

Research Articles: Behavioral/Cognitive

Neural pattern similarity differentially relates to memory performance in younger and older adults

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

Cite as: J. Neurosci 2019; 10.1523/JNEUROSCI.0197-19.2019

Received: 23 January 2019

Revised: 26 July 2019

Accepted: 29 July 2019

This Early Release article has been peer-reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.jneurosci.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

1 **Neural pattern similarity differentially relates to memory performance**
2 **in younger and older adults**

3
4 Verena R. Sommer¹, Yana Fandakova¹, Thomas H. Grandy¹, Yee Lee Shing^{1,2}, Markus
5 Werkle-Bergner¹, and Myriam C. Sander¹

6
7 *¹Center for Lifespan Psychology, Max Planck Institute for Human Development,*
8 *14195 Berlin, Germany*

9 *²Institute of Psychology, Goethe University Frankfurt, 60629 Frankfurt am Main,*
10 *Germany*

11
12 Corresponding authors: Verena R. Sommer (vsommer@mpib-berlin.mpg.de) or Myriam C.
13 Sander (sander@mpib-berlin.mpg.de); Lentzeallee 94, 14195 Berlin, Germany

14
15
16 Abbreviated title: Age differences in neural similarity and memory

17
18 Conflict of interest: The authors declare no competing financial interests.

19
20
21 Number of pages: 37

22 Number of figures: 5

23 Number of words for Abstract: 235, Introduction: 650, Discussion: 1500

24 Acknowledgements:

25 This work is part of the MERLIN studies conducted within the “Cognitive and Neural Dynamics of
26 Memory across the Lifespan” (CONMEM) project at the Center for Lifespan Psychology, Max Planck
27 Institute for Human Development. The research was partially financed by the Max Planck Society.
28 MWB’s work was supported by a grant from the German Research Foundation (DFG, WE 4269/3-1,
29 YLS as Co-PI) as well as an Early Career Research Fellowship 2017–2019 awarded by the Jacobs
30 Foundation. YLS and MCS were each supported via Minerva Research Groups awarded by the Max
31 Planck Society. YLS is funded by the European Union (ERC-2018-StGPiVOTAL-758898) and a
32 Fellowship from the Jacobs Foundation (JRF 2018-2020). VRS is a fellow of the International Max
33 Planck Research School on the Life Course. We thank Beate Muehlroth and Xenia Grande for
34 organizing data collection, Kristina Günther for help in participant recruitment, Julia Delius for
35 editorial assistance, Michael Krause for support with cluster computing, and all student assistants
36 who helped with data collection. We are grateful to all members of the CONMEM project for helpful
37 feedback on the analysis. Finally, we thank all study participants for their time.

38 **Abstract**

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

57 **Significance statement**

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

67 Introduction

68 A long-standing hypothesis in the cognitive neuroscience of aging holds that neural representations
69 become less specific with advancing age, with detrimental effects on cognitive performance (S.-C. Li
70 et al., 2001). Previous neuroimaging studies have shown reduced neural distinctiveness between
71 different stimulus items or categories in older compared to younger adults (Carp et al., 2011; Goh et
72 al., 2010; Koen et al., 2019; D. Park et al., 2004; J. Park et al., 2010, 2012; Payer et al., 2006; St-
73 Laurent et al., 2014), whereby different definitions and measures of distinctiveness impede
74 comparability between studies (see also Multivariate EEG analysis and Discussion). More
75 importantly, most of these studies did not provide evidence for the direct link between reduced
76 neural distinctiveness and behavior, either by not assessing performance or by assessing it
77 separately. An exception is a recent functional magnetic resonance imaging (fMRI) study by Koen et
78 al. (2019) that showed an age-invariant association between individual neural category selectivity
79 during encoding (measured as differences between preferred and non-preferred stimuli) and
80 recognition performance (see also Abdulrahman et al., 2017, for a link between task context
81 reinstatement and performance). However, memory-related differences in distinctiveness on the
82 item-level were not investigated. Such a subsequent memory approach was taken by Zheng et al.
83 (2017) who showed stronger item-specific representations (defined as higher similarity of fMRI
84 patterns across item repetitions than between different items) for later remembered compared to
85 not remembered items which explained age-related memory performance differences.

86 Surprisingly, the hypothesis of the cognitive aging literature suggesting that reduced neural
87 specificity underlies cognitive decline is in stark contrast to the prevalent evidence in general
88 memory research that increased neural similarity is actually advantageous for performance: In
89 young adult samples, various studies have shown that the representational similarity between
90 different items is positively related to memory for these items (Davis et al., 2014a; LaRocque et al.,
91 2013; Lu et al., 2015; Wagner et al., 2016), which is in line with cognitive and computational models
92 (Clark and Gronlund, 1996; Gillund and Shiffrin, 1984). Global similarity may support memory by
93 capturing regularities (LaRocque et al., 2013) and creating familiarity (Davis et al., 2014a).

94 To date, most studies have used fMRI to assess neural representations, prioritizing the spatial
95 distribution of representational patterns over their temporal dynamics. In contrast, time-sensitive
96 magneto-/electroencephalography (M/EEG) measurements are able to identify the temporal
97 distribution and oscillatory dynamics in which information is encoded in neural patterns as well as
98 the processing stages at which representational similarity supports performance. For example, Lu et
99 al. (2015) showed that at approximately 420–580 ms after stimulus onset, global spatiotemporal
100 EEG pattern similarity was higher for later remembered than for not remembered symbols. In
101 addition, concurrent power increases and decreases in different frequency bands have consistently
102 been related to memory performance (Hanslmayr and Staudigl, 2014). Beyond the relevance of
103 power in single frequency bands, recent scalp (Kerrén et al., 2018; Michelmann et al., 2016, 2018)
104 and intracranial EEG studies (Staresina et al., 2016; Zhang et al., 2015) have demonstrated the
105 importance of considering the rich information profile carried by a wide range of frequencies for
106 item-specific neural signatures. However, there are no previous reports on the relation of the
107 similarity between these dynamic time–frequency patterns to later memory success for the studied
108 items, neither in young nor in older adults.

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

117

118 **Materials and Methods**

119 **Experimental design**

120 The research presented here comprises data from two associated studies that investigated age-
121 related differences in associative memory encoding, consolidation, and retrieval (Fandakova et al.,
122 2018; Muehlroth et al., 2019; Sander et al., 2019). Despite subsequent procedural differences, an
123 identical picture–word association task paradigm during which EEG was recorded was at the core of
124 both studies. In this task, participants were asked to memorize scene–word pairs by applying a
125 previously trained mnemonic imagery strategy. Specifically, they were instructed to imagine the
126 scene and word content together in a unique and memorable mental image. Stimuli consisted of
127 color photographs of indoor and outdoor scenes randomly paired with concrete German nouns (4–8
128 letters). During the initial study phase, scenes and words were presented next to each other on a
129 black background for 4 s. After studying a pair, participants indicated on a four-point scale how well
130 they were able to integrate the presented scene and word. Young and older adults studied 440 and
131 280 pairs, respectively. During the subsequent cued recall phase, scenes served as cues for
132 participants to verbally recall the associated word. Recall time was not constrained. After each trial,
133 the correct scene–word pair was presented again for 3 s and subjects were instructed to restudy the
134 pair, independent of previous retrieval success. This recall and restudy phase was repeated one
135 more time for the older adults (similarly to Daselaar et al., 2006; Duverne et al., 2008; J. Li et al.,
136 2004; Morcom et al., 2007). Finally, both young and older participants underwent a final cued recall
137 round in which no feedback was presented.

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

141 Since older adults often remember less and need more repetitions to learn the same information
142 as young adults (e.g., J. Li et al., 2004), the numbers of to-be-studied pairs as well as recall
143 repetitions were adjusted between age groups in order to achieve comparable recall success of
144 about half of the studied items. It can be assumed that an equivalent relative amount of information

145 remembered by both groups indicates that the task was similarly difficult for them. These kinds of
146 age-adapted procedures help to identify memory-relevant age differences in brain activity without
147 the influence of confounding variables that correlate with age (Rugg and Morcom, 2005), and thus
148 unconfound task and age difference. Here, extensive pilot experiments showed that the reported
149 numbers of pairs for young and older adults as well as one additional recall and feedback phase for
150 older adults produced the desired results. The adequacy of the chosen number of pairs and
151 repetitions for producing the desired performance levels was recently confirmed by a replication in
152 an independent (third) sample of younger and older adults (Fandakova et al., 2019).

