

Neural Pattern Similarity Differentially Relates to Memory Performance in Younger and Older Adults

Verena R. Sommer,¹ Yana Fandakova,¹ Thomas H. Grandy,¹ Yee Lee Shing,^{1,2} Markus Werkle-Bergner,¹ and Myriam C. Sander¹

¹Center for Lifespan Psychology, Max Planck Institute for Human Development, 14195 Berlin, Germany, and ²Institute of Psychology, Goethe University Frankfurt, 60629 Frankfurt am Main, Germany

Age-related memory decline is associated with changes in neural functioning, but little is known about how aging affects the quality of information representation in the brain. Whereas a long-standing hypothesis of the aging literature links cognitive impairments to less distinct neural representations in old age (“neural dedifferentiation”), memory studies have shown that overlapping neural representations of different studied items are beneficial for memory performance. In an electroencephalography (EEG) study, we addressed the question whether distinctiveness or similarity between patterns of neural activity supports memory differentially in younger and older adults. We analyzed between-item neural pattern similarity in 50 younger (19–27 years old) and 63 older (63–75 years old) male and female human adults who repeatedly studied and recalled scene–word associations using a mnemonic imagery strategy. We compared the similarity of spatiotemporal EEG frequency patterns during initial encoding in relation to subsequent recall performance. The within-person association between memory success and pattern similarity differed between age groups: For older adults, better memory performance was linked to higher similarity early in the encoding trials, whereas young adults benefited from lower similarity between earlier and later periods during encoding, which might reflect their better success in forming unique memorable mental images of the joint picture–word pairs. Our results advance the understanding of the representational properties that give rise to subsequent memory, as well as how these properties may change in the course of aging.

Key words: aging; EEG; episodic memory; representational similarity

Significance Statement

Declining memory abilities are one of the most evident limitations for humans when growing older. Despite recent advances of our understanding of how the brain represents and stores information in distributed activation patterns, little is known about how the quality of information representation changes during aging and thus affects memory performance. We investigated how the similarity between neural representations relates to subsequent memory in younger and older adults. We present novel evidence that the interaction of pattern similarity and memory performance differs between age groups: Older adults benefited from higher similarity during early encoding, whereas young adults benefited from lower similarity between early and later encoding. These results provide insights into the nature of memory and age-related memory deficits.

Introduction

A long-standing hypothesis in the cognitive neuroscience of aging holds that neural representations become less specific with

advancing age, with detrimental effects on cognitive performance (Li SC et al., 2001). Previous neuroimaging studies have shown reduced neural distinctiveness between different stimulus items or categories in older compared with younger adults (Park et al., 2004, 2010, 2012; Payer et al., 2006; Goh et al., 2010; Carp et al.,

Received Jan. 23, 2019; revised July 26, 2019; accepted July 29, 2019.

Author contributions: V.R.S., M.W.-B., and M.C.S. analyzed data; V.R.S. wrote the first draft of the paper; V.R.S., Y.F., T.H.G., Y.L.S., M.W.-B., and M.C.S. edited the paper; V.R.S. and M.C.S. wrote the paper; Y.F., T.H.G., Y.L.S., M.W.-B., and M.C.S. designed research; Y.F., Y.L.S., M.W.-B., and M.C.S. performed research.

This work is part of the MERLIN studies conducted within the “Cognitive and Neural Dynamics of Memory across the Lifespan” (CONMEM) project at the Center for Lifespan Psychology, Max Planck Institute for Human Development. The research was partially financed by the Max Planck Society. M.W.-B.’s work was supported by a grant from the German Research Foundation (DFG, WE 4269/3-1), with Y.L.S. as co-PI, as well as an Early Career Research Fellowship 2017–2019 awarded by the Jacobs Foundation. Y.L.S. and M.C.S. were each supported via Minerva Research Groups awarded by the Max Planck Society. Y.L.S. is funded by the European Union (ERC-2018-STGPIVOTAL-758898) and a fellowship from the Jacobs Foundation (JRF 2018–2020). V.R.S. is a fellow of the

International Max Planck Research School on the Life Course. We thank Beate Muehlroth and Xenia Grande for organizing data collection, Kristina Günther for help in participant recruitment, Julia Delius for editorial assistance, Michael Krause for support with cluster computing, all student assistants who helped with data collection, members of the CONMEM project for helpful feedback on the analysis, and all study participants for their time.

The authors declare no competing financial interests.

Correspondence should be addressed to Verena R. Sommer at vsommer@mpib-berlin.mpg.de or Myriam C. Sander at sander@mpib-berlin.mpg.de.

<https://doi.org/10.1523/JNEUROSCI.0197-19.2019>

Copyright © 2019 the authors

2011; St-Laurent et al., 2014; Koen et al., 2019), whereby different definitions and measures of distinctiveness impede comparability between studies (see also “Multivariate EEG analysis” and Discussion). More importantly, most of these studies did not provide evidence for the direct link between reduced neural distinctiveness and behavior, either by not assessing performance or by assessing it separately. An exception is a recent functional magnetic resonance imaging (fMRI) study by Koen et al. (2019) showing an age-invariant association between individual neural category selectivity during encoding (measured as differences between preferred and nonpreferred stimuli) and recognition performance, (for a link between task context reinstatement and performance, see also Abdulrahman et al., 2017). However, memory-related differences in distinctiveness on the item level were not investigated. Such a subsequent memory approach was taken by Zheng et al. (2018), who showed stronger item-specific representations (defined as higher similarity of fMRI patterns across item repetitions than between different items) for later remembered compared with not remembered items, which explained age-related memory performance differences.

Surprisingly, the hypothesis of the cognitive aging literature suggesting that reduced neural specificity underlies cognitive decline is in stark contrast to the prevalent evidence in general memory research that increased neural similarity is actually advantageous for performance: In young adult samples, various studies have shown that the representational similarity between different items is positively related to memory for these items (LaRocque et al., 2013; Davis et al., 2014a; Lu et al., 2015; Wagner et al., 2016), which is consistent with cognitive and computational models (Gillund and Shiffrin, 1984; Clark and Gronlund, 1996). Global similarity may support memory by capturing regularities (LaRocque et al., 2013) and creating familiarity (Davis et al., 2014a).

To date, most studies have used fMRI to assess neural representations, prioritizing the spatial distribution of representational patterns over their temporal dynamics. In contrast, time-sensitive magneto-/electroencephalography (M/EEG) measurements are able to identify the temporal distribution and oscillatory dynamics in which information is encoded in neural patterns as well as the processing stages at which representational similarity supports performance. For example, Lu et al. (2015) showed that at ~420–580 ms after stimulus onset, global spatio-temporal EEG pattern similarity was higher for later remembered than for not remembered symbols. In addition, concurrent power increases and decreases in different frequency bands have consistently been related to memory performance (Hanslmayr and Staudigl, 2014). Beyond the relevance of power in single frequency bands, recent scalp (Michelmann et al., 2016, 2018; Kerrén et al., 2018) and intracranial EEG studies (Zhang et al., 2015; Staresina et al., 2016) have demonstrated the importance of considering the rich information profile carried by a wide range of frequencies for item-specific neural signatures. However, there are no previous reports on the relation of the similarity between these dynamic time–frequency patterns to later

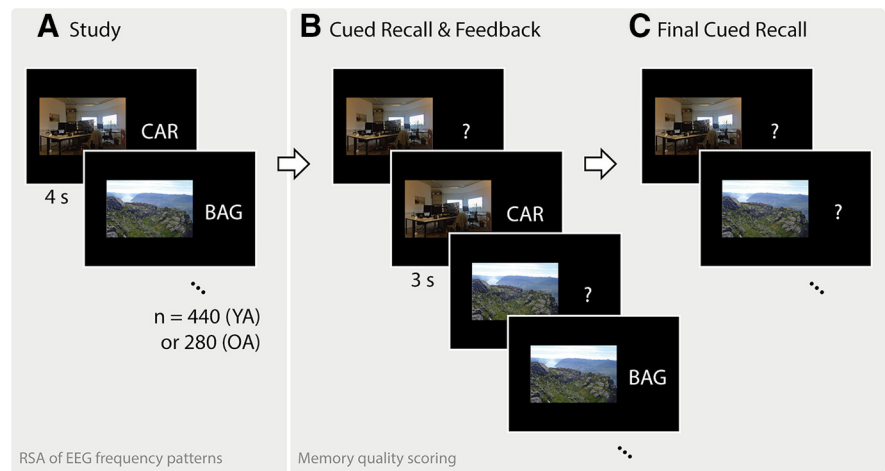


Figure 1. Memory task paradigm (cf. Fandakova et al., 2018). **A**, In the study phase, participants were asked to associate 440 (young adults; YA) or 280 (older adults; OA) scene–word pairs using an imagery strategy. RSA was conducted on EEG data during this phase. **B**, During the cued recall and feedback phase, the scene was presented as a cue to verbally recall the associated word. Subsequently, the original pair was presented again for restudy. The cued recall and feedback phase was performed once for younger and twice for older adults. **C**, During final recall, no feedback was provided. Scene–word pairs were sorted into three memory quality categories based on recall performance in phases B and C (see Fig. 2).

memory success for the studied items either in young or older adults.

