

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/150733/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Sommer, Tobias, Hennies, Nora, Lewis, Penelope A. ORCID: <https://orcid.org/0000-0003-1793-3520> and Alink, Arjen 2022. The assimilation of novel information into schemata and its efficient consolidation. *Journal of Neuroscience* 42 (30) , pp. 5916-5929. 10.1523/jneurosci.2373-21.2022 file

Publishers page: <https://doi.org/10.1523/jneurosci.2373-21.2022>  
<<https://doi.org/10.1523/jneurosci.2373-21.2022>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



*Research Articles: Behavioral/Cognitive*

## The assimilation of novel information into schemata and its efficient consolidation

<https://doi.org/10.1523/JNEUROSCI.2373-21.2022>

**Cite as:** J. Neurosci 2022; 10.1523/JNEUROSCI.2373-21.2022

Received: 1 December 2021

Revised: 7 June 2022

Accepted: 11 June 2022

---

*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](http://www.jneurosci.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

1     **The assimilation of novel information into schemata and its**  
2                                   **efficient consolidation**

3                   Short title: Assimilation and consolidation of schema-related facts  
4           Tobias Sommer<sup>1†</sup>, Nora Hennies<sup>1</sup>, Penelope A. Lewis<sup>2\*</sup>, Arjen Alink<sup>1\*</sup>

5  
6     <sup>1</sup> Institute of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg,  
7     Germany

8     <sup>2</sup> CUBRIC, School of Psychology, Cardiff University, Cardiff CF24 4HQ, UK

9     <sup>†</sup> corresponding author: [tsommer@uke.de](mailto:tsommer@uke.de)

10    \* these author contributed equally

11  
12

13    No or pages: 29

14    No of figures: 6

15    No of table: 1

16    No of words: abstract 207, introduction 650, discussion 1478

17    No conflicts of interest.

18    No acknowledgement

19 **Abstract**

20 Schemata enhance memory formation for related novel information. This is true even  
21 when this information is neutral with respect to schema-driven expectations. This  
22 assimilation of novel information into schemata has been attributed to more effective  
23 organizational processing that leads to more referential connections with the activated  
24 associative schema network. Animal data suggest that systems consolidation of novel  
25 assimilated information is also accelerated. In the current study, we used both multivariate  
26 and univariate fMRI analyses to provide further support for these proposals and to  
27 elucidate the neural underpinning of these processes. 28 Participants (5 male) over-  
28 learned fictitious schemata for seven weeks and then encoded novel related and control  
29 facts in the scanner. These facts were retrieved both immediately and two weeks later,  
30 also in the scanner. Our results conceptually replicate previous findings with respect to  
31 enhanced vmPFC-hippocampus coupling during encoding of novel related information and  
32 point to a prior knowledge-effect that is distinct from situations where novel information is  
33 experienced as congruent or incongruent with a schema. Moreover, the combination of  
34 both multi- and univariate results further specified the proposed contributions of the  
35 vmPFC, precuneus and angular gyrus-network to the more efficient encoding of schema-  
36 related information. In addition, our data provide further evidence for more efficient  
37 systems consolidation of such novel schema-related and potentially assimilated  
38 information.

39

40 **Significance Statement**

41 Our prior-knowledge in a certain domain, often termed schema, heavily influences whether  
42 and how we form memories for novel information that can be related to them. The results  
43 of the current study show how a ventromedial prefrontal-precuneal-angular network  
44 contributes to the more efficient encoding of novel related information. Furthermore, the  
45 observed increase in prefrontal-hippocampal coupling during this process points to a  
46 critical distinction from the previously described mechanisms supporting the encoding of  
47 information that is experienced as congruent with schema-driven expectations. In addition,  
48 we find further support for the proposal based on animal data that prior-knowledge  
49 enhances also the consolidation of schema-related information.

## 50 Introduction

51 Whether and how we encode, consolidate and later retrieve novel information is heavily  
52 influenced by our prior-knowledge (Alba and Hasher, 1983; Gilboa and Marlatte, 2017).  
53 The impact of prior-knowledge on memory has been studied in the framework of schemata  
54 (Bartlett, 1932) which are previously acquired and continuously developing associative  
55 networks (Ghosh and Gilboa, 2014).