153

154 *** **Figure 1** ***

155

156 **Subjects**

157 The original sample of study 1 (Fandakova et al., 2018) consisted of 30 healthy young adults and 44
158 healthy older adults. Due to technical failures, one young adult and three older adults did not
159 complete the study. Study 2 (Muehlroth et al., 2019) involved 34 healthy young adults and 41
160 healthy older adults, with 4 younger and 4 older participants not completing the experiment for
161 technical reasons. Due to missing or noisy EEG data, we additionally excluded 9 younger and 15
162 older adults, resulting in a total of 50 younger adults and 63 older adults across both studies, who
163 are included in the analyses presented here (young adults: $M(SD)$ age = 24.3(2.5) years, 19–27 years,
164 27 female, 23 male; old adults: $M(SD)$ age = 70.4(2.6) years, 63–75 years, 33 female, 30 male).

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

172

173 **Behavioral analysis**

174 During the cued recall phases, participants had to verbally recall the word associated with the
175 presented image. We report the proportion of correctly recalled words. False responses occurred
176 rarely and were treated as no responses. Following the rationale of a subsequent memory analysis
177 (Paller and Wagner, 2002) we sorted all trials according to whether the associated word was
178 successfully recalled during the experiment or not. Items that were not remembered after repeated
179 encoding were assumed to have only created a weak memory trace, that was not sufficient for
180 successful recall (although maybe strong enough for successful recognition, see Fandakova et al.,
181 2018). Importantly, given the repeated recall phases, we were able to further differentiate
182 successfully recalled items, distinguishing those that were immediately learned from those that
183 were only acquired later in the experiment. We refer to those items as high memory quality and
184 medium memory quality items, respectively (see Figure 2). Because the pattern similarity between
185 items of a given memory quality was computed (see Multivariate EEG analysis), a certain number of
186 trials in that quality category was required. Due to close-to-floor performance of older adults in
187 their initial recall phase (16 older adults recalled only one or no item), we only started scoring older
188 adults' performance in the second recall phase. To keep the scoring of stimulus pairs as evincing
189 high, medium, or low memory quality comparable across age groups, items that were recalled
190 successfully in the final recall cycle were divided into those that were also already recalled in the
191 previous cycle (high quality) and those that were only remembered in the last recall (medium
192 quality) in contrast to never-recalled items (low quality). In other words, memory performance in
193 older adults' very first recall phase was omitted for memory quality scoring. For both age groups,
194 the few items that were remembered in an earlier but not later recall (i.e., forgotten), were excluded
195 from further EEG analyses (see Results and Figure 4). All EEG analyses were conducted on the
196 activity patterns elicited during the first learning phase such that all pairs were novel to the
197 participants and no retrieval-related processes could influence the evoked activity patterns.

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

207

208 *** **Figure 2** ***

209

210 **EEG recording and preprocessing**

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

218 EEG data preprocessing was performed with the Fieldtrip software package (developed at the F. C.
219 Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands;
220 <http://fieldtrip.fcdonders.nl>; RRID: SCR 004849) and custom MATLAB code (The MathWorks Inc.,
221 Natick, MA, USA; RRID: SCR 001622). Data were downsampled to 250 Hz and an independent
222 component analysis was used to correct for eye blink, (eye) movement, and heartbeat artifacts (Jung
223 et al., 2000). Artifact components were automatically detected, visually checked, and removed from
224 the data. For analyses, the EEG was demeaned, re-referenced to the mathematically linked mastoids,

225 and band-pass filtered (0.2–100 Hz; fourth order Butterworth). Following the FASTER procedure
226 (Nolan et al., 2010), automatic artifact correction was performed for the remaining artifacts. Excluded
227 channels were interpolated with spherical splines (Perrin et al., 1989). Finally, data epochs of 4
228 seconds were extracted from -1 s to 3 s with respect to the onset of the scene–word presentation
229 during the study phase (Figure 1A).

230

231 **EEG analysis**

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

238 To counter the effect of intrinsically high correlations between frequency patterns due to the
239 1/frequency power spectrum (Schönauer et al., 2017), we removed the mean background noise
240 spectrum from the log-transformed TFRs following previously established procedures (i.e., as
241 suggested by the “Better oscillation detection” (BOSC) method; Caplan et al., 2001; Kosciessa et al.,
242 2018; Whitten et al., 2011). Because of structured noise, correlations between different activity
243 patterns are usually very high and almost never at or below zero, meaning that the true null-
244 distribution is higher than zero. For detailed discussions of these issues (in fMRI), see Allefeld et al.
245 (2016); Cai et al. (2016).

246

247 **Multivariate EEG analysis**

248 In the aging literature, different measures of neural distinctiveness (also called specificity, selectivity,
249 differentiation, fidelity) have been used, for instance, the differences in univariate activation levels to
250 preferred and non-preferred stimuli (e.g., D. Park et al., 2004), increased similarity (e.g., St-Laurent et
251 al., 2014) or reduced discriminability (e.g., J. Park et al., 2010) between multivariate neural activity

252 patterns, or the difference between within-category and between-category representational similarity
253 (e.g., Carp et al., 2011). Reduced neural distinctiveness in older compared to younger adults has been
254 observed in encoding and retrieval phases between different memory tasks (Carp et al., 2010; St-
255 Laurent et al., 2011), in the reinstatement of encoding task context during retrieval (Abdulrahman et al.,
256 2017; but compare Wang et al., 2015), between different stimulus categories (Carp et al., 2011; Koen et
257 al., 2019; D. Park et al., 2004, J. Park et al., 2010, 2012; Payer et al., 2006), and between different
258 individual stimuli (Goh et al., 2010; St-Laurent et al., 2014). In turn, neural similarity in the general
259 memory literature has been quantified by distance measures based on correlations (e.g., Davis et al.,
260 2014a) or directly as (usually Pearson) correlation (e.g., LaRocque et al., 2013; Lu et al., 2015; Wagner
261 et al., 2016) between activation patterns.

262 In the current paper, EEG data were analyzed using representational similarity analysis (RSA;
263 Kriegeskorte et al., 2008). RSA assesses the resemblance of patterns of neural activity, with similar
264 patterns assumed to represent mutual information and/or processes. Similarity was measured as
265 Pearson correlation, which is insensitive to absolute power and variance of the time–frequency
266 representations. Similarity and distinctiveness were defined as inverses of each other.

267 Although the pattern of neural activity elicited by a stimulus is commonly defined as the neural
268 representation of that stimulus (Carp et al., 2011; S.-C. Li and Sikström, 2002), the measured activity
269 pattern does not only contain information of that stimulus but also about the context, the current task
270 etc. Furthermore, activity patterns cannot keep apart the content of a memory (the memory
271 representation in the original sense; e.g., Tulving and Bower, 1974) and the underlying processes of, for
272 example, encoding it (if these are distinct entities). However, the term “neural/memory
273 representations” usually denotes the respective activation patterns, and the two terms are therefore
274 used synonymously in this paper.

275 In the current study, we investigated frequency-transformed EEG activity patterns (see EEG
276 analysis). In addition to their spatial and temporal domains, the (often) strong oscillatory nature of
277 electrochemical brain signals allows information to be encoded in their frequency, power, and phase
278 dimensions, which are largely independent of each other (Cohen, 2011). Oscillations reflect

279 rhythmic and synchronous fluctuations in the excitability of neural populations that have been
280 shown to be functionally relevant for cognition (Buzsáki and Draguhn, 2004; Wang, 2010). Our
281 decision to examine EEG frequency patterns is largely based on findings of recent studies that have
282 demonstrated the importance of the rich information profile carried by a wide range of frequencies
283 for item-specific neural signatures (Kerrén et al., 2018; Michelmann et al., 2016; Staresina et al.,
284 2016; Zhang et al., 2015).