To our knowledge, the apparent conflict between the observed beneficial effect of global similarity in memory studies with young adults and the potentially detrimental effect of decreasing distinctiveness in the aging literature has not been explicitly addressed. Here, we aimed to resolve the question whether distinctiveness or similarity (which we define as each other’s inverse) between patterns of neural activity is beneficial for memory performance by systematically investigating the relation between representational similarity and memory performance in young and older adults. For this, we examined the similarity of EEG frequency patterns elicited when encoding scene–word pairs in relation to age and subsequent recall performance.

Materials and Methods

Experimental design

The research presented here comprises data from two associated studies that investigated age-related differences in associative memory encoding, consolidation, and retrieval (Fandakova et al., 2018; Muehlroth et al., 2019; Sander et al., 2019). Despite subsequent procedural differences, an identical picture–word association task paradigm during which EEG was recorded was at the core of both studies (Fig. 1). In this task, participants were asked to memorize scene–word pairs by applying a previously trained mnemonic imagery strategy. Specifically, they were instructed to imagine the scene and word content together in a unique and memorable mental image. Stimuli consisted of color photographs of indoor and outdoor scenes randomly paired with concrete German nouns (4–8 letters). During the initial study phase, scenes and words were presented next to each other on a black background for 4 s. After studying a pair, participants indicated on a four-point scale how well they were able to integrate the presented scene and word. Young and older adults studied 440 and 280 pairs, respectively. During the subsequent cued recall phase, scenes served as cues for participants to verbally recall the associated word. Recall time was not constrained. After each trial, the correct scene–word pair was presented again for 3 s and subjects were instructed to restudy the pair, independent of previous retrieval success. This recall and restudy phase was repeated one more time for the older adults (similarly to Li J et al., 2004; Daselaar et al., 2006; Morcom et al., 2007; Duverne et al., 2008). Finally, both young and older participants underwent a final cued recall round in which no feedback was presented.

After each phase, we asked participants to indicate on a four-point scale how often they used the instructed imagery strategy or other specific memory strategies to memorize a pair. For a detailed description of the study design and stimulus selection, see Fandakova et al. (2018).

Because older adults often remember less and need more repetitions to learn the same information as young adults (Li J et al., 2004), the numbers of to-be-studied pairs as well as recall repetitions were adjusted between age groups to achieve comparable recall success of approximately half of the studied items. It can be assumed that an equivalent relative amount of information remembered by both groups indicates that the task was similarly difficult for them. These kinds of age-adapted procedures help to identify memory-relevant age differences in brain activity without the influence of confounding variables that correlate with age (Rugg and Morcom, 2005) and thus unconfound task and age difference. Here, extensive pilot experiments showed that the reported numbers of pairs for young and older adults as well as one additional recall and feedback phase for older adults produced the desired results. The adequacy of the chosen number of pairs and repetitions for producing the desired performance levels was recently confirmed by a replication in an independent (third) sample of younger and older adults (Fandakova et al., 2019).

Subjects

The original sample of Study 1 (Fandakova et al., 2018) consisted of 30 healthy young adults and 44 healthy older adults. Due to technical failures, one young adult and three older adults did not complete the study. Study 2 (Muehlroth et al., 2019) involved 34 healthy young adults and 41 healthy older adults, with 4 younger and 4 older participants not completing the experiment for technical reasons. Due to missing or noisy EEG data, we additionally excluded 9 younger and 15 older adults, resulting in a total of 50 younger adults and 63 older adults across both studies, who are included in the analyses presented here (young adults: $M(SD)$ age = 24.3(2.5) years, 19–27 years, 27 female, 23 male; old adults: $M(SD)$ age = 70.4(2.6) years, 63–75 years, 33 female, 30 male).

All participants were right-handed native German speakers, reported normal or corrected-to-normal vision, no history of psychiatric or neurological disease, and no use of psychiatric medication. We screened older adults with the Mini-Mental State Examination (MMSE; Folstein et al., 1975) and none had a score below the threshold of 26 points. Both studies were approved by the ethics committee of the Deutsche Gesellschaft für Psychologie and took place at the Max Planck Institute for Human Development in Berlin, Germany. All participants gave written consent to take part in the experiment.

Behavioral analysis

During the cued recall phases, participants had to verbally recall the word associated with the presented image. We report the proportion of correctly recalled words. False responses occurred rarely and were treated as no responses. Following the rationale of a subsequent memory analysis (Paller and Wagner, 2002) we sorted all trials according to whether the associated word was successfully recalled during the experiment or not. Items that were not remembered after repeated encoding were assumed to have only created a weak memory trace, that was not sufficient for successful recall (although maybe strong enough for successful recognition, see Fandakova et al., 2018). Importantly, given the repeated recall phases, we were able to further differentiate successfully recalled items, distinguishing those that were immediately learned from those that were only acquired later in the experiment. We refer to those items as high memory quality and medium memory quality items, respectively (Fig. 2). Because the pattern similarity between items of a given memory quality was computed (see “Multivariate EEG analysis”), a certain number of trials in that quality category was required. Due to close-to-floor performance of older adults in their initial recall phase (16 older adults recalled only one or no item), we only started scoring older adults’ performance in the second recall phase. To keep the scoring of stimulus pairs as evincing high, medium, or low memory quality comparable across age groups, items that were recalled successfully in the final recall cycle were divided into those that were also already recalled in the previous cycle (high quality) and those that were only remembered in the last recall (medium quality) in

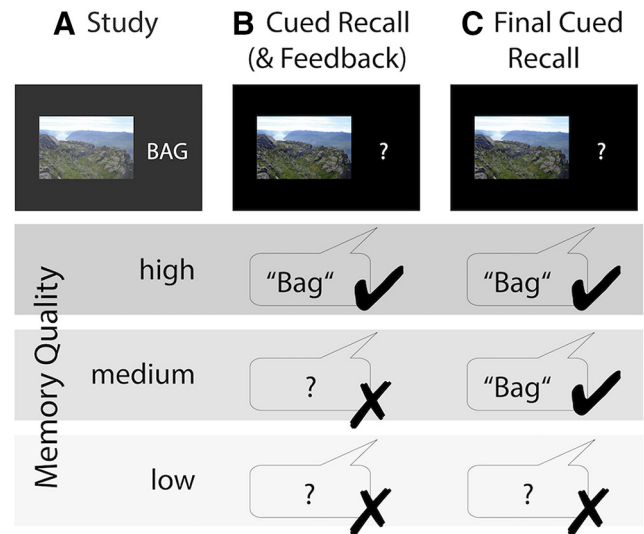


Figure 2. Scoring of stimulus pairs into high, medium, or low memory quality categories based on learning history. For both younger and older adults, items that were correctly recalled in the last recall cycle (C) as well as the previous one (B) were scored as high memory quality items. Pairs that were solely recalled in the final recall were scored as medium memory quality items, and items that were never correctly recalled were scored as low memory quality items. Not depicted: Items that were recalled in the earlier but not later recall were excluded. Older adults performed one more cued recall and restudy cycle (between A and B) that was not included in item scoring due to close-to-floor performance. Note that wrong and missing responses were treated equally.

contrast to never-recalled items (low quality). In other words, memory performance in older adults’ very first recall phase was omitted for memory quality scoring. For both age groups, the few items that were remembered in an earlier but not later recall (i.e., forgotten), were excluded from further EEG analyses (see Results and Fig. 3). All EEG analyses were conducted on the activity patterns elicited during the first learning phase such that all pairs were novel to the participants and no retrieval-related processes could influence the evoked activity patterns.

The fact that the current study design did not allow us to include older adults’ first recall attempt because performance was too low is a limitation as we cannot completely rule out the possibility that the obtained age effects arise from the different memory quality scoring for young and older adults. However, subjecting both age groups to identical procedures in the current study (for example, by also omitting young adults’ first recall) eliminates the strength of our approach, which is the ability to differentiate more fine-grained differences in the memory fate of the stimulus material, which are already observable in the EEG patterns during first encounter. This is the great advantage of our study design compared with the usual contrast of subsequently remembered and not remembered items (see also Discussion).

EEG recording and preprocessing

EEG was recorded continuously with BrainVision amplifiers from 61 Ag/AgCl electrodes embedded in an elastic cap. Three additional electrodes were placed at the outer canthi (horizontal electrooculograph, EOG) and below the left eye (vertical EOG) to monitor eye movements. During recording, all electrodes were referenced to the right mastoid electrode, and the left mastoid electrode was recorded as an additional channel. The EEG was recorded with a pass-band of 0.1 to 250 Hz and digitized with a sampling rate of 1000 Hz. During preparation, electrode impedances were kept below 5 k Ω .

