lenth, R. (2020). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.5.1 https://CRAN.R-project.org/package=emmeans.

Maass, A., et al. (2019). Alzheimer's pathology targets distinct memory networks in the ageing brain. *Brain, 142*(8), 2492–2509. https://doi.org/10.1093/brain/awz154

Naspi, L., Hoffman, P., Devereux, B., Thejll-Madsen, T., Doumas, L. A., & Morcom, A. (2020). *Multiple dimensions of semantic and perceptual similarity contribute to mnemonic discrimination*. https://doi.org/10.31234/osf.io/qt5wc

Newman, M. C., & Kaszniak, A. W. (2000). Spatial memory and aging: Performance on a human analog of the Morris Water Maze. *Aging, Neuropsychology, and Cognition, 7*(2), 86–93. https://doi.org/10.1076/1382-5585(200006)7:2;1-u;ft086

Nyberg, L., Lövden, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences, 16*(5), 292–305. https://doi.org/10.1016/j.tics.2012.04.005

Pidgeon, L. M., & Morcom, A. M. (2014). Age-related increases in false recognition: The role of perceptual and conceptual similarity. *Frontiers in Aging Neuroscience, 6*. https://doi.org/10.3389/fnagi.2014.00283

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews. Neuroscience, 13*(10), 713–726. https://doi.org/10.1038/nrn3338

Reagh, Z. M., et al. (2016). Greater loss of object than spatial mnemonic discrimination in aged adults. *Hippocampus, 26*(4), 417–422. https://doi.org/10.1002/hipo.22562

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). https://doi.org/10.1016/j.neuron.2018.01.039, 1187–1198.e4.

Rosenbaum, R. S., Winocur, G., Binns, M. A., & Moscovitch, M. (2012). Remote spatial memory in aging: All is not lost. *Frontiers in Aging Neuroscience, 4*, 25. https://doi.org/10.3389/fnagi.2012.00025

Ritchey, M., Libby, L. A., & Ranganath, C. (2015). Progress in brain research. *Progress in Brain Research, 219*, 45–64. https://doi.org/10.1016/bs.pbr.2015.04.001

Rodgers, M. K., Sindone, J. A., & Moffat, S. D. (2012). Effects of age on navigation strategy. *Neurobiology of Aging, 33*(1). https://doi.org/10.1016/j.neurobiolaging.2010.07.021, 202.e15–202.e22.

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L.-G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20*(1), 3–18. https://doi.org/10.1037/0882-7974.20.1.3

Schaie, K. W. (2005). *Developmental influences on adult intelligence* (pp. 309–317). https://doi.org/10.1093/acprof:oso/9780195156737.003.0013

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

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, C. E. L., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences, 98*(22), 12760–12766. https://doi.org/10.1073/pnas.221462998

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., 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

Thomas, K. A., & Clifford, S. (2017). Validity and Mechanical Turk: An assessment of exclusion methods and interactive experiments. *Computers in Human Behavior, 77*, 184–197. https://doi.org/10.1016/j.chb.2017.08.038

Toner, C. K., Pirogovsky, E., Kirwan, C. B., & Gilbert, P. E. (2009). Visual object pattern separation deficits in nondemented older adults. *Learning & Memory (Cold Spring Harbor, N.Y.), 16*(5), 338–342. https://doi.org/10.1101/lm.1315109

Vieweg, P., Stangl, M., Howard, L. R., & Wolbers, T. (2015). Changes in pattern completion – a key mechanism to explain age-related recognition memory deficits? *Cortex; A Journal Devoted to the Study of the Nervous System and Behavior, 64*, 343–351. https://doi.org/10.1016/j.cortex.2014.12.007

Walters, K., Christakis, D. A., & Wright, D. R. (2018). Are Mechanical Turk worker samples representative of health status and health behaviors in the U.S.? *PLoS One, 13*(6), Article e0198835. https://doi.org/10.1371/journal.pone.0198835

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