56 A prominent line of research on the prior-knowledge effect investigates how the  
57 congruency of information with schema-driven expectations affects memory, e.g.  
58 remembering a palm-tree vs. a polar bear at a beach (van Kesteren et al., 2012; Greve et  
59 al., 2019). However, prior-knowledge enhances also memory for novel, previously  
60 unknown - and hence expectation-neutral - information that can be related to it, for  
61 instance when we learn new facts in our academic field (Brandt et al., 2005; Witherby and  
62 Carpenter, 2021). The vmPFC together with the ventral precuneus/retrosplenial cortex  
63 (vPC/RSC) are involved also in this effect of prior-knowledge where the contribution of the  
64 hippocampus and its coupling with the vmPFC remain so far inconsistent across studies  
65 (Tse et al., 2011; van Kesteren et al., 2014; Brod et al., 2016; Liu et al., 2016; Sommer,  
66 2017).

67 The mnemonic advantage for novel schema-related but expectation-neutral  
68 information has been attributed to the activation of schema-knowledge. This allows more  
69 effective organizational processing leading to assimilation into the associative structure  
70 (Ericsson and Kintsch, 1995). In particular, the newly encoded information might be  
71 integrated via spreading activations into the existing associative network resulting in  
72 referential connections and the association with appropriate retrieval cues (Long and Prat,  
73 2002).

74 These processes underlying assimilation would not necessarily result in enhanced  
75 mean activity in the involved brain areas compared to encoding of schema-unrelated  
76 information. The first goal of the current study was therefore to find support for such  
77 processes by using multivariate representational similarity analyses (RSA). In particular,  
78 after participants had acquired knowledge of experimental schemata (Fig. 1 A) we  
79 contrasted the similarity of activity patterns during three encoding rounds of novel schema-  
80 related with those of tightly matched control facts. By this means we tested the hypothesis  
81 that the rapid integration of schema-related facts into the associative structures results in  
82 more consistent representation across encoding rounds (i.e. greater pattern robustness;  
83 Xue et al., 2010; Bruett et al., 2020). Moreover, we tested whether the more effective

84 organizational processing would be reflected in more consistent encoding operations  
85 across novel schema-related facts. Finally, we tested the hypothesis that assimilation  
86 should be evident in higher similarity between encoding of novel schema-related facts and  
87 the retrieval of overlearned schema-knowledge.

88 A second goal of the current study was to provide further evidence with respect to  
89 hippocampal-vmPFC coupling during encoding of novel schema-related but expectation-  
90 neutral information. In the two studies observing stronger coupling participants encoded  
91 arbitrary associations (Liu et al., 2016; Sommer, 2017) whereas in the third students  
92 learned real facts, i.e. associations of an already taught with a new term, in their or another  
93 discipline (van Kesteren et al., 2014). The diverging results might be caused by the fact  
94 that in the other discipline both terms forming a new fact were novel (Carpenter et al.,  
95 2018) or that only in the latter study meaningful information was encoded. Both factors  
96 were addressed in the current paradigm with the goal to better understand hippocampal-  
97 vmPFC coupling during encoding of novel schema-related information.

98 A series of animal studies showed that prior-knowledge not only influences  
99 encoding but also accelerates systems consolidation of novel schema-related expectation-  
100 neutral information, i.e. randomly paired flavor-location associations (Tse et al., 2007). In  
101 the aforementioned study using a similar design, we showed that also in humans schema-  
102 related information might be more rapidly transferred from hippocampal to cortical retrieval  
103 (Sommer, 2017). Using meaningful knowledge structures we also showed that nightly  
104 replay might underlie this effect (Hennies et al., 2016). A third goal of the current study was  
105 therefore to use those meaningful associative structures to conceptually replicate our  
106 previous findings and to provide further evidence that novel schema-related expectation-  
107 neutral information is transferred more rapidly to neocortical retrieval. To test this  
108 hypothesis, participants' memory for novel facts was tested in the MRI scanner  
109 immediately after encoding and 2 weeks later.

## 110 **Methods**

### 111 ***Participants***

112 Thirty-two native Germans (mean age: 25.90 years; SD: 3.71 years; 6 males; randomly  
113 assigned to one of the two schemata) participated in the study. All had normal or  
114 corrected-to-normal vision, and no history of neurological and psychiatric disorders.  
115 Participants were required to have no knowledge beyond that of basic schooling at the  
116 primary level in the two schema categories (arthropods and cell biology) and no particular  
117 interest in biology, medicine, chemistry, and zoology. Informed consent was obtained from

118 all participants before the study, which was approved by the Ethics Committee of the  
119 Hamburg Medical Association. The first participant could not be analyzed due to data loss.  
120 Two participants dropped out of the study during the schema-knowledge acquisition. For  
121 one participant, retrieval results are missing due to technical failure, leaving 28 participants  
122 (5 male, 14 for each of the two schemata).