EEG data preprocessing was performed with the Fieldtrip software package (developed at the F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands; <http://fieldtrip.fcdonders.nl>; RRID: SCR_004849) and custom MATLAB code (The MathWorks; RRID: SCR_001622). Data were down-sampled to 250 Hz and an independent component analysis was used to correct for eye blink, (eye) movement, and heartbeat artifacts (Jung et al., 2000). Artifact components were automati-

cally detected, visually checked, and removed from the data. For analyses, the EEG was demeaned, re-referenced to the mathematically linked mastoids, and band-pass filtered (0.2–100 Hz; fourth order Butterworth). Following the FASTER procedure (Nolan et al., 2010), automatic artifact correction was performed for the remaining artifacts. Excluded channels were interpolated with spherical splines (Perrin et al., 1989). Finally, data epochs of 4 s were extracted from -1 to 3 s with respect to the onset of the scene-word presentation during the study phase (Fig. 1A).

EEG analysis

Time–frequency representations (TFRs) of the data were derived using a multitaper approach. For the low frequencies (2–20 Hz), we used Hanning tapers with a fixed width of 500 ms, resulting in frequency steps of 2 Hz. For higher frequencies (25–100 Hz), we used discrete prolate spheroidal sequences (DPSS) tapers with a width of 400 ms in steps of 5 Hz with seven Slepian tapers, resulting in ± 10 Hz smoothing. In this way, we obtained a TFR for each trial and electrode. Trial lengths were reduced to -0.752 to 3 s relative to stimulus onset.

To counter the effect of intrinsically high correlations between frequency patterns due to the $1/\text{frequency}$ power spectrum (Schönauer et al., 2017), we removed the mean background noise spectrum from the log-transformed TFRs following previously established procedures, i.e., as suggested by the better oscillation detection (BOSC) method; (Caplan et al., 2001; Kosciessa et al., 2018; Whitten et al., 2011). Because of structured noise, correlations between different activity patterns are usually very high and almost never at or below zero, meaning that the true null distribution is higher than zero. For detailed discussions of these issues (in fMRI), see Allefeld et al. (2016); Cai et al. (2016).

Multivariate EEG analysis

In the aging literature, different measures of neural distinctiveness (also called specificity, selectivity, differentiation, fidelity) have been used, for instance, the differences in univariate activation levels to preferred and nonpreferred stimuli (Park et al., 2004), increased similarity (St-Laurent et al., 2014) or reduced discriminability (Park et al., 2010) between multivariate neural activity patterns, or the difference between within-category and between-category representational similarity (Carp et al., 2011). Reduced neural distinctiveness in older compared with younger adults has been observed in encoding and retrieval phases between different memory tasks (Carp et al., 2010; St-Laurent et al., 2011), in the reinstatement of encoding task context during retrieval (Abdulrahman et al., 2017; but compare Wang et al., 2016), between different stimulus categories (Park et al., 2004, 2010, 2012; Payer et al., 2006; Carp et al., 2011; Koen et al., 2019), and between different individual stimuli (Goh et al., 2010; St-Laurent et al., 2014). In turn, neural similarity in the general memory literature has been quantified by distance measures based on correlations (e.g., Davis et al., 2014a) or directly as (usually Pearson) correlation (e.g., LaRocque et al., 2013; Lu et al., 2015; Wagner et al., 2016) between activation patterns.

In the present study, EEG data were analyzed using representational similarity analysis (RSA; Kriegeskorte et al., 2008). RSA assesses the resemblance of patterns of neural activity, with similar patterns assumed to represent mutual information and/or processes. Similarity was measured as Pearson correlation, which is insensitive to absolute power and variance of the TFRs. Similarity and distinctiveness were defined as inverses of each other.

Although the pattern of neural activity elicited by a stimulus is commonly defined as the neural representation of that stimulus (Li and Sikström, 2002; Carp et al., 2011), the measured activity pattern does not only contain information of that stimulus but also about the context, the current task etc. Furthermore, activity patterns cannot keep apart the content of a memory (the memory representation in the original sense;

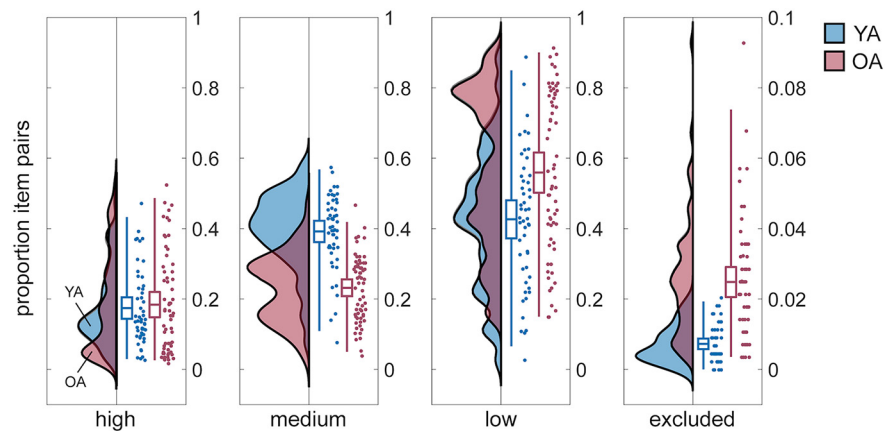


Figure 3. Proportion of item pairs with high, medium, and low memory quality as well as proportion of excluded items for 50 young adults (YA; blue) and 63 older adults (OA; red). Group distributions as un-mirrored violin plots (probability density functions), boxplots with means and 95% confidence intervals, whiskers with second and 98th percentiles, and individual data points (horizontally jittered) (modified from Allen et al., 2019). Note that the y -axis for excluded items differs from that of the other categories. YA studied 440 pairs and OA studied 280 pairs.

Tulving and Bower, 1974) and the underlying processes of, for example, encoding it (if these are distinct entities). However, the term “neural/memory representations” usually denotes the respective activation patterns, and the two terms are therefore used synonymously herein.

In the current study, we investigated frequency-transformed EEG activity patterns (see “EEG analysis”). In addition to their spatial and temporal domains, the (often) strong oscillatory nature of electrochemical brain signals allows information to be encoded in their frequency, power, and phase dimensions, which are largely independent of each other (Cohen, 2011). Oscillations reflect rhythmic and synchronous fluctuations in the excitability of neural populations that have been shown to be functionally relevant for cognition (Buzsáki and Draguhn, 2004; Wang, 2010). Our decision to examine EEG frequency patterns is largely based on findings of recent studies that have demonstrated the importance of the rich information profile carried by a wide range of frequencies for item-specific neural signatures (Zhang et al., 2015; Michelmann et al., 2016; Staresina et al., 2016; Kerrén et al., 2018).

We analyzed between-item representational similarity during the first encoding phase in relation to memory quality. “Item” or “stimulus” always refers to a scene–word pair. Figure 4 illustrates the procedure for analyzing the similarity between stimulus-specific spatiotemporal frequency representations. RSA was conducted for each participant and EEG channel independently. Stimuli were grouped according to high, medium, and low memory quality (Fig. 2). To determine whether between-item representational similarities differed as a function of memory quality, we correlated the noise-corrected and log-transformed frequency patterns of every item with the frequency patterns of all other items of the same memory quality. That is, for each participant we ran three similarity analyses, namely for high, medium, and low memory quality items. To use the same number of items for each RSA of a given participant, we reduced them to the number of items available in the condition with the least items. For example, if there were 50 items with high, 180 items with medium, and 210 items with low memory quality for a given participant, the number of items used in the RSAs of medium and low quality items, respectively, was reduced to 50 as well. Note that the category containing the fewest items was in most cases the group of high memory quality items (except for six younger and six older participants). We randomly sampled the respective number of items from all available trials of the respective memory quality. As the actual measure of similarity, we used pairwise Pearson correlations between the corresponding frequency patterns. In each of these correlations, every pair of frequency vectors (with 26 frequency bins) of all time points from the two respective trials were correlated with each other (470 time points, from 752 ms before stimulus onset to 3000 ms after stimulus onset). The resulting time–time similarity matrices were Fisher’s z transformed. To prevent