285 We analyzed between-item representational similarity during the first encoding phase in relation
286 to memory quality. "Item" or "stimulus" always refers to a scene-word pair. Figure 3 illustrates the
287 procedure for analyzing the similarity between stimulus-specific spatiotemporal frequency
288 representations. RSA was conducted for each participant and EEG channel independently. Stimuli
289 were grouped according to high, medium, and low memory quality (see Figure 2). In order to
290 examine whether between-item representational similarities differed as a function of memory
291 quality, we correlated the noise-corrected and log-transformed frequency patterns of every item
292 with the frequency patterns of all other items of the same memory quality. That is, for each
293 participant we ran three similarity analyses, namely for high, medium, and low memory quality
294 items. In order to use the same number of items for each RSA of a given participant, we reduced
295 them to the number of items available in the condition with the least items. For example, if there
296 were 50 items with high, 180 items with medium, and 210 items with low memory quality for a
297 given participant, the number of items used in the RSAs of medium and low quality items,
298 respectively, was reduced to 50 as well. Note that the category containing the fewest items was in
299 most cases the group of high memory quality items (except for 6 younger and 6 older participants).
300 We randomly sampled the respective number of items from all available trials of the respective
301 memory quality. As the actual measure of similarity, we employed pairwise Pearson correlations
302 between the corresponding frequency patterns. In each of these correlations, every pair of
303 frequency vectors (with 26 frequency bins) of all time points from the two respective trials were
304 correlated with each other (470 time points, from 752 ms before stimulus onset to 3000 ms after
305 stimulus onset). The resulting time-time similarity matrices were Fisher z-transformed. In order to

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

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

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

321 Representational similarity analyses were computed parallelized on a high-performance computing
322 cluster. All computations and statistics were conducted with Matlab (The MathWorks, Inc., RRID: SCR
323 001622) versions R2014b or R2016b. The Matlab-based Fieldtrip Toolbox (Maris and Oostenveld,
324 2007; Oostenveld et al., 2011; RRID: SCR _004849) was used to perform time–frequency
325 transformations and cluster-based permutation analyses.

326

327 ***** Figure 3 *****

328

329 **Statistical analysis**

330 *Memory performance, imagery ratings, and strategy use*

331 We analyzed the relationship between age group and the number of items in the three memory
332 quality categories (high, medium, low) by conducting a Chi-Squared test. For post-hoc analyses, we

333 computed two-sided independent samples t -tests in order to test for age differences in the
334 proportion of items within each memory quality category (high, medium, low, as well as
335 forgotten/excluded) and the proportion of items remembered in the final recall task. The imagery
336 ratings after each trial were analyzed by computing frequencies of how often which ratings were
337 given for items of each memory quality. The strength of the relationship between imagery rating
338 and memory quality on the group and within-person level was tested by conducting non-parametric
339 Goodman and Kruskal's Gamma correlations for ordinally scaled data. The association between
340 these individual Gamma correlations and the individual effect of pattern similarity and memory
341 quality (regression coefficients; see *Age and memory quality effects in the identified clusters*) was
342 further analyzed using Pearson correlations. To compare younger and older adults' overall strategy
343 use in the first encoding phase (post-encoding strategy questionnaire), we used the Wilcoxon rank
344 sum test to examine differences in their median responses of how often they used the imagery
345 strategy.

346

347 *Differences in representational similarity*

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

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

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

367 We regarded clusters whose test statistic exceeded the 97.5th percentile for its respective
368 reference probability distribution as significant. If such clusters were obtained, we furthermore
369 assessed the time–time intervals and the topographic distributions of the channels showing when
370 and where, respectively, the differences were reliable. The clusters that were identified for each age
371 group were further examined for age and memory quality effects (see below). In addition, we tested
372 for main age group differences in a separate permutation analysis using independent samples t -
373 tests.

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

382

383 *Age and memory quality effects in the identified clusters*

384 To explore potential age differences more closely, we further investigated the relationship between
385 pattern similarity and memory quality by conducting independent samples regression coefficient t -
386 statistics for each participant. We extracted and averaged the individual z -transformed regression

387 coefficients within the time–time–electrode clusters that were identified in younger and older adults
388 (see above). For both clusters and age groups we performed one-sample *t*-tests to test whether the
389 correlation coefficients come from a distribution with a mean different from zero. Furthermore, we
390 tested for differences between the age groups in both clusters using independent samples *t*-tests.

391

392 **Code Accessibility**

393 Custom Matlab code of the main analyses as well as control analyses are available on a public Gitlab
394 repository.

395

396 **Results**

397 **Memory performance and strategy use**

398 During the cued recall phases, participants had to respond verbally with the word they had
399 previously learned to associate with the presented image. We sorted the trials according to whether
400 recall was successful, and when, into high, medium, and low memory quality items (see Methods). A
401 Chi-Squared test revealed a significant association between memory quality and age ($\chi^2(2) = 19.71$,
402 $p = 0.000053$). Post-hoc *t*-tests furthermore showed that the proportion of high memory quality
403 items did not differ between younger adults and older adults ($M(\text{younger adults}) = 0.17$, $SD(\text{younger}$
404 $\text{adults}) = 0.11$, $M(\text{older adults}) = 0.18$, $SD(\text{older adults}) = 0.15$; $t(111) = -0.4$, $p = 0.69$, two-sample *t*-
405 test; see Figure 4). In contrast, the proportion of items with medium memory quality was
406 significantly larger for younger than older participants ($M(\text{younger adults}) = 0.39$, $SD(\text{younger}$
407 $\text{adults}) = 0.11$, $M(\text{older adults}) = 0.23$, $SD(\text{older adults}) = 0.09$; $t(111) = 8.48$, $p = 10^{-13}$), while older
408 adults had a significantly higher proportion of low memory items ($M(\text{younger adults}) = 0.43$,
409 $SD(\text{younger adults}) = 0.19$, $M(\text{older adults}) = 0.56$, $SD(\text{older adults}) = 0.23$; $t(111) = -3.31$, $p =$
410 0.0012). Note that in older adults we observed a higher proportion of items that were remembered
411 in an early but not later recall phase, i.e., that were forgotten in the course of the experiment
412 ($M(\text{younger adults}) = 0.007$, $SD(\text{younger adults}) = 0.005$, $M(\text{older adults}) = 0.025$, $SD(\text{older adults}) =$
413 0.02 ; $t(111) = -7.04$, $p = 1.6 \times 10^{-10}$). Those item pairs were excluded from further analyses.

414 Our experimental procedure was successful in inducing variability in memory performance such
415 that both groups could remember approximately half of the studied items: Young adults successfully
416 recalled on average 56.64 % (SD = 10.7) and older adults successfully recalled on average 41.6 %
417 (SD = 12.06) of the items (440 and 280, respectively). However, our procedure did not completely
418 eliminate age differences since young adults still performed significantly better than older
419 participants in the final recall task ($t(111) = 3.82, p = 0.0002$, two-sample t -test).

420 Due to the different number of items that younger and older adults studied in the course of the
421 experiment and the fact that the number of items included in the RSA was reduced based on the
422 smallest memory quality category (usually high quality), the number of items that were compared
423 to each other in the RSA also differed between groups: The median number of items included in the
424 RSA was 48 (range 10 to 101) for younger adults and 32 (range 5 to 79) for older adults. The groups
425 differed significantly in their respective item numbers ($z = 3.76, p = 0.0002$; Wilcoxon rank sum
426 test) which, however, did not affect group differences in pattern similarity (control analysis code
427 **are** available on Gitlab).

428 After the first study phase was completed, we asked participants to indicate on a four-point scale
429 how often they had used specific memory strategies for the task (1: almost always, 4: almost never).
430 With regard to the imagery strategy, young adults indicated that they had used it significantly more
431 often than older adults did (younger adults: median = 1.5, min = 1, max = 3; older adults: median = 2,
432 min = 1, max = 4; $z = -5.09, p = 0.0000004$, Wilcoxon rank sum test).