### 123 **Stimuli**

124 The aim was to experimentally construct two fictitious schemata that fulfilled the four  
125 previously identified criteria, i.e. an associative network structure, basis on multiple  
126 episodes, lack of unit detail, and adaptability, (Ghosh and Gilboa, 2014) and the  
127 neuroscientific definition, i.e. consolidated cortical representation (van Kesteren et al.,  
128 2012). At the same time, we aimed to minimize previously described confounds when pre-  
129 existing, e.g. academic or other expert, knowledge is used: i) higher  
130 curiosity/interest/motivation to learn novel facts in the domain of expertise and ii) the  
131 possibility that novel schema-related information had been known before (Witherby and  
132 Carpenter, 2021). Moreover, in the area of expertise participants have a higher familiarity  
133 with the names of the concepts the novel information has to be associated with, e.g. when  
134 “blossom end rot” has to be associated with calcium deficiency the term “blossom end rot”  
135 is for experts familiar but novel for control participants (Carpenter et al., 2018). Such  
136 greater cue familiarity results in a processing advantage and in differences in (meta-)  
137 memory (Chua et al., 2012). Importantly, as the goal of the current study was to  
138 characterize the effects of prior-knowledge on memory due to more effective  
139 organizational processing and not to in-/congruency, we aimed to minimize the influence of  
140 schema driven expectations with respect to the novel facts. However, we also aimed to  
141 use not arbitrary associative structures as in our previous and the animal studies (Tse et  
142 al., 2007, 2011; Sommer, 2017) but more ecological valid, i.e. meaningful experimental  
143 schemata. Novel facts related to such meaningful knowledge structures can be only  
144 neutral with respect to specific expectations based on prior experience (as for instance  
145 palm-trees and not polar bears are expected at a beach scenery) but are still consistent  
146 with more general knowledge. To stay with the above example, even if there is no specific  
147 schema-driven expectation for the cause for “blossom end rot” the novel fact “calcium  
148 deficiency” is similar to many other possibilities, e.g. sodium deficiency or a fungal attack,  
149 that are generally consistent with the student’s knowledge about phytology. The absence  
150 of specific schema-driven expectations results at the same time in more detailed, less  
151 generalized and gist-like memories (Tse et al., 2007, 2007; Sommer, 2017) which diverges

152 from the more typical schema studies on the effect of schema-congruency where higher  
153 false alarm rates are observed (Rojahn and Pettigrew, 1992; van Kesteren et al., 2012).

#### 154 *Schemata*

155 The stimuli from our previous study were translated into German and modified for the  
156 purpose of the current study (Hennies et al., 2016). The schemata were constructed with  
157 four hierarchical levels (Fig.1 A). The two schemata were parallel in structure and  
158 contained facts at each level. In particular, each schema comprised facts about the  
159 category (arthropods or cells), its two sub-categories (ant and crabs, cell-types and  
160 organelles), their three families each (e.g. symbiotic, hunter and weaver ants) and detailed  
161 characteristics of the 12 individual exemplars (two in each family). The detailed facts  
162 (approximately 25 facts) about anatomy, habitat, food preferences, and behavioral  
163 characteristics for each of the 12 exemplars comprised the main part of the 'arthropods'-  
164 schema. For the 'cells'-schema a matched number of facts for each of the 12 exemplars  
165 existed (see Hennies et al., 2016 for family labels and pictures of the exemplars of the  
166 'cells'-schema). In addition, each exemplar was presented on 12 different pictures. Novel  
167 names were invented for all exemplars in order to avoid large differences in length and  
168 complexity of the names. The whole hierarchical associative structure and all facts, i.e. the  
169 pictures of the exemplars, related to its nodes and elements were considered as schema -  
170 similar to studies employing pre-existing academic or trivial real-world knowledge.