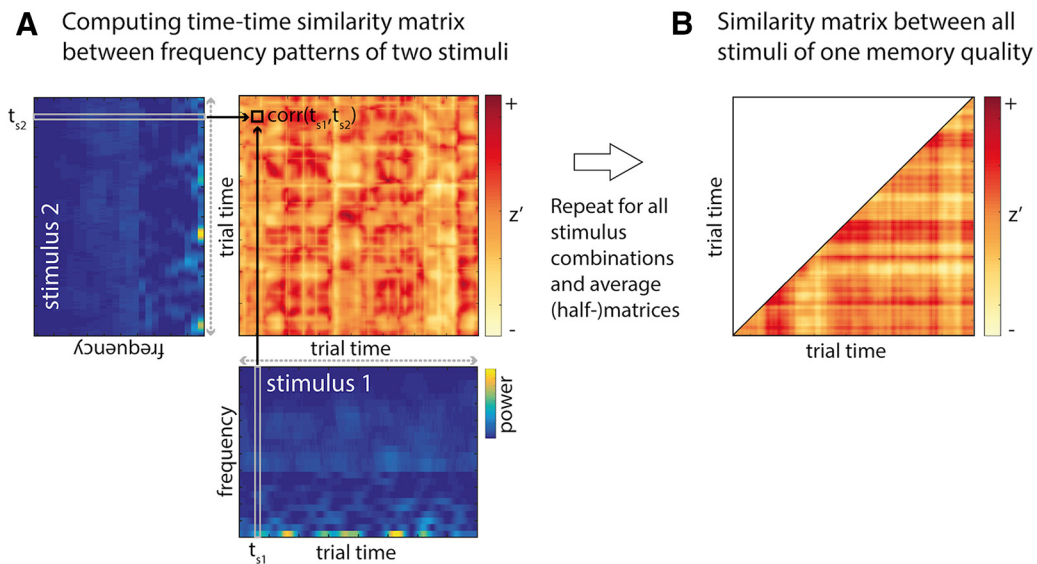


Figure 4. Spectral RSA methodology. **A**, Frequency vector from every time point (i.e., column) of the noise-corrected and log-transformed time–frequency pattern (from one electrode) corresponding to stimulus 1 (bottom) is Pearson-correlated with the vectors from every time point of stimulus 2 (left; rotated by 90°). For illustration, sample vectors of stimulus 1 (t_{s1}) and stimulus 2 (t_{s2}) are highlighted. Correlating these two vectors gives one correlation coefficient, i.e., one coordinate (indicated by black box) on a matrix with time on both axes. Computing all pairwise time vector correlations results in a time–time similarity matrix representing the similarity of those two frequency patterns at all time point combinations. This procedure is repeated for all items of a certain memory quality (i.e., similarity of stimulus 1 with all others, stimulus 2 with all others, etc.). **B**, Averaging across all similarity matrices yields the mean similarity matrix showing the pattern similarity among all items of the same memory quality. Only one triangle and the diagonal of the matrix are relevant because the similarity of every frequency pattern pair is computed twice, resulting in an identical correlation coefficient on both sides of the diagonal. Similarity is quantified as the Fisher’s z -transformed Pearson correlation coefficient (z'). Not depicted: This procedure was repeated for all 60 electrodes, the three memory quality categories, and all subjects.

bias toward the randomly picked items, the item sampling was repeated 20 times. Finally, the matrices were averaged to obtain one between-item similarity matrix for each scene–word pair, which indicates the similarity of this pair to all other pairs of the same memory quality. The similarity matrices of all items within one memory quality were then again averaged to obtain the mean similarity matrices between all high, medium, and low memory quality items, respectively. This procedure was performed separately for each of the 60 scalp electrodes.

The resulting similarity matrices contain the time dimension on both the x - and the y -axis, revealing the frequency pattern similarity not only at identical within-trial time points (diagonal) but also between all combinations of time points (in analogy to the temporal generalization method; Cichy et al., 2014; King and Dehaene, 2014). This enables us to determine whether certain parts of the memory representations were similar to each other at different times during encoding of the respective scene–word pairs.

Because the similarity of any two items is computed twice and thus the identical correlation coefficients appear twice, namely on both sides of the diagonal, the similarity matrix was reduced to only one of the triangles plus the diagonal.

RSAs were computed parallelized on a high-performance computing cluster. All computations and statistics were conducted with MATLAB (The MathWorks, RRID:SCR_001622) versions R2014b or R2016b. The MATLAB-based Fieldtrip Toolbox (Maris and Oostenveld, 2007; Oostenveld et al., 2011; RRID:SCR_004849) was used to perform time–frequency transformations and cluster-based permutation analyses.

Statistical analysis

Memory performance, imagery ratings, and strategy use. We analyzed the relationship between age group and the number of items in the three memory quality categories (high, medium, low) by conducting a χ^2 test. For *post hoc* analyses, we computed two-sided independent samples t tests to test for age differences in the proportion of items within each memory quality category (high, medium, low, as well as forgotten/excluded) and the proportion of items remembered in the final recall task. The imagery ratings after each trial were analyzed by computing frequencies of how often which ratings were given for items of each memory quality. The strength of the relationship between imagery rating and

memory quality on the group and within-person level was tested by conducting nonparametric Goodman and Kruskal’s Gamma correlations for ordinal scaled data. The association between these individual Gamma correlations and the individual effect of pattern similarity and memory quality (regression coefficients; see “Age and memory quality effects in the identified clusters”) was further analyzed using Pearson correlations. To compare younger and older adults’ overall strategy use in the first encoding phase (post-encoding strategy questionnaire), we used the Wilcoxon rank-sum test to examine differences in their median responses of how often they used the imagery strategy.

Differences in representational similarity. Within both groups, we tested for differences in the representational similarity matrices between different memory quality categories (i.e., low < medium < high) by conducting nonparametric, cluster-based, random permutation tests (Fieldtrip Toolbox; Maris and Oostenveld, 2007). Univariate two-sided, dependent samples regression coefficient t statistics were calculated for the time–time similarity matrices at all channels. Clusters were formed by grouping neighboring channel \times time \times time samples with a p -value below 0.05 (spatially and temporally). The respective test statistic was then determined as the sum of all t -values within a cluster. The Monte Carlo method was used to compute the reference distribution for the summed cluster-level t values. Samples were repeatedly (1000 \times) assigned into three groups and the differences between these random groups were contrasted to the differences between the three actual conditions (high, medium, and low memory quality). The t statistic was computed for every repetition and the t -values summed for each cluster. The t values were z transformed for further analysis.

In addition to the linear regression of all three memory qualities mentioned above, we also compared each pair of memory quality categories using a two-sided, dependent samples t test in the permutation analysis (1000 permutations).

We examined overall age differences in the level of between-item pattern similarity independently of memory success by conducting independent samples t statistics within a cluster-based permutation analysis. For simplicity, similarity matrices were averaged across one time dimension (y).

We regarded clusters with test statistic exceeding the 97.5th percentile for their respective reference probability distribution as significant. If

such clusters were obtained, we furthermore assessed the time–time intervals and the topographic distributions of the channels showing when and where, respectively, the differences were reliable. The clusters that were identified for each age group were further examined for age and memory quality effects (see below). In addition, we tested for main age group differences in a separate permutation analysis using independent samples t tests.

To demonstrate that the effects obtained for the young adult group and the older adult group appeared at different times during stimulus encoding, we formally contrasted the times at which the clusters showed significant differences. For this, we extracted the most extreme z -value (z -transformed regression coefficients) within the respective cluster from each subject and compared their coordinates in time–time space. We fitted two models to test whether it was more likely that the time points come from an identical multivariate normal distribution (single model) or from two distinct distributions (two-group model). We then compared the two models using a χ^2 test for model comparison with the null hypothesis that both models fit equally well.

Age and memory quality effects in the identified clusters. To explore potential age differences more closely, we further investigated the relationship between pattern similarity and memory quality by conducting independent samples regression coefficient t statistics for each participant. We extracted and averaged the individual z -transformed regression coefficients within the time–time–electrode clusters that were identified in younger and older adults (see above). For both clusters and age groups, we performed one-sample t tests to determine whether the correlation coefficients come from a distribution with a mean different from zero. Furthermore, we tested for differences between the age groups in both clusters using independent-samples t tests.

Code accessibility

Custom MATLAB code of the main analyses as well as control analyses are available on <https://osf.io/p7v3s/>.

Results

Memory performance and strategy use

During the cued recall phases, participants had to respond verbally with the word they had previously learned to associate with the presented image. We sorted the trials according to whether recall was successful, and when, into high, medium, and low memory quality items (see Materials and Methods). A χ^2 test revealed a significant association between memory quality and age ($\chi^2_{(2)} = 19.71, p < 0.001$). *Post hoc t* tests furthermore showed that the proportion of high memory quality items did not differ between younger adults and older adults ($M_{(\text{younger adults})} = 0.17, SD_{(\text{younger adults})} = 0.11, M_{(\text{older adults})} = 0.18, SD_{(\text{older adults})} = 0.15; t_{(111)} = -0.40, p = 0.690$, two-sample t test; Fig. 3). In contrast, the proportion of items with medium memory quality was significantly larger for younger than older participants ($M_{(\text{younger adults})} = 0.39, SD_{(\text{younger adults})} = 0.11, M_{(\text{older adults})} = 0.23, SD_{(\text{older adults})} = 0.09; t_{(111)} = 8.48, p < 0.001$), whereas older adults had a significantly higher proportion of low memory items ($M_{(\text{younger adults})} = 0.43, SD_{(\text{younger adults})} = 0.19, M_{(\text{older adults})} = 0.56, SD_{(\text{older adults})} = 0.23; t_{(111)} = -3.31, p = 0.001$). Note that in older adults we observed a higher proportion of items that were remembered in an early but not later recall phase, that is, those that were forgotten in the course of the experiment ($M_{(\text{younger adults})} = 0.007, SD_{(\text{younger adults})} = 0.005, M_{(\text{older adults})} = 0.025, SD_{(\text{older adults})} = 0.02; t_{(111)} = -7.04, p < 0.001$). Those item pairs were excluded from further analyses.