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

435

436 ***** Figure 4 *****

437

438 **Representational similarity**

439 Calculation of between-item representational similarity was based on the initial encoding phase
440 (Figure 1A). To identify whether high pattern similarity or high pattern distinctiveness during

441 learning was beneficial for later memory success, we sorted all items according to subsequent
442 memory performance and correlated the evoked spatiotemporal frequency pattern of each item
443 with every other item in the same memory quality category. The resulting mean similarity matrices
444 over all channels and scene–word pairs are shown in Figure 5A. These matrices display the
445 similarity of the frequency representations at all possible within-trial time point combinations
446 (-0.752 s to 3 s relative to stimulus onset at 0). In contrast, the diagonals of the similarity matrices
447 (also plotted separately in Figure 5B) only show the similarity between items at identical time
448 points and facilitate a visual comparison of the time courses of representational similarities for the
449 different memory quality categories and age groups. Although this omits much of the similarity
450 information, elevated similarities do occur largely along the diagonal. Note that the diagonals are
451 only plotted for illustration purposes and all statistical tests were performed on the complete
452 matrices as presented in Figure 5A.

453

454 *Older adults generally exhibit higher representational similarity than young adults*

455 Just before stimulus onset, similarity increased in both age groups and reached a peak around the
456 time of onset (cf. Figure 5A and B). Elevated similarity occurred mainly between identical trial time
457 points (diagonal) with slightly more persistent activity (elevated off-diagonal similarity) in older
458 adults compared to young adults. Irrespective of later memory success, between-item pattern
459 similarity was generally higher in older adults than in young adults (cf. Figure 5A and B; averaged
460 across the whole time–time matrix and all 60 channels: $M(\text{younger adults}) = 0.21$, $SD(\text{younger}$
461 $\text{adults}) = 0.065$, $M(\text{older adults}) = 0.25$, $SD(\text{older adults}) = 0.068$; 5000 cluster permutations, $p =$
462 0.0016). Furthermore, the level of pattern similarity and final memory performance were negatively
463 correlated across age groups ($r = -0.22$, $p = 0.02$; Pearson correlation). This is in line with previous
464 “dedifferentiation” findings and suggests that also on the within-person level better remembered
465 items should be less similar to each other. However, an across-group correlation may completely
466 differ from a within-group or even within-person correlation (Simpson's Paradox; Kievit et al.,

467 2013). Therefore, we further investigated the association of pattern similarity and memory quality
468 on the within-group and individual level.

469

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

471 Within both age groups, we tested for differences in the levels of representational similarity between
472 scene-word pairs of different memory quality by conducting linear regressions (low < medium <
473 high). We controlled for multiple comparisons by using non-parametric cluster-based permutation
474 tests. In both age groups we identified a cluster with a Monte Carlo p -value below 0.025, which
475 indicates a reliable linear relationship between representational similarity and memory quality
476 (young adults: $p = 0.02$; older adults: $p = 0.003$; see Figure 5C). Importantly, the direction of this
477 relationship differed between groups: while the relation between similarity and memory quality was
478 positive in older adults (low < medium < high), it was negative in young adults (low > medium > high)
479 (Figure 5E).

480 The cluster obtained in older adults included most of the diagonal from 50 ms to 830 ms after
481 stimulus onset and extended off-diagonally to 470 ms before and 1240 ms after stimulus onset (Figure
482 5C). Elevated similarity along the diagonal indicates similarity between neural representational
483 patterns at identical trial time points, whereas off-diagonal time windows suggest similar activation
484 patterns at different trial time points. The larger the distance of a coordinate from the diagonal, the
485 more distant are the compared time points in the respective frequency patterns. Differences between
486 memory quality categories were reliable in most (49 out of 60) occipital, parietal, temporal, and
487 central electrodes in older adults (Figure 5D).

488 In contrast to the cluster found in older adults, an off-diagonal cluster was identified for young
489 adults, in which low memory quality items displayed significantly more similarity than medium and
490 high memory quality items (Figure 5C). Compared to older adults, where differences between memory
491 qualities were found to be most pronounced between early and neighboring trial time points, i.e., close
492 to the diagonal, the off-diagonal cluster identified in young adults indicated that differences occurred
493 at later and more distant trial time points. Specifically, differences were found between earlier (450

494 ms to 1400 ms after stimulus onset) and later time points (2640 ms to 2800 ms after onset) and at 34
495 mainly parietal-occipital and central electrodes (Figure 5D). Despite the relatively poor spatial
496 resolution in EEG, the large electrode clusters in both young and older adults indicate that the
497 encoding-related patterns of neural activity that proved to be indicative of subsequent memory were
498 broadly distributed across the brain rather than specific to a particular region.

499 Additional analyses of pairwise comparison of the three memory quality categories instead of
500 linear regression resulted in a significant cluster that extended over similar time–time intervals and
501 electrodes only for high versus medium memory quality items in younger adults (high vs medium: $p =$
502 0.008; high vs low: $p = 0.03$; medium vs low: $p = 0.6$; 1000 cluster permutations), and high versus
503 medium as well as high versus low quality in older adults (high vs medium: $p = 0.004$; high vs low: $p =$
504 0.006; medium vs low: $p = 0.3$; 1000 cluster permutations).

505 In order to demonstrate that the effects obtained for the young adult group and the older adult group
506 appeared at different times during stimulus encoding, we extracted the most extreme z -value (z -
507 transformed regression coefficients) within the respective cluster from each subject and compared
508 their coordinates in time–time space. We fitted two models to test whether it was more likely that the
509 time points come from an identical multivariate normal distribution (single model) or from two
510 distinct distributions (two-group model). We then compared the two models using a Chi-Squared test
511 for model comparison with the null hypothesis that both models fit equally well. The two models
512 differed in model fit ($p = 10^{-54}$), and the two-group model showed significantly better fit. This
513 demonstrates that the effects obtained for young and older adults appeared at different times during
514 stimulus encoding.

515

516 *Age and memory quality effects in the identified clusters*

517 The cluster-based analyses reported above suggested differential memory-related representational
518 similarity in younger and older adults. To explore potential age differences in more depth, we
519 additionally tested for a linear relationship between representational similarity and memory quality in
520 each participant by conducting individual linear regressions. We then extracted and averaged the

521 individual z-transformed regression coefficients within each time–time–electrode cluster (see Figure
522 5E). In the young-adult cluster only the mean regression coefficients of the young adults differed from
523 zero (young adults: $t(49) = -3.42, p = 0.0013$; older adults: $t(62) = 1.79, p = 0.08$; one-sample t -tests)
524 and vice versa in the older-adult cluster (young adults: $t(49) = 0.75, p = 0.46$; older adults: $t(62) = 5.27,$
525 $p = 0.000002$). In both clusters the regression coefficients differed significantly between younger and
526 older adults (young-adult cluster: $M(\text{young adults}) = -0.27, SD(\text{young adults}) = 0.57, M(\text{older adults}) =$
527 $0.086, SD(\text{older adults}) = 0.38, t(111) = -4.03, p = 0.0001$; older-adults cluster: $M(\text{young adults}) =$
528 $0.058, SD(\text{young adults}) = 0.55, M(\text{older adults}) = 0.29, SD(\text{older adults}) = 0.43, t(111) = -2.5, p =$
529 0.014 ; independent samples t -tests) implying that age differences do exist in the relation between
530 representational similarity and memory quality in these clusters.

531

532 ***** Figure 5 *****

533

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

535 After each study trial, participants indicated on a four-point scale how well they were able to
536 integrate the presented scene and word (1: not well, 4: very well). We calculated the frequencies of
537 how often each rating was given by each participant. To test the strength of the relationship
538 between participants' imagery ratings and memory quality (see Memory performance and strategy
539 use) on the group and within-person level, we conducted non-parametric Goodman and Kruskal's
540 Gamma correlations for ordinally scaled data. For both groups, we obtained significant positive
541 relationships showing that higher imagery ratings were given to items of higher memory quality
542 (young adults: $\gamma = 0.28, z = 32.04, p = 10^{-223}$; older adults: $\gamma = 0.13, z = 11.04, p = 10^{-27}$). The
543 individual z -values from the within-person correlations of young and older adults differed
544 significantly ($t = 7.08, p = 10^{-10}$; two sample t -test) indicating a stronger link between imagery
545 ratings and memory success in young adults.