#### 171 *Novel, schema-related facts*

172 For each of the exemplars of both schemata six additional facts were created. Each facts  
173 existed in two, equally likely alternatives (e.g. Styga is 2 or 4 cms long, NIV contains  
174 *copper* or *nickel*, Fig. 1B). One version, which was randomly chosen for each participant,  
175 was encoded as novel schema-related (SR) fact, the other one served as lure for the 2-  
176 alterantive forced-choice memory tests. This design ensured that participants could not  
177 guess the correct response based on their schema-knowledge and that the novel schema-  
178 related facts were indifferent with respect to schema congruency and expectations. The  
179 SR facts for the 'arthropods'-schema served as non-schema related (NS) control facts for  
180 the participants who acquired knowledge for the 'cells'-schema and vice versa. The facts  
181 were kept vague to minimize participants guessing about what type of exemplars the  
182 control facts were about. The novel facts for both schemata were counterbalanced for the  
183 number of words, number of syllables, and numerical values. None of the facts were



184 longer than eight words or 14 syllables so that all facts could easily be read within the  
185 presentation time.

186 **Procedure**

187 This experiment was realized with custom-written scripts using Cogent 2000, developed by  
188 the Cogent 2000 team at the Functional Imaging Laboratory and the Institute for Cognitive  
189 Neuroscience (University College, London).

190 *Acquisition of structured, associative schema-knowledge*

191 In a first session in the institute, participants were randomly assigned to one of the two  
192 schemata and performed a test on their prior-knowledge in that category. This pretest  
193 involved one picture and 6 statements about each of the 12 category exemplars (using  
194 their real not the invented names). When a picture was presented the participant could  
195 select the name of the exemplar from 6 response options or select a “?” to indicate not  
196 knowing. In response to the statements participants could select “true”, “false” or “?”. All  
197 questions were presented without a time limit. No subjects achieved more than the cut-off  
198 of 20% correct responses and therefore all could take part in the experiment.

199 The acquisition of schema-knowledge started after the pre-test in the first session  
200 and involved seven more sessions in the institute, which were separated by seven days.  
201 There were very few exceptions, when participants could not make that day, they came  
202 within around 2 days. Sessions were ~1 h long (only session 1 took 2 h), but this could  
203 vary as participants completed most tasks in a self-paced way. Between the sessions at  
204 the institute, participants deepened their schema-knowledge by working through  
205 homework. The first two sessions in the institute and the corresponding homework will be  
206 described more in detail in order to give an impression how acquisition of structured  
207 schema-knowledge was achieved.

208 In the first session participants started to learn general background information  
209 about their schema (on the category and sub-category level), that was presented self-  
210 paced on the computer. Participant's assigned to the 'arthropods'-schema learned for  
211 instance that arthropods have an exoskeleton, a segmented body, compound eyes, that  
212 the exoskeletons are based on chitin and vary with respect to their stability, that ants have  
213 two antenna, two mandibles, live in colonies etc. Participants did at the end of the first  
214 session a multiple-choice test on the acquired knowledge with 32 questions with 87.7%  
215 ( $\pm 10\%$ ) accuracy (Fig. 1 B). In the following homework the schema facts from session 1  
216 and additional facts were presented as reading material. In addition, the families and

217 exemplars were introduced including the first facts about both hierarchy levels. Moreover,  
218 each exemplar was presented on 2 pictures from different perspective. Participants were  
219 asked to work through the material and to answer open questions in writing to facilitate  
220 and deepen the learning. They were told, that their answers were evaluated in the next  
221 session in the institute.

222         In the second session in the institute, participants were asked to free recall the facts  
223 about the exemplars followed by the multiple choice and ten open questions about the  
224 schema background. They then repeated learning the background facts followed again by  
225 multiple choice and open questions. In the following homework, they repeated the general  
226 background facts as well as learning more facts about the families and exemplars by  
227 reading and responding to open question. After this homework they had learned all  
228 relevant schema-related facts the first time. From there on, this knowledge was repeated  
229 and deepened.

230         In the remaining 6 sessions in the institute and the 5 homework sessions between  
231 them participants continued in the same way to recapitulate the facts about all hierarchy  
232 levels of their schema, including 12 pictures of each exemplar and their knowledge was  
233 tested using various multiple choice and open questions as well as picture naming tests. At  
234 the end of each session participants' schema-knowledge was assessed with multiple  
235 choice tests for which they did not receive feedback. Based on their performance in these  
236 tests the homework sessions were individually adjusted.