Our experimental procedure was successful in inducing variability in memory performance such that both groups could remember approximately half of the studied items: Young adults successfully recalled on average 56.64% (SD = 10.70) and older adults successfully recalled on average

41.60% (SD = 12.06) of the items (440 and 280, respectively). However, our procedure did not completely eliminate age differences since young adults still performed significantly better than older participants in the final recall task ($t_{(111)} = 3.82, p < 0.001$, two-sample t test).

Due to the different number of items that younger and older adults studied in the course of the experiment and the fact that the number of items included in the RSA was reduced based on the smallest memory quality category (usually high quality), the number of items that were compared with each other in the RSA also differed between groups: The median number of items included in the RSA was 48 (range 10–101) for younger adults and 32 (range 5–79) for older adults. The groups differed significantly in their respective item numbers ($z = 3.76, p < 0.001$; Wilcoxon rank-sum test), which, however, did not affect group differences in pattern similarity (control analysis code are available on <https://osf.io/p7v3s/>).

After the first study phase was completed, we asked participants to indicate on a four-point scale how often they had used specific memory strategies for the task (1: almost always, 4: almost never). With regard to the imagery strategy, young adults indicated that they had used it significantly more often than older adults did (younger adults: median = 1.5, min = 1, max = 3; older adults: median = 2, min = 1, max = 4; $z = -5.09, p < 0.001$, Wilcoxon rank-sum test).

We further analyzed the relationship among memory quality, imagery rating, and representational similarity (see below).

Representational similarity

Calculation of between-item representational similarity was based on the initial encoding phase (Fig. 1A). To identify whether high pattern similarity or high pattern distinctiveness during learning was beneficial for later memory success, we sorted all items according to subsequent memory performance and correlated the evoked spatiotemporal frequency pattern of each item with every other item in the same memory quality category. The resulting mean similarity matrices over all channels and scene–word pairs are shown in Figure 5A. These matrices display the similarity of the frequency representations at all possible within-trial time point combinations (–0.752 to 3 s relative to stimulus onset at 0). In contrast, the diagonals of the similarity matrices (also plotted separately in Fig. 5B) only show the similarity between items at identical time points and facilitate a visual comparison of the time courses of representational similarities for the different memory quality categories and age groups. Although this omits much of the similarity information, elevated similarities do occur largely along the diagonal. Note that the diagonals are only plotted for illustration purposes and all statistical tests were performed on the complete matrices as presented in Figure 5A.

Older adults generally exhibit higher representational similarity than young adults

Just before stimulus onset, similarity increased in both age groups and reached a peak around the time of onset (cf. Fig. 5A,B). Elevated similarity occurred mainly between identical trial time points (diagonal) with slightly more persistent activity (elevated off-diagonal similarity) in older adults compared with young adults. Regardless of later memory success, between-item pattern similarity was generally higher in older adults than in young adults (cf. Fig. 5A,B; averaged across the whole time–time matrix and all 60 channels: $M_{(\text{younger adults})} = 0.21, SD_{(\text{younger adults})} =$

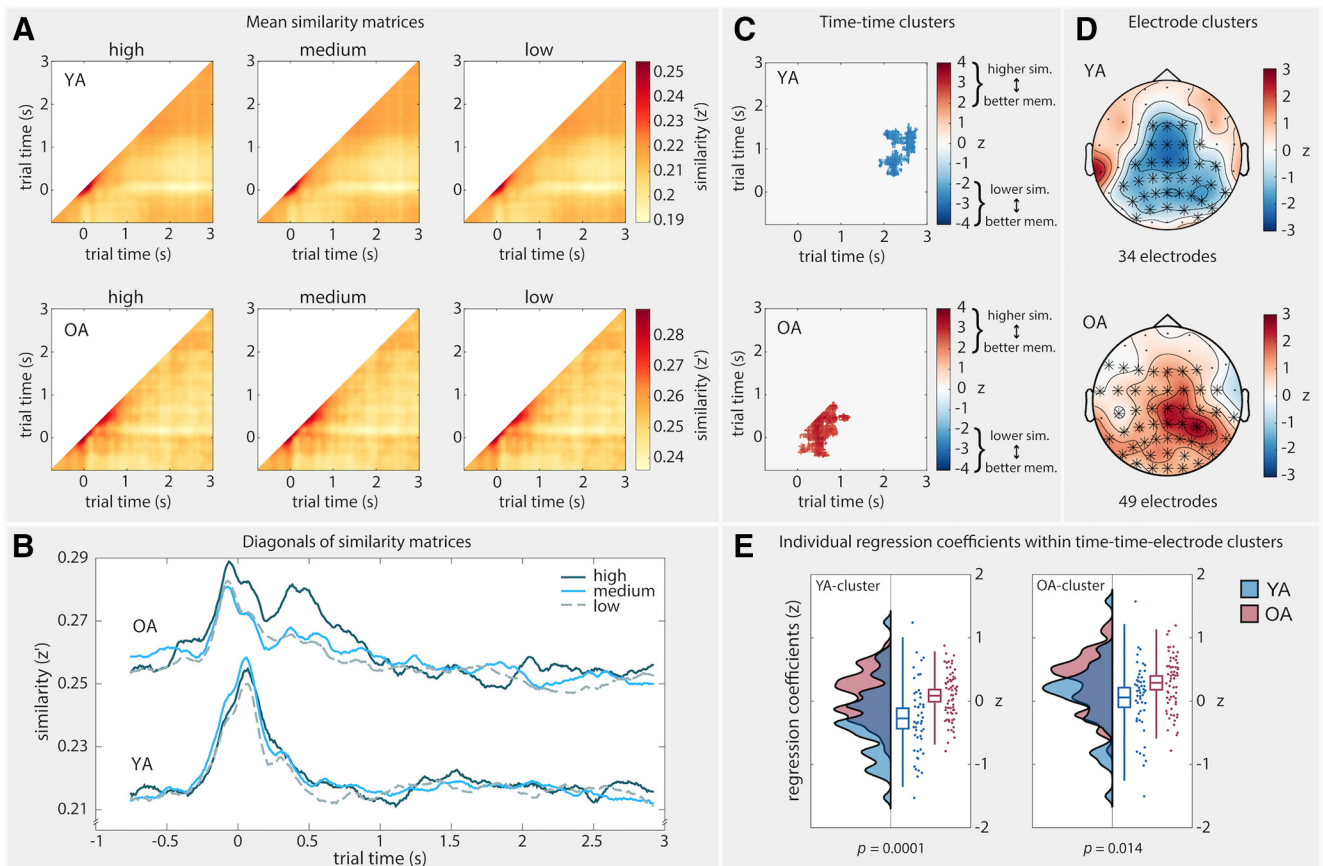


Figure 5. Between-item pattern similarities and statistics. Similarity is quantified as Fisher’s z -transformed Pearson correlation coefficient (z). On time axes, zero denotes stimulus onset. **C** and **D** show results from cluster-based permutation analyses for each age group, **E** shows results from individual regression analyses (see Materials and Methods). **A**, Mean time–time similarity matrices across all 60 channels and items within each memory quality category (high, medium, low) for all 50 young adults (YA; top) and 63 older adults (OA; bottom). Note that the scales differ between age groups. **B**, Diagonals from the time–time similarity matrices (see **A**). **C**, Time–time clusters (masked z -scores) in which the three memory quality categories differ significantly within each age group (averaged across reliable electrodes, see **D**). Positive z -values (red) reveal a positive relationship between pattern similarity and memory quality (higher similarity is associated with higher memory quality), negative z -values (blue) reveal a negative relationship (lower similarity is associated with higher memory quality). **D**, Topographic representations of the electrode clusters that revealed reliable differences between memory quality categories within each age group (averaged across reliable time windows, see **C**). **E**, z -transformed regression coefficients extracted from time–time–electrode clusters identified in YA (left) and OA (right) (see **C** and **D**). Group distributions (probability density functions), boxplots with means and 95% confidence intervals, whiskers with second and 98th percentiles, and individual data points (horizontally jittered) for YA (blue) and OA (red) (modified from Allen et al., 2019). p -values are given for group differences within each cluster (independent-samples t tests). Note the difference between z' (Fisher’s z -transformed correlation coefficients) and z (z -transformed regression coefficients).

0.065, $M_{(\text{older adults})} = 0.25$, $SD_{(\text{older adults})} = 0.068$; 5000 cluster permutations, $p < 0.001$). Furthermore, the level of pattern similarity and final memory performance were negatively correlated across age groups ($r = -0.22$, $p = 0.020$; Pearson correlation). This is in line with previous “dedifferentiation” findings and suggests that also on the within-person level better remembered items should be less similar to each other. However, an across-group correlation may completely differ from a within-group or even within-person correlation (Simpson’s paradox; Kievit et al., 2013). Therefore, we further investigated the association of pattern similarity and memory quality on the within-group and individual level.