546 We further analyzed the association between these individual gamma correlations and the
547 individual regression coefficients from the representational similarity analyses (see *Age and*

548 *memory quality effects in the identified clusters*). Across both groups (but not within either group),
549 individual z-values from the gamma correlations and the individual regression coefficients (Figure
550 5E) showed a negative association ($r = -0.27$, $p = 0.005$, Pearson correlation). This means that the
551 lower (more negative) the individual regression coefficient (lower similarity in higher memory
552 quality; i.e., the effect seen in young adults), the stronger was the link between imagery rating and
553 memory quality. Equivalently, the higher (more positive) the individual regression coefficient
554 (higher similarity in higher memory quality; i.e., the effect seen in older adults), the weaker was the
555 link between imagery rating and memory quality.

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

563

564

565 **Discussion**

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

572 For older adults, between-item representational similarity was generally higher compared to young
573 adults, supporting the "dedifferentiation" hypothesis of declining neural distinctiveness with age
574 (Baltes and Lindenberger, 1997; Carp et al., 2011; S.-C. Li et al., 2004; D. Park et al., 2004, J. Park et al.,

575 2012; Payer et al., 2006; St-Laurent et al., 2014). Previous studies suggested that the loss of neural
576 specificity in old age may underlie age-related cognitive impairments. This was, for example,
577 supported by the finding that neural distinctiveness and fluid intelligence were correlated (D. Park et
578 al., 2010). However, most previous studies were not able to link neural item specificity directly with
579 participants' performance since memory for the items themselves was not assessed. By measuring
580 between-item representational similarity during the encoding phase of an associative memory task
581 and sorting the trials according to subsequent memory performance, we were able to relate measures
582 of neural distinctiveness during encoding directly to later recall success. Notably, results obtained
583 from multivariate analyses like those carried out here mostly reflect within-subject variability rather
584 than differences between individuals (Davis et al. 2014b). We found that although older adults
585 remembered significantly fewer items and revealed overall higher between-item pattern similarity
586 than younger adults, on the within-subject level, items represented with high similarity to other items
587 were actually those that older adults remembered best.

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

595 Beneficial effects of representational similarity for memory have been reported before (Davis et al.,
596 2014a; LaRocque et al., 2013; Lu et al., 2015; Visser et al., 2013; Wagner et al., 2016; Ye et al., 2016)
597 and have been located to medial temporal lobe regions, whereas pattern distinctiveness supported
598 memory in the hippocampus (LaRocque et al., 2013). These pattern separation computations were
599 shown to be impaired for older adults (Shing et al., 2011; Wilson et al., 2006; Yassa et al., 2011). While
600 high distinctiveness may be beneficial for memory performance to prevent false memories, high
601 similarity may support mnemonic decisions by capturing regularities across experiences (LaRocque et

602 al., 2013) and by giving rise to a feeling of familiarity (Davis et al., 2014a). Higher pattern similarity
603 may also reflect more consistent processing that facilitates associative memory formation (Wagner et
604 al., 2016).

605 Since neural activity patterns contain information on both content and processes (**cf.** Multivariate
606 EEG analysis), the age differences reported here could reflect differences in the similarity of memory
607 representations, processes, or both. The observed benefit of early high pattern similarity in older
608 adults may indeed reflect similarities in processing, e.g., increased attention to the stimuli and/or gist
609 extraction. Trials in which similar, memory-beneficial processes occur would be associated with
610 higher between-item pattern similarity and they would have a higher recall probability. However, our
611 findings may also refer to age differences on the representational level: A tendency toward more
612 generalized memories (Koutstaal and Schacter, 1997; Koutstaal et al., 2001; Tun et al., 1998) is often
613 reported for older adults and may also be associated with increased neural similarity. In our study, age
614 differences in the subjective judgements of imagery strategy use during encoding suggest that older
615 adults did indeed rely more on encoding the general gist of scene–word pairs while young adults more
616 often used the imagery strategy to create and encode unique details (**cf.** Hertzog et al., 2012).
617 Moreover, imagery and memory success were more strongly associated in young compared to older
618 adults, and more strongly linked to the association of pattern similarity and memory quality. Older
619 adults' benefit from successful early gist extraction may thus be reflected in increased early similarity,
620 whereas young adults' formation of mental images with distinct details may be reflected in increased
621 later dissimilarity.

622 The negative relationship between pattern similarity and memory performance in younger adults
623 that we report in the current study contrasts with other memory studies that showed a positive
624 relation, namely for recognition memory (LaRocque et al., 2013; Lu et al., 2015; Ye et al., 2016),
625 memory confidence and categorization (Davis et al., 2014a), fear memory (Visser et al., 2013), and
626 associative memory (Wagner et al., 2016). This could be due to the fact that most previous studies
627 showed a beneficial effect of neural similarity for performance in *recognition* tasks (but compare
628 Wagner et al., 2016), in which a sense of familiarity due to high pattern similarity (Davis et al., 2014a;

629 Gillund and Shiffrin, 1984) can be sufficient. In contrast to that, *recall* tasks as adopted in the current
630 study typically require retrieval of specific details of the studied items (Craik and Tulving, 1975).
631 Therefore, the observed benefit of distinct neural activation patterns for young adults' performance
632 here may be due to the deployed intentional learning task in which participants were explicitly
633 instructed to form very distinct mental images of the corresponding scene–word pairs. Furthermore,
634 similarity of event-related potentials such as that observed by Lu et al. (2015) may result in different
635 effects than in the time–frequency domain.

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

650 So far, the prevailing evidence on the relationship between representational similarity and memory
651 has been based on fMRI studies and therefore lacks insights into the temporal dynamics of pattern
652 similarity during the formation of memory representations. Here, we demonstrate the advantage of
653 dissociating different parts within the trial time course that reveal distinctions in the way
654 representational similarity relates to memory performance of younger and older adults. Furthermore,
655 the present study provides further evidence for the high relevance of the rich neural signatures

656 offered by a wide range of frequencies and across multiple topographical sites for memory encoding
657 and extends previous research with similar approaches (cf. Kerrén et al., 2018; Michelmann et al.,
658 2016; Staresina et al., 2016; Zhang et al., 2015).

659 The question remains how between-item similarity links to within-item similarity, i.e. item-specific
660 representational stability (across item repetitions) and reinstatement (between encoding and
661 retrieval). Recent research suggests that within-item similarity benefits memory performance (Lu et
662 al., 2015; Xue, 2018; Xue et al., 2010) and declines during aging (St-Laurent et al., 2014; Zheng et al.,
663 2017). Understanding the mutual influences of between-item similarity, pattern stability, and pattern
664 reinstatement may be crucial to complete our comprehension of how memories are represented and
665 processed in the brain across the lifespan.

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

675

676 **References**

- 677 Abdulrahman, H., Fletcher, P.C., Bullmore, E., and Morcom, A.M. (2017). Dopamine and memory
678 dedifferentiation in aging. *NeuroImage* 153, 211–220.
- 679 Allefeld, C., Gørgen, K., and Haynes, J.D. (2016). Valid population inference for information-based
680 imaging: From the second-level t-test to prevalence inference. *NeuroImage* 141, 378–392.
- 681 Allen, M., Poggiali, D., Whitaker, K., Marshall, T.R., and Kievit, R.A. (2019). Raincloud plots: a multi-
682 platform tool for robust data visualization. *Wellcome Open Research* 4, 63.
- 683 Baltes, P.B., and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and
684 cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychology*
685 *and Aging* 12, 12–21.
- 686 Buzsáki, G., and Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science* 304, 1926–
687 1929.
- 688 Cai, D.J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., Wei, B., Veshkini, M., La-Vu, M., Lou, J., et al.
689 (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* 534,
690 115–118.
- 691 Caplan, J.B., Madsen, J.R., Raghavachari, S., and Kahana, M.J. (2001). Distinct patterns of brain
692 oscillations underlie two basic parameters of human maze learning. *Journal of Neurophysiology* 86,
693 368–380.
- 694 Carp, J., Gmeindl, L., and Reuter-Lorenz, P.A. (2010). Age differences in the neural representation of
695 working memory revealed by multi-voxel pattern analysis. *Frontiers in Human Neuroscience* 4, 217.
- 696 Carp, J., Park, J., Polk, T.A., and Park, D.C. (2011). Age differences in neural distinctiveness revealed by
697 multi-voxel pattern analysis. *NeuroImage* 56, 736–743.