237         Participants showed close to ceiling performance in the multiple-choice and picture  
238 naming tests in all session in the institute and participants achieved the cut-off of 85%  
239 correct responses after the final session (Fig 1 B). A final test on schema-knowledge (24  
240 questions) about 14 days after the last session in the institute and directly after the  
241 delayed retrieval of novel schema-related and control facts in the MR scanner (Fig 1B)  
242 showed that participants had successfully acquired schema-knowledge. Participants  
243 reached 91.8% ( $\pm 1.4\%$ ) high confidence correct responses and across all confidence  
244 levels 96.9% ( $\pm 0.8\%$ ) correct responses (Fig 2B). The response times for high confidence  
245 schema-knowledge retrieval was similar to high confidence immediate retrieval of novel  
246 schema-related facts and substantially faster than their delayed high confidence retrieval,  
247 which occurred in the same session (Fig. 2 D).

248 *Learning exemplar names of the control schema*

249 Differences in familiarity with the names of the schema and control exemplars could affect  
250 memory formation independent of the hierarchical, associative knowledge structure that  
251 was acquired only for the schema (Chua et al., 2012). The goal was therefore to minimize  
252 potential differences in familiarity between the names of schema and control exemplars  
253 prior to learning novel facts. Starting in session 2, participant were learning and writing  
254 down the names of the control exemplars in random order without any information about  
255 the nature of them. Participants expected to be tested on their memory for these 12 words  
256 in the next session. Participants were asked to recapitulate the names of the control  
257 exemplars in each of the following homework sessions and were tested for them in each of  
258 the sessions in the institute. As intended, this procedure resulted in high familiarity and  
259 perfect memory for the names of the exemplars of the control schema. However, this tight  
260 experimental control also probably established to some degree semanticized and  
261 arbitrarily interconnected representations of the control exemplar names, which potentially  
262 reduces the observable prior knowledge effect.

263 *Encoding and retrieval of novel schema-related and control facts*

264 Encoding and retrieval of schema-related (SR) and not schema-related (NS) control facts  
265 took place in the MR scanner. Three encoding rounds for SR and NS facts were followed  
266 by four additional encoding rounds outside of the scanner only for NS facts to reach similar  
267 performance in the following immediate memory test in the MR scanner. Similar memory in  
268 the immediate retrieval test for novel schema-related and control facts is a critical  
269 prerequisite to relate potential differences in forgetting until the delayed memory test to  
270 divergent consolidation trajectories. About two weeks later ( $15.17 \pm 1.95$  days) there was a  
271 delayed retrieval test followed by the aforementioned retrieval of schema-knowledge.

272         The encoding and immediate retrieval rounds took place on the day following the  
273 eighth session in the institute. After arrival, participants wrote down the names of the 12  
274 schema and the 12 control exemplars in order to activate their schema-knowledge.  
275 Participants were instructed to memorize each novel fact carefully, to focus only on the  
276 facts that was presented at the time, and to give each facts equal memorization effort.  
277 Participants were also informed about how they would be tested in the retrieval test. In  
278 each of the three encoding rounds all novel 72 SR and 72 NS facts (6 for each exemplar)  
279 were presented in pseudorandom order (consecutive facts were from different category  
280 members). Each encoding round was distributed across two fMRI runs. Each trial started

281 with the presentation of a fact for 4s (Fig. 1 B). Participants were asked in the first round to  
282 indicate within this time interval whether they will remember the fact (i.e. judgement of  
283 learning; van Buuren et al., 2014; Witherby and Carpenter, 2021) and in the two  
284 subsequent rounds whether they remembered the fact (judgement of memory). Therefore,  
285 the tasks differed between the first and the later encoding rounds, which we considered  
286 acceptable because implicitly there is always an unavoidable difference between the first  
287 and subsequent encounters with new information. To foreshadow the results of our across  
288 round representational similarity analyses (Figs. 5 and 6), similarity between the first two  
289 rounds was rather greater which speaks against a major influence of the different tasks on  
290 fMRI results. They had for each hand one button box, pressing the left index finger to  
291 indicate 'yes, pressing the right for 'no'. The presentation of the facts was followed by a  
292 fixation cross for 500ms, followed by an active baseline task. In this task the fixation cross  
293 changed its color from white to either red or blue and participants were instructed to press  
294 a corresponding button as quickly as possible. The baseline task was introduced to  
295 prevent rehearsal of the previously encoded novel facts and to reduce activity in the  
296 default mode network in the implicit baseline (Stark and Squire, 2001). This baseline  
297 lasted between 2 - 4 s (jittered) before the next stimulus was presented. After the 3  
298 encoding rounds, participants left the scanner and encoded the NS facts for 4 additional  
299 rounds because we aimed for equal performance in the immediate memory test for NR  
300 and NS facts. The previous study and further piloting showed that to achieve comparable  
301 performance NS items need to be presented twice as many times as SR items (Hennies et  
302 al., 2016). The procedure was the same as before.