Representational similarity differentially relates to memory performance in younger and older adults

Within both age groups, we tested for differences in the levels of representational similarity between scene–word pairs of different memory quality by conducting linear regressions (low < medium < high). We controlled for multiple comparisons by using nonparametric cluster-based permutation tests. In both age groups we identified a cluster with a Monte Carlo p -value below 0.025, which indicates a reliable linear relationship between rep-

resentational similarity and memory quality (young adults: $p = 0.020$; older adults: $p = 0.003$; see Fig. 5C). Importantly, the direction of this relationship differed between groups: while the relation between similarity and memory quality was positive in older adults (low < medium < high), it was negative in young adults (low > medium > high) (Fig. 5E).

The cluster obtained in older adults included most of the diagonal from 50 to 830 ms after stimulus onset and extended off-diagonally to 470 ms before and 1240 ms after stimulus onset (Fig. 5C). Elevated similarity along the diagonal indicates similarity between neural representational patterns at identical trial time points, whereas off-diagonal time windows suggest similar activation patterns at different trial time points. The larger the distance of a coordinate from the diagonal, the more distant are the compared time points in the respective frequency patterns. Differences between memory quality categories were reliable in most (49 of 60) occipital, parietal, temporal, and central electrodes in older adults (Fig. 5D).

In contrast to the cluster found in older adults, an off-diagonal cluster was identified for young adults, in which low memory quality items displayed significantly more similarity than medium and high memory quality items (Fig. 5C). Compared with

older adults, where differences between memory qualities were found to be most pronounced between early and neighboring trial time points, that is, close to the diagonal, the off-diagonal cluster identified in young adults indicated that differences occurred at later and more distant trial time points. Specifically, differences were found between earlier (450–1400 ms after stimulus onset) and later time points (2640–2800 ms after onset) and at 34 mainly parietal-occipital and central electrodes (Fig. 5D). Despite the relatively poor spatial resolution in EEG, the large electrode clusters in both young and older adults indicate that the encoding-related patterns of neural activity that proved to be indicative of subsequent memory were broadly distributed across the brain rather than specific to a particular region.

Additional analyses of pairwise comparison of the three memory quality categories instead of linear regression resulted in a significant cluster that extended over similar time–time intervals and electrodes only for high versus medium memory quality items in younger adults (high vs medium: $p = 0.008$; high vs low: $p = 0.030$; medium vs low: $p = 0.600$; 1000 cluster permutations), and high versus medium as well as high versus low quality in older adults (high vs medium: $p = 0.004$; high vs low: $p = 0.006$; medium vs low: $p = 0.300$; 1000 cluster permutations).

To demonstrate that the effects obtained for the young adult group and the older adult group appeared at different times during stimulus encoding, we extracted the most extreme z -value (z -transformed regression coefficients) within the respective cluster from each subject and compared their coordinates in time–time space. We fitted two models to test whether it was more likely that the time points come from an identical multivariate normal distribution (single model) or from two distinct distributions (two-group model). We then compared the two models using a χ^2 test for model comparison with the null hypothesis that both models fit equally well. The two models differed in model fit ($p < 0.001$), and the two-group model showed significantly better fit. This demonstrates that the effects obtained for young and older adults appeared at different times during stimulus encoding.

Age and memory quality effects in the identified clusters

The cluster-based analyses reported above suggested differential memory-related representational similarity in younger and older adults. To explore potential age differences in more depth, we additionally tested for a linear relationship between representational similarity and memory quality in each participant by conducting individual linear regressions. We then extracted and averaged the individual z -transformed regression coefficients within each time–time–electrode cluster (Fig. 5E). In the young-adult cluster only the mean regression coefficients of the young adults differed from zero (young adults: $t_{(49)} = -3.42$, $p = 0.001$; older adults: $t_{(62)} = 1.79$, $p = 0.078$; one-sample t tests) and vice versa in the older-adult cluster (young adults: $t_{(49)} = 0.75$, $p = 0.457$; older adults: $t_{(62)} = 5.27$, $p < 0.001$). In both clusters the regression coefficients differed significantly between younger and older adults (young-adult cluster: $M_{(\text{young adults})} = -0.27$, $SD_{(\text{young adults})} = 0.57$, $M_{(\text{older adults})} = 0.086$, $SD_{(\text{older adults})} = 0.38$, $t_{(111)} = -4.03$, $p < 0.001$; older-adults cluster: $M_{(\text{young adults})} = 0.058$, $SD_{(\text{young adults})} = 0.55$, $M_{(\text{older adults})} = 0.29$, $SD_{(\text{older adults})} = 0.43$, $t_{(111)} = -2.5$, $p = 0.014$; independent samples t tests), implying that age differences do exist in the relation between representational similarity and memory quality in these clusters.

Stronger links among pattern similarity, memory quality, and imagery ratings in young adults

After each study trial, participants indicated on a four-point scale how well they were able to integrate the presented scene and word (1: not well, 4: very well). We calculated the frequencies of how often each rating was given by each participant. To test the strength of the relationship between participants' imagery ratings and memory quality (see "Memory performance and strategy use") on the group and within-person level, we conducted non-parametric Goodman and Kruskal's Gamma correlations for ordinally scaled data. For both groups, we obtained significant positive relationships showing that higher imagery ratings were given to items of higher memory quality (young adults: $\gamma = 0.28$, $z = 32.04$, $p < 0.001$; older adults: $\gamma = 0.13$, $z = 11.04$, $p < 0.001$). The individual z -values from the within-person correlations of young and older adults differed significantly ($t = 7.08$, $p < 0.001$; two-sample t test) indicating a stronger link between imagery ratings and memory success in young adults.

We further analyzed the association between these individual Gamma correlations and the individual regression coefficients from the RSAs (see "Age and memory quality effects in the identified clusters"). Across both groups (but not within either group), individual z -values from the Gamma correlations and the individual regression coefficients (Fig. 5E) showed a negative association ($r = -0.27$, $p = 0.005$, Pearson correlation). This means that the lower (more negative) the individual regression coefficient (lower similarity in higher memory quality; i.e., the effect seen in young adults), the stronger was the link between imagery rating and memory quality. Equivalently, the higher (more positive) the individual regression coefficient (higher similarity in higher memory quality; i.e., the effect seen in older adults), the weaker was the link between imagery rating and memory quality.

These *post hoc* analyses underline our interpretation of the main results showing an age-dependent effect of between-item representational similarity and memory. We suggest that older adults' benefit from more similar activation patterns may reflect their reliance on gist extraction whereas young adults' benefit from distinct patterns reflects the encoding of more specific details (see Discussion). It seems likely that implementing the imagery strategy allowed the younger participants to create novel, salient mental images from the rather common and similar stimuli, as reflected in more distinct memory representations (McDaniel and Einstein, 1986).

Discussion

The present study aimed to reconcile an evident tension between theories relating neural pattern similarity and memory in the fields of cognitive neuroscience and cognitive aging research. We addressed the central question whether high pattern similarity or high pattern distinctiveness benefits memory performance. To this end, we computed the similarity between the EEG frequency patterns elicited during encoding of different scene–word pairs at each electrode and related this measure of between-pair similarity to the subsequent recall performance of younger and older adults.

For older adults, between-item representational similarity was generally higher compared with young adults, supporting the "dedifferentiation" hypothesis of declining neural distinctiveness with age (Baltes and Lindenberger, 1997; Park et al., 2004, 2012; Li SC et al., 2004; Payer et al., 2006; Carp et al., 2011; St-Laurent et al., 2014). Previous studies suggested that the loss of neural specificity in old age may underlie age-related cognitive impair-

ments. This was, for example, supported by the finding that neural distinctiveness and fluid intelligence were correlated (Park et al., 2010). However, most previous studies were not able to link neural item specificity directly with participants' performance since memory for the items themselves was not assessed. By measuring between-item representational similarity during the encoding phase of an associative memory task and sorting the trials according to subsequent memory performance, we were able to relate measures of neural distinctiveness during encoding directly to later recall success. Notably, results obtained from multivariate analyses like those performed here mostly reflect within-subject variability rather than differences between individuals (Davis et al., 2014b). We found that although older adults remembered significantly fewer items and revealed overall higher between-item pattern similarity than younger adults, on the within-subject level, items represented with high similarity to other items were actually those that older adults remembered best.

Specifically, based on their learning history, we sorted the studied pairs into high, medium, and low memory quality items and, on the within-subject level, measured the linear relationship between the level of representational similarity and memory quality. Importantly, the direction of this relationship and the time window in which representational similarity mattered for subsequent memory performance differed between younger and older participants: For older adults, high similarity early during encoding supported memory performance. For young adults, low similarity between earlier and later time points benefited memory performance.