- 698 Cichy, R.M., Pantazis, D., and Oliva, A. (2014). Resolving human object recognition in space and time.
699 *Nature Neuroscience* *17*, 455–462.
- 700 Clark, S.E., and Gronlund, S.D. (1996). Global matching models of recognition memory: How the models
701 match the data. *Psychonomic Bulletin & Review* *3*, 37–60.
- 702 Cohen, M.X. (2011). It's about Time. *Frontiers in Human Neuroscience* *5*, 2.
- 703 Craik, F.I.M., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory.
704 *Journal of Experimental Psychology: General* *104*, 268–294.
- 705 Daselaar, S.M., Fleck, M.S., Dobbins, I.G., Madden, D.J., and Cabeza, R. (2006). Effects of healthy aging on
706 hippocampal and rhinal memory functions: an event-related fMRI study. *Cerebral Cortex* *16*, 1771–
707 1782.
- 708 Davis, T., Xue, G., Love, B.C., Preston, A.R., and Poldrack, R.A. (2014a). Global neural pattern similarity as
709 a common basis for categorization and recognition memory. *Journal of Neuroscience* *34*, 7472–7484.
- 710 Davis, T., LaRocque, K.F., Mumford, J.A., Norman, K.A., Wagner, A.D., and Poldrack, R.A. (2014b). What do
711 differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level
712 variance impact fMRI analysis. *NeuroImage* *97*, 271–283.
- 713 Duverne, S., Habibi, A., and Rugg, M.D. (2008). Regional specificity of age effects on the neural correlates
714 of episodic retrieval. *Neurobiology of Aging* *29*, 1902–1916.
- 715 Fandakova, Y., Sander, M.C., Grandy, T.H., Cabeza, R., Werkle-Bergner, M., and Shing, Y.L. (2018). Age
716 differences in false memory: The importance of retrieval monitoring processes and their modulation by
717 memory quality. *Psychology and Aging* *33*, 119–133.
- 718 Fandakova, Y., Werkle-Bergner, M., and Sander, M.C. (2019). (Only) time can tell: Age differences in false
719 memory are magnified at longer delays. *PsyArXiv* doi:10.31234/osf.io/eh5x7.

- 720 Folstein, M.F., Folstein, S.E., and McHugh, P.R. (1975). "Mini-mental state": A practical method for
721 grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12, 189–198.
- 722 Gillund, G., and Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. *Psychological*
723 *Review* 91, 1–67.
- 724 Goh, J.O., Suzuki, A., and Park, D.C. (2010). Reduced neural selectivity increases fMRI adaptation with
725 age during face discrimination. *NeuroImage* 51, 336–344.
- 726 Hanslmayr, S., and Staudigl, T. (2014). How brain oscillations form memories — A processing based
727 perspective on oscillatory subsequent memory effects. *NeuroImage* 85, 648–655.
- 728 Hertzog, C., Price, J., and Dunlosky, J. (2012). Age differences in the effects of experimenter-instructed
729 versus self-generated strategy use. *Experimental Aging Research* 38, 42–62.
- 730 Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., and Sejnowski, T.J. (2000).
731 Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178.
- 732 Kerrén, C., Linde-Domingo, J., Hanslmayr, S., and Wimber, M. (2018). An optimal oscillatory phase for
733 pattern reactivation during memory retrieval. *Current Biology* 28, 3383–3392.
- 734 Kievit, R.A., Frankenhuis, W.E., Waldorp, L.J., and Borsboom, D. (2013). Simpson's paradox in
735 psychological science: A practical guide. *Frontiers in Psychology* 4, 513.
- 736 King, J.R., and Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal
737 generalization method. *Trends in Cognitive Sciences* 18, 203–210.
- 738 Koen, J.D., Hauck, N., and Rugg, M.D. (2019). The relationship between age, neural differentiation, and
739 memory performance. *The Journal of Neuroscience* 39, 149–162.
- 740 Kosciessa, J.Q., Grandy, T.H., Garrett, D.D., and Werkle-Bergner, M. (2018). Single-trial characterization
741 of neural rhythms: potentials and challenges. *BioRxiv* doi:10.1101/356089.

- 742 Koutstaal, W., and Schacter, D.L. (1997). Gist-based false recognition of pictures in older and younger
743 adults. *Journal of Memory and Language* 37, 555–583.
- 744 Koutstaal, W., Schacter, D.L., and Brenner, C. (2001). Dual task demands and gist-based false recognition
745 of pictures in younger and older adults. *Journal of Memory and Language* 44, 399–426.
- 746 Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting
747 the branches of systems neuroscience. *Frontiers in Systems Neuroscience* 2,4.
- 748 LaRocque, K.F., Smith, M.E., Carr, V.A., Witthoft, N., Grill-Spector, K., and Wagner, A.D. (2013). Global
749 similarity and pattern separation in the human medial temporal lobe predict subsequent memory.
750 *Journal of Neuroscience* 33, 5466–5474.
- 751 Li, S.-C., and Sikström, S. (2002). Integrative neurocomputational perspectives on cognitive aging,
752 neuromodulation, and representation. *Neuroscience & Biobehavioral Reviews* 26, 795–808.
- 753 Li, J., Morcom, A.M., and Rugg, M.D. (2004a). The effects of age on the neural correlates of successful
754 episodic retrieval: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience* 4, 279–293.
- 755 Li, S.-C., Lindenberger, U., Sikström, S., Lindenberger, U., Sikstrom, S., and Sikstrom, S. (2001). Aging
756 cognition: From neuromodulation to representation. *Trends in Cognitive Sciences* 5, 479–486.
- 757 Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., and Baltes, P.B. (2004b).
758 Transformations in the couplings among intellectual abilities and constituent cognitive processes
759 across the life span. *Psychological Science* 15, 155–163.
- 760 Lu, Y., Wang, C., Chen, C., and Xue, G. (2015). Spatiotemporal neural pattern similarity supports episodic
761 memory. *Current Biology* 25, 780–785.
- 762 Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of*
763 *Neuroscience Methods* 164, 177–190.

- 764 McDaniel, M.A., and Einstein, G.O. (1986). Bizarre imagery as an effective memory aid: The importance
765 of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition* *12*, 54–65.
- 766 Michelmann, S., Bowman, H., and Hanslmayr, S. (2016). The temporal signature of memories:
767 Identification of a general mechanism for dynamic memory replay in humans. *PLoS Biology* *14*,
768 e1002528.
- 769 Michelmann, S., Bowman, H., and Hanslmayr, S. (2018). Replay of stimulus-specific temporal patterns
770 during associative memory formation. *Journal of Cognitive Neuroscience* *30*, 1577–1589.
- 771 Morcom, A.M., Li, J., and Rugg, M.D. (2007). Age Effects on the Neural Correlates of Episodic Retrieval:
772 Increased Cortical Recruitment with Matched Performance. *Cerebral Cortex* *17*, 2491–2506.
- 773 Muehlroth, B.E., Sander, M.C., Fandakova, Y., Grandy, T.H., Rasch, B., Shing, Y.L., and Werkle-Bergner, M.
774 (2019). Precise Slow Oscillation–Spindle Coupling Promotes Memory Consolidation in Younger and
775 Older Adults. *Scientific Reports* *9*, 1940.
- 776 Nagel, I.E., Preuschhof, C., Li, S.-C., Nyberg, L., Bäckman, L., Lindenberger, U., and Heekeren, H.R. (2009).
777 Performance level modulates adult age differences in brain activation during spatial working memory.
778 *Proceedings of the National Academy of Sciences* *106*, 22552–22557.
- 779 Nolan, H., Whelan, R., and Reilly, R.B. (2010). FASTER: Fully automated statistical thresholding for EEG
780 artifact rejection. *Journal of Neuroscience Methods* *192*, 152–162.
- 781 Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open source software for
782 advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and*
783 *Neuroscience* *2011*, 156869.
- 784 Paller, K.A., and Wagner, A.D. (2002). Observing the transformation of experience into memory. *Trends*
785 *in Cognitive Sciences* *6*, 93–102.