303 Participants entered then the scanner again for the immediate retrieval test. All 72  
304 SR and 72 NS facts were tested in a two-alternative forced choice task distributed across  
305 two fMRI runs. Each trial started with the presentation of a fixation cross in the middle of  
306 the screen between 2-4 sec (jittered). This was followed by the presentation of a fact  
307 together with its alternative fact for 6 s. Participants had to choose the correct answer by  
308 pressing the corresponding button on one of the two button boxes with the right or left  
309 hand within the 6 s interval. Participants had to indicate their confidence by choosing  
310 between the options sure, unsure, guess. After each trial the same active baseline task  
311 that was employed during the encoding followed. Approximately 2 weeks later participants  
312 returned for the delayed retrieval test in the scanner, which was identical to the first test,  
313 but the order of facts was re-randomized. After retrieving all 72 SR and 72 NS facts  
314 distributed across 2 fMRI runs participants were tested in a final fMRI run on 24 of

315 schema-knowledge facts that were tested in a similar format, i.e. forced choice with  
316 confidence ratings.

317 It should be noted that the lure of a novel fact during retrieval (e.g. '4 cm' when  
318 'Styga is 2 cm long' was learned) was not systematically associated as a novel fact with  
319 another exemplar (e.g. '4 cm' was not necessarily a novel fact related to another  
320 exemplar). This lack of a systematic using each novel fact as both, lure and target might  
321 preclude an unambiguous interpretation of the results as noted by a reviewer: The novel  
322 schema-related and control facts might have been more familiar than the lures during the  
323 recognition tests and the responses could reflect familiarity driven item memory and not  
324 assimilation into the schema. However, although we did not systematically used each lure  
325 as target in a different recognition trial many lures or at least very similar responses  
326 appeared in more than one trial which reduces this potential confound (a list of all retrieval  
327 questions and the response alternatives can be found at <https://osf.io/aj28h/>).

328 As described participants encoded and retrieved ('Did you remember?') the control  
329 facts in seven rounds (compared to three rounds for schema-related facts) and  
330 immediately retrieved them then. This many rounds was necessary to reach similar  
331 performance in the immediate memory test, which was critical for the second goal of the  
332 study, i.e. to test prior knowledge effects on consolidation. However, on the other hand the  
333 repetitive encoding of object-location associations, respectively word lists has been  
334 recently shown to result itself in accelerated system consolidation that is stabilized during  
335 sleep (Brodt et al., 2016, 2018; Himmer et al., 2019). In the current study this effect might  
336 be even stronger due to the over-learned and consolidated names of the control  
337 exemplars. In other words, the tight experimental control might also have resulted in  
338 consolidated memory traces for the novel control facts – even if via a different mechanism  
339 as hypothesized for the novel schema-related facts.

#### 340 ***fMRI data acquisition***

341 Functional magnetic resonance imaging (fMRI) was performed on a 3 T system (Siemens  
342 Trio) with a 32-channel head coil. An echo planar imaging T2\*-sensitive sequence in 50  
343 contiguous axial slices (3 x 3 x 2.8 mm) with 1 mm gap; TR, 2.96 s; TE, 30 ms; PAT factor  
344 2; flip angle, 80°; matrix 64 x 64) was employed. High resolution (1 x 1 x 1 mm voxel size)  
345 T1-weighted structural MRIs were acquired for each subject using a 3D MPRAGE  
346 sequence as part of the first scanning session.

347 ***fMRI data preprocessing***

348 Functional imaging data were processed using the Statistical Parametric Mapping 12  
349 software (SPM12, Wellcome Department of Cognitive Neurology, London, UK;  
350 <http://www.fil.ion.ucl.ac.uk/spm>). Functional images were realigned and unwarped to  
351 correct for susceptibility-by-movement artifacts. The anatomical image were coregistered  
352 to the mean functional image of that participant. The anatomical images were then  
353 transformed into standard stereotaxic space using DARTEL as implemented in SPM12  
354 and the deformation field applied to the functional images of the same participant.  
355 Functional images were smoothed with