Beneficial effects of representational similarity for memory have been reported before (LaRocque et al., 2013; Visser et al., 2013; Davis et al., 2014a; Lu et al., 2015; Wagner et al., 2016; Ye et al., 2016) and have been located to medial temporal lobe regions, whereas pattern distinctiveness supported memory in the hippocampus (LaRocque et al., 2013). These pattern separation computations were shown to be impaired for older adults (Wilson et al., 2006; Shing et al., 2011; Yassa et al., 2011). While high distinctiveness may be beneficial for memory performance to prevent false memories, high similarity may support mnemonic decisions by capturing regularities across experiences (LaRocque et al., 2013) and by giving rise to a feeling of familiarity (Davis et al., 2014a). Higher pattern similarity may also reflect more consistent processing that facilitates associative memory formation (Wagner et al., 2016).

Because neural activity patterns contain information on both content and processes (cf. multivariate EEG analysis), the age differences reported here could reflect differences in the similarity of memory representations, processes, or both. The observed benefit of early high pattern similarity in older adults may indeed reflect similarities in processing, for example, increased attention to the stimuli and/or gist extraction. Trials in which similar, memory-beneficial processes occur would be associated with higher between-item pattern similarity and they would have a higher recall probability. However, our findings may also refer to age differences on the representational level: A tendency toward more generalized memories (Koutstaal and Schacter, 1997; Tun et al., 1998; Koutstaal et al., 2001) is often reported for older adults and may also be associated with increased neural similarity. In our study, age differences in the subjective judgements of imagery strategy use during encoding suggest that older adults did indeed rely more on encoding the general gist of scene–word pairs while young adults more often used the imagery strategy to create and encode unique details (cf. Hertzog et al., 2012). Moreover, imagery and memory success were more strongly associated

in young compared with older adults, and more strongly linked to the association of pattern similarity and memory quality. Older adults' benefit from successful early gist extraction may thus be reflected in increased early similarity, whereas young adults' formation of mental images with distinct details may be reflected in increased later dissimilarity.

The negative relationship between pattern similarity and memory performance in younger adults that we report in the current study contrasts with other memory studies that showed a positive relation, namely for recognition memory (LaRocque et al., 2013; Lu et al., 2015; Ye et al., 2016), memory confidence and categorization (Davis et al., 2014a), fear memory (Visser et al., 2013), and associative memory (Wagner et al., 2016). This could be due to the fact that most previous studies showed a beneficial effect of neural similarity for performance in recognition tasks (but see Wagner et al., 2016), in which a sense of familiarity due to high pattern similarity (Gillund and Shiffrin, 1984; Davis et al., 2014a) can be sufficient. In contrast to that, recall tasks as adopted in the current study typically require retrieval of specific details of the studied items (Craik and Tulving, 1975). Therefore, the observed benefit of distinct neural activation patterns for young adults' performance here may be due to the deployed intentional learning task in which participants were explicitly instructed to form very distinct mental images of the corresponding scene–word pairs. Furthermore, similarity of event-related potentials such as that observed by Lu et al. (2015) may result in different effects than in the time–frequency domain.

The current study used an age-adapted procedure with adjusted numbers of items and repetitions to identify memory-relevant age differences in neural patterns. Although procedural differences may have contributed to the observed age differences in pattern similarities, we argue that avoiding differences in task difficulty, a typical confound in age-comparative studies (Rugg and Morcom, 2005) which have shown to be reflected in differences in brain activity (e.g., Nagel et al., 2009) outweigh this concern. In fact, minimizing this confound enables us to conclude that the identified differences between groups are indeed related to age. Nevertheless, it is a limitation that we cannot completely rule out the possibility that the different effects identified in the two groups arise from the different memory quality scoring procedures that were necessary to appropriately handle the age-related performance differences. It is possible that both effects may play an important role for memory encoding in the two age groups but the early similarity seems to be more critical for older adults whereas the later dissimilarity may be more critical for young adults. Alternatively, it is possible that the differences in memory from first to second recall arise from unmeasured differentiation during the second encoding phase.

So far, the prevailing evidence on the relationship between representational similarity and memory has been based on fMRI studies and therefore lacks insights into the temporal dynamics of pattern similarity during the formation of memory representations. Here, we demonstrate the advantage of dissociating different parts within the trial time course that reveal distinctions in the way representational similarity relates to memory performance of younger and older adults. Furthermore, the present study provides further evidence for the high relevance of the rich neural signatures offered by a wide range of frequencies and across multiple topographical sites for memory encoding and extends previous research with similar approaches (Zhang et al., 2015; Michelmann et al., 2016; Staresina et al., 2016; cf. Kerrén et al., 2018).

The question remains how between-item similarity links to within-item similarity, i.e., item-specific representational stability (across item repetitions) and reinstatement (between encoding and retrieval). Recent research suggests that within-item similarity benefits memory performance (Xue et al., 2010; Lu et al., 2015; Xue, 2018) and declines during aging (St-Laurent et al., 2014; Zheng et al., 2018). Understanding the mutual influences of between-item similarity, pattern stability, and pattern reinstatement may be crucial to complete our comprehension of how memories are represented and processed in the brain across the lifespan.

In summary, we provide critical new evidence countering the assumption that a decrease in neural distinctiveness underlies age differences in memory. Although older adults showed generally higher between-item representational similarity and performed worse on the memory task, they actually best remembered the items with the highest peak in pattern similarity early during encoding. Moreover, we show that young adults benefited from eliciting distinct memory representations later during the encoding trial, which presumably reflects the implementation of the imagery strategy for scene–word binding. The work presented here extends our knowledge about between-item pattern similarity as a memory-relevant representational property. In particular, it shows how its relation to cognitive performance may change in the course of aging.

References

- Abdulrahman H, Fletcher PC, Bullmore E, Morcom AM (2017) Dopamine and memory dedifferentiation in aging. *Neuroimage* 153:211–220.
- Allefeld C, Görgen K, Haynes JD (2016) Valid population inference for information-based imaging: from the second-level t test to prevalence inference. *Neuroimage* 141:378–392.
- Allen M, Poggiali D, Whitaker K, Marshall TR, Kievit RA (2019) Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Research* 4:63.
- Baltes PB, Lindenberger U (1997) Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychol Aging* 12:12–21.
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926–1929.
- Cai DJ, Aharoni D, Shuman T, Shobe J, Biane J, Song W, Wei B, Veshkini M, La-Vu M, Lou J, Flores SE, Kim I, Sano Y, Zhou M, Baumgaertel K, Lavi A, Kamata M, et al. (2016) A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* 534:115–118.
- Caplan JB, Madsen JR, Raghavachari S, Kahana MJ (2001) Distinct patterns of brain oscillations underlie two basic parameters of human maze learning. *J Neurophysiol* 86:368–380.
- Carp J, Gmeindl L, Reuter-Lorenz PA (2010) Age differences in the neural representation of working memory revealed by multi-voxel pattern analysis. *Front Hum Neurosci* 4:217.
- Carp J, Park J, Polk TA, Park DC (2011) Age differences in neural distinctiveness revealed by multi-voxel pattern analysis. *Neuroimage* 56:736–743.
- Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time. *Nat Neurosci* 17:455–462.
- Clark SE, Gronlund SD (1996) Global matching models of recognition memory: How the models match the data. *Psychon Bull Rev* 3:37–60.
- Cohen MX (2011) It's about time. *Front Hum Neurosci* 5:2.
- Craik FIM, Tulving E (1975) Depth of processing and the retention of words in episodic memory. *J Exp Psychol Gen* 104:268–294.
- Daselaar SM, Fleck MS, Dobbins IG, Madden DJ, Cabeza R (2006) Effects of healthy aging on hippocampal and rhinal memory functions: an event-related fMRI study. *Cereb Cortex* 16:1771–1782.
- Davis T, Xue G, Love BC, Preston AR, Poldrack RA (2014a) Global neural pattern similarity as a common basis for categorization and recognition memory. *J Neurosci* 34:7472–7484.
- Davis T, LaRocque KF, Mumford JA, Norman KA, Wagner AD, Poldrack RA (2014b) What do differences between multi-voxel and univariate analysis mean? how subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage* 97:271–283.
- Duvernois S, Habibi A, Rugg MD (2008) Regional specificity of age effects on the neural correlates of episodic retrieval. *Neurobiol Aging* 29:1902–1916.
- Fandakova Y, Sander MC, Grandy TH, Cabeza R, Werkle-Bergner M, Shing YL (2018) Age differences in false memory: the importance of retrieval monitoring processes and their modulation by memory quality. *Psychol Aging* 33:119–133.
- Fandakova Y, Werkle-Bergner M, Sander MC (2019) (Only) time can tell: age differences in false memory are magnified at longer delays. *PsyArXiv*. <https://www.doi.org/10.31234/osf.io/eh5x7>.
- Folstein MF, Folstein SE, McHugh PR (1975) “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12:189–198.
- Gillund G, Shiffrin RM (1984) A retrieval model for both recognition and recall. *Psychol Rev* 91:1–67.
- Goh JO, Suzuki A, Park DC (2010) Reduced neural selectivity increases fMRI adaptation with age during face discrimination. *Neuroimage* 51:336–344.
- Hanslmayr S, Staudigl T (2014) How brain oscillations form memories — A processing based perspective on oscillatory subsequent memory effects. *Neuroimage* 85:648–655.
- Hertzog C, Price J, Dunlosky J (2012) Age differences in the effects of experimenter-instructed versus self-generated strategy use. *Exp Aging Res* 38:42–62.
- Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ (2000) Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37:163–178.
- Kerrén C, Linde-Domingo J, Hanslmayr S, Wimber M (2018) An optimal oscillatory phase for pattern reactivation during memory retrieval. *Curr Biol* 28:3383–3392.e6.
- Kievit RA, Frankenhuis WE, Waldorp LJ, Borsboom D (2013) Simpson's paradox in psychological science: a practical guide. *Front Psychol* 4:513.
- King JR, Dehaene S (2014) Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci* 18:203–210.
- Koen JD, Hauck N, Rugg MD (2019) The relationship between age, neural differentiation, and memory performance. *J Neurosci* 39:149–162.
- Kosciessa JQ, Grandy TH, Garrett DD, Werkle-Bergner M (2018) Single-trial characterization of neural rhythms: potentials and challenges. Available from: <https://www.biorxiv.org/content/10.1101/356089v1>.
- Koutstaal W, Schacter DL (1997) Gist-based false recognition of pictures in older and younger adults. *J Mem Lang* 37:555–583.
- Koutstaal W, Schacter DL, Brenner C (2001) Dual task demands and gist-based false recognition of pictures in younger and older adults. *J Mem Lang* 44:399–426.
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4.
- LaRocque KF, Smith ME, Carr VA, Witthoft N, Grill-Spector K, Wagner AD (2013) Global similarity and pattern separation in the human medial temporal lobe predict subsequent memory. *J Neurosci* 33:5466–5474.
- Li J, Morcom AM, Rugg MD (2004) The effects of age on the neural correlates of successful episodic retrieval: an ERP study. *Cogn Affect Behav Neurosci* 4:279–293.
- Li SC, Lindenberger U, Sikström S (2001) Aging cognition: from neuro-modulation to representation. *Trends Cogn Sci* 5:479–486.
- Li SC, Sikström S (2002) Integrative neurocomputational perspectives on cognitive aging, neuromodulation, and representation. *Neurosci Biobehav Rev* 26:795–808.
- Li SC, Lindenberger U, Hommel B, Aschersleben G, Prinz W, Baltes PB (2004) Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychol Sci* 15:155–163.
- Lu Y, Wang C, Chen C, Xue G (2015) Spatiotemporal neural pattern similarity supports episodic memory. *Curr Biol* 25:780–785.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190.
- McDaniel MA, Einstein GO (1986) Bizarre imagery as an effective memory aid: the importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12:54–65.
- Michelmann S, Bowman H, Hanslmayr S (2018) Replay of stimulus-specific