- 786 Park, D.C., Polk, T.A., Park, R., Minear, M., Savage, A., and Smith, M.R. (2004). Aging reduces neural
787 specialization in ventral visual cortex. *Proceedings of the National Academy of Sciences of the United*
788 *States of America* *101*, 13091–13095.
- 789 Park, J., Carp, J., Hebrank, A., Park, D.C., and Polk, T.A. (2010). Neural specificity predicts fluid processing
790 ability in older adults. *Journal of Neuroscience* *30*, 9253–9259.
- 791 Park, J., Carp, J., Kennedy, K.M., Rodrigue, K.M., Bischof, G.N., Huang, C.-M., Rieck, J.R., Polk, T.A., and Park,
792 D.C. (2012). Neural broadening or neural attenuation? Investigating age-related dedifferentiation in the
793 face network in a large lifespan sample. *Journal of Neuroscience* *32*, 2154–2158.
- 794 Payer, D., Marshuetz, C., Sutton, B., Hebrank, A., Welsh, R.C., and Park, D.C. (2006). Decreased neural
795 specialization in old adults on a working memory task. *Neuroreport* *17*, 487–491.
- 796 Perrin, F., Pernier, J., Bertrand, O., and Echallier, J.F. (1989). Spherical splines for scalp potential and
797 current density mapping. *Electroencephalography and Clinical Neurophysiology* *72*, 184–187.
- 798 Rugg, M.D., and Morcom, A.M. (2005). The Relationship Between Brain Activity, Cognitive Performance,
799 and Aging: The Case of Memory. In *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral*
800 *Aging*, (New York, NY, US: Oxford University Press), pp. 132–154.
- 801 Sander, M.C., Fandakova, Y., Grandy, T.H., Shing, Y.L., and Werkle-Bergner, M. (2019). Oscillatory
802 mechanisms of successful memory formation in younger and older adults are related to structural
803 integrity. *BioRxiv*. doi:10.1101/530121.
- 804 Schönauer, M., Alizadeh, S., Jamalabadi, H., Abraham, A., Pawlizki, A., and Gais, S. (2017). Decoding
805 material-specific memory reprocessing during sleep in humans. *Nature Communications* *8*, 15404.
- 806 Shing, Y.L., Rodrigue, K.M., Kennedy, K.M., Fandakova, Y., Bodammer, N., Werkle-Bergner, M.,
807 Lindenberger, U., and Raz, N. (2011). Hippocampal subfield volumes: Age, vascular risk, and correlation
808 with associative memory. *Frontiers in Aging Neuroscience* *3*, 2.

- 809 Staresina, B.P., Michelmann, S., Bonnefond, M., Jensen, O., Axmacher, N., and Fell, J. (2016). Hippocampal
810 pattern completion is linked to gamma power increases and alpha power decreases during recollection.
811 *eLife* 5, e17397.
- 812 St-Laurent, M., Abdi, H., Burianová, H., and Grady, C.L. (2011). Influence of aging on the neural correlates
813 of autobiographical, episodic, and semantic memory retrieval. *Journal of Cognitive Neuroscience* 23,
814 4150–4163.
- 815 St-Laurent, M., Abdi, H., Bondad, A., and Buchsbaum, B.R. (2014). Memory reactivation in healthy aging:
816 Evidence of stimulus-specific dedifferentiation. *Journal of Neuroscience* 34, 4175–4186.
- 817 Tulving, E., and Bower, G.H. (1974). The Logic of Memory Representations. *Psychology of Learning and*
818 *Motivation* 8, 265–301.
- 819 Tun, P.A., Wingfield, A., Rosen, M.J., and Blanchard, L. (1998). Response latencies for false memories:
820 gist-based processes in normal aging. *Psychology and Aging* 13, 230–241.
- 821 Visser, R.M., Scholte, H.S., Beemsterboer, T., and Kindt, M. (2013). Neural pattern similarity predicts
822 long-term fear memory. *Nature Neuroscience* 16, 388–390.
- 823 Wagner, I.C., van Buuren, M., Bovy, L., and Fernandez, G. (2016). Parallel engagement of regions
824 associated with encoding and later retrieval forms durable memories. *Journal of Neuroscience* 36,
825 7985–7995.
- 826 Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition.
827 *Physiological Reviews* 90, 1195–1268.
- 828 Wang, T.H., Johnson, J.D., de Chastelaine, M., Donley, B.E., and Rugg, M.D. (2015). The Effects of Age on
829 the Neural Correlates of Recollection Success, Recollection-Related Cortical Reinstatement, and Post-
830 Retrieval Monitoring. *Cerebral Cortex* 26, 1698–1714.

831 Whitten, T.A., Hughes, A.M., Dickson, C.T., and Caplan, J.B. (2011). A better oscillation detection method
832 robustly extracts EEG rhythms across brain state changes: The human alpha rhythm as a test case.
833 *NeuroImage* 54, 860–874.

834 Wilson, I.A., Gallagher, M., Eichenbaum, H., and Tanila, H. (2006). Neurocognitive aging: prior memories
835 hinder new hippocampal encoding. *Trends in Neurosciences* 29, 662–670.

836 Xue, G. (2018). The neural representations underlying human episodic memory. *Trends in Cognitive*
837 *Sciences* 22, 544–561.

838 Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J.A., and Poldrack, R.A. (2010). Greater neural pattern
839 similarity across repetitions is associated with better memory. *Science* 330, 97–101.

840 Yassa, M.A., Mattfeld, A.T., Stark, S.M., and Stark, C.E.L. (2011). Age-related memory deficits linked to
841 circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences of the*
842 *United States of America* 108, 8873–8878.

843 Ye, Z., Zhu, B., Zhuang, L., Lu, Z., Chen, C., and Xue, G. (2016). Neural global pattern similarity underlies
844 true and false memories. *Journal of Neuroscience* 36, 6792–6802.

845 Zhang, H., Fell, J., Staresina, B.P., Weber, B., Elger, C.E., and Axmacher, N. (2015). Gamma power
846 reductions accompany stimulus-specific representations of dynamic events. *Current Biology* 25, 635–
847 640.

848 Zheng, L., Gao, Z., Xiao, X., Ye, Z., Chen, C., and Xue, G. (2017). Reduced fidelity of neural representation
849 underlies episodic memory decline in normal aging. *Cerebral Cortex* 28, 2283–2296.

850

851

852

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

861
862 **Figure 2:** Scoring of stimulus pairs into high, medium, or low memory quality categories based on
863 learning history. For both younger and older adults, items that were correctly recalled in the last recall
864 cycle (C) as well as the previous one (B) were scored as high memory quality items. Pairs that were
865 solely recalled in the final recall were scored as medium memory quality items. And items that were
866 never correctly recalled were scored as low memory quality items. Not depicted: Items that were
867 recalled in the earlier but not later recall were excluded. Older adults performed one more cued recall
868 and restudy cycle (between A and B) that was not included in item scoring due to close-to-floor
869 performance. Note that wrong and missing responses were treated equally.

870
871 **Figure 3:** Spectral representational similarity analysis methodology. **A.** The frequency vector from
872 every time point (i.e., column) of the noise-corrected and log-transformed time–frequency pattern
873 (from one electrode) corresponding to stimulus 1 (bottom) is Pearson-correlated with the vectors from
874 every time point of stimulus 2 (left; rotated by 90°). For illustration, sample vectors of stimulus 1 (t_{s1})
875 **and** stimulus 2 (t_{s2}) are highlighted. Correlating these two vectors gives one correlation coefficient, i.e.,
876 one coordinate (indicated by black box) on a matrix with time on both axes. Computing all pairwise
877 time vector correlations results in a time–time similarity matrix representing the similarity of those two
878 frequency patterns at all time point combinations. This procedure is repeated for all items of a certain
879 memory quality (i.e., similarity of stimulus 1 with all others, stimulus 2 with all others, etc.). **B.**

880 Averaging across all similarity matrices yields the mean similarity matrix showing the pattern similarity
881 among all items of the same memory quality. Only one triangle and the diagonal of the matrix are
882 relevant because the similarity of every frequency pattern pair is computed twice, resulting in an
883 identical correlation coefficient on both sides of the diagonal. Similarity is quantified as the Fisher z -
884 transformed Pearson correlation coefficient (z'). Not depicted: This procedure was repeated for all 60
885 electrodes, the three memory quality categories, and all subjects.