- temporal patterns during associative memory formation. *J Cogn Neurosci* 30:1577–1589.
- Michelmann S, Bowman H, Hanslmayr S (2016) The temporal signature of memories: Identification of a general mechanism for dynamic memory replay in humans. *PLoS Biol* 14:e1002528.
- Morcom AM, Li J, Rugg MD (2007) Age effects on the neural correlates of episodic retrieval: increased cortical recruitment with matched performance. *Cereb Cortex* 17:2491–2506.
- Muehlroth BE, Sander MC, Fandakova Y, Grandy TH, Rasch B, Shing YL, Werkle-Bergner M (2019) Precise slow oscillation-spindle coupling promotes memory consolidation in younger and older adults. *Sci Rep* 9:1940.
- Nagel IE, Preuschhof C, Li SC, Nyberg L, Bäckman L, Lindenberger U, Heekeren HR (2009) Performance level modulates adult age differences in brain activation during spatial working memory. *Proc Natl Acad Sci U S A* 106:22552–22557.
- Nolan H, Whelan R, Reilly RB (2010) FASTER: fully automated statistical thresholding for EEG artifact rejection. *J Neurosci Methods* 192:152–162.
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Paller KA, Wagner AD (2002) Observing the transformation of experience into memory. *Trends Cogn Sci* 6:93–102.
- Park DC, Polk TA, Park R, Minear M, Savage A, Smith MR (2004) Aging reduces neural specialization in ventral visual cortex. *Proc Natl Acad Sci U S A* 101:13091–13095.
- Park J, Carp J, Hebrank A, Park DC, Polk TA (2010) Neural specificity predicts fluid processing ability in older adults. *J Neurosci* 30:9253–9259.
- Park J, Carp J, Kennedy KM, Rodrigue KM, Bischof GN, Huang CM, Rieck JR, Polk TA, Park DC (2012) Neural broadening or neural attenuation? Investigating age-related dedifferentiation in the face network in a large lifespan sample. *J Neurosci* 32:2154–2158.
- Payer D, Marshuetz C, Sutton B, Hebrank A, Welsh RC, Park DC (2006) Decreased neural specialization in old adults on a working memory task. *Neuroreport* 17:487–491.
- Perrin F, Pernier J, Bertrand O, Echallier JF (1989) Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* 72:184–187.
- Rugg MD, Morcom AM (2005) The relationship between brain activity, cognitive performance, and aging: the case of memory. In: *Cognitive neuroscience of aging: linking cognitive and cerebral aging* (Cabeza R, Nyberg L, Park D, eds), pp 132–154. New York: OUP.
- Sander MC, Fandakova Y, Grandy TH, Shing YL, Werkle-Bergner M (2019) Oscillatory mechanisms of successful memory formation in younger and older adults are related to structural integrity. Available from: <https://www.biorxiv.org/content/10.1101/530121v1>.
- Schönauer M, Alizadeh S, Jamalabadi H, Abraham A, Pawlizki A, Gais S (2017) Decoding material-specific memory reprocessing during sleep in humans. *Nat Commun* 8:15404.
- Shing YL, Rodrigue KM, Kennedy KM, Fandakova Y, Bodammer N, Werkle-Bergner M, Lindenberger U, Raz N (2011) Hippocampal subfield volumes: age, vascular risk, and correlation with associative memory. *Front Aging Neurosci* 3:2.
- Staresina BP, Michelmann S, Bonnefond M, Jensen O, Axmacher N, Fell J (2016) Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection. *eLife* 5:e17397.
- St-Laurent M, Abdi H, Burianová H, Grady CL (2011) Influence of aging on the neural correlates of autobiographical, episodic, and semantic memory retrieval. *J Cogn Neurosci* 23:4150–4163.
- St-Laurent M, Abdi H, Bondad A, Buchsbaum BR (2014) Memory reactivation in healthy aging: evidence of stimulus-specific dedifferentiation. *J Neurosci* 34:4175–4186.
- Tulving E, Bower GH (1974) The logic of memory representations. *Psychology of Learning and Motivation* 8:265–301.
- Tun PA, Wingfield A, Rosen MJ, Blanchard L (1998) Response latencies for false memories: gist-based processes in normal aging. *Psychol Aging* 13:230–241.
- Visser RM, Scholte HS, Beemsterboer T, Kindt M (2013) Neural pattern similarity predicts long-term fear memory. *Nat Neurosci* 16:388–390.
- Wagner IC, van Buuren M, Bovy L, Fernández G (2016) Parallel engagement of regions associated with encoding and later retrieval forms durable memories. *J Neurosci* 36:7985–7995.
- Wang TH, Johnson JD, de Chastelaine M, Donley BE, Rugg MD (2016) The effects of age on the neural correlates of recollection success, recollection-related cortical reinstatement, and post-retrieval monitoring. *Cereb Cortex* 26:1698–1714.
- Wang XJ (2010) Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev* 90:1195–1268.
- Whitten TA, Hughes AM, Dickson CT, Caplan JB (2011) A better oscillation detection method robustly extracts EEG rhythms across brain state changes: the human alpha rhythm as a test case. *Neuroimage* 54:860–874.
- Wilson IA, Gallagher M, Eichenbaum H, Tanila H (2006) Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends Neurosci* 29:662–670.
- Xue G (2018) The neural representations underlying human episodic memory. *Trends Cogn Sci* 22:544–561.
- Xue G, Dong Q, Chen C, Lu Z, Mumford JA, Poldrack RA (2010) Greater neural pattern similarity across repetitions is associated with better memory. *Science* 330:97–101.
- Yassa MA, Mattfeld AT, Stark SM, Stark CE (2011) Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proc Natl Acad Sci U S A* 108:8873–8878.
- Ye Z, Zhu B, Zhuang L, Lu Z, Chen C, Xue G (2016) Neural global pattern similarity underlies true and false memories. *J Neurosci* 36:6792–6802.
- Zhang H, Fell J, Staresina BP, Weber B, Elger CE, Axmacher N (2015) Gamma power reductions accompany stimulus-specific representations of dynamic events. *Curr Biol* 25:635–640.
- Zheng L, Gao Z, Xiao X, Ye Z, Chen C, Xue G (2018) Reduced fidelity of neural representation underlies episodic memory decline in normal aging. *Cereb Cortex* 28:2283–2296.