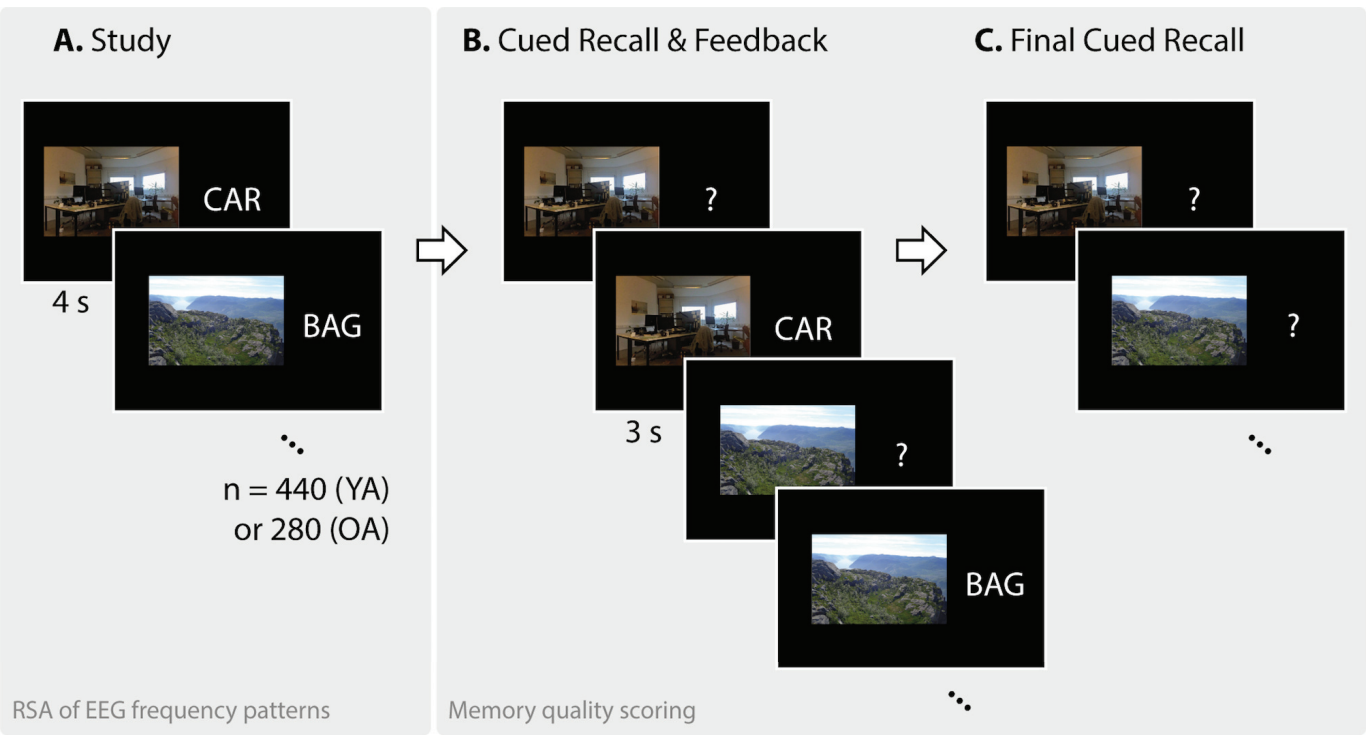
886

887 **Figure 4:** Proportion of item pairs with high, medium, and low memory quality as well as proportion of
888 excluded items for 50 young adults (YA; blue) and 63 older adults (OA; red). Group distributions as un-
889 mirrored violin plots (probability density functions), boxplots with means and 95% confidence
890 intervals, whiskers with 2nd and 98th percentiles, and individual data points (horizontally jittered)
891 (modified from Allen et al., 2019). Note that the y-axis for excluded items differs from that of the other
892 categories. YA studied 440 pairs and OA studied 280 pairs.

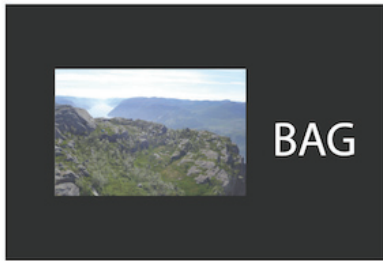
893

894 **Figure 5:** Between-item pattern similarities and statistics. Similarity is quantified as Fisher z -
895 transformed Pearson correlation coefficient (z'). On time axes, zero denotes stimulus onset. C and D
896 show results from cluster-based permutation analyses for each age group, E shows results from
897 individual regression analyses (see Methods). **A.** Mean time–time similarity matrices across all 60
898 channels and items within each memory quality category (high, medium, low) for all 50 young adults
899 (YA; top) and 63 older adults (OA; bottom). Note that the scales differ between age groups. **B.** Diagonals
900 from the time–time similarity matrices (see A). **C.** Time–time clusters (masked z -scores) in which the
901 three memory quality categories differ significantly within each age group (averaged across reliable
902 electrodes, see D). Positive z -values (red) reveal a positive relationship between pattern similarity and
903 memory quality (higher similarity is associated with higher memory quality), negative z -values (blue)
904 reveal a negative relationship (lower similarity is associated with higher memory quality). **D.**
905 Topographic representations of the electrode clusters that revealed reliable differences between
906 memory quality categories within each age group (averaged across reliable time windows, see C). **E.** Z -

907 transformed regression coefficients extracted from time–time–electrode clusters identified in YA (left)
908 and OA (right) (see C and D). Group distributions (probability density functions), boxplots with means
909 and 95% confidence intervals, whiskers with 2nd and 98th percentiles, and individual data points
910 (horizontally jittered) for YA (blue) and OA (red) (modified from Allen et al., 2019). *P*-values are given
911 for group differences within each cluster (independent samples *t* -tests). *Note the difference between z'*
912 *(Fisher z-transformed correlation coefficients) and z (z-transformed regression coefficients).*



A. Study



B. Cued Recall (& Feedback)

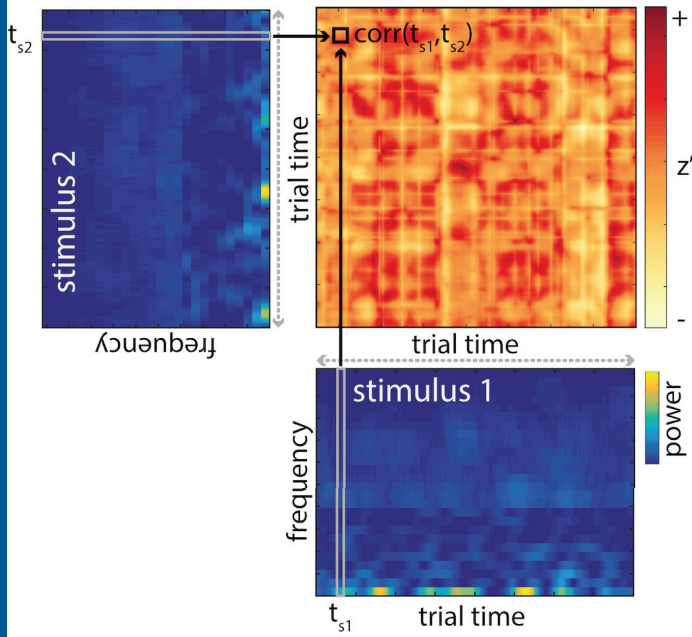


C. Final Cued Recall



Memory Quality	high	"Bag" ✓	"Bag" ✓
	medium	? ✗	"Bag" ✓
	low	? ✗	? ✗

A. Computing time-time similarity matrix between frequency patterns of two stimuli



Repeat for all stimulus combinations and average (half-)matrices

B. Similarity matrix between all stimuli of one memory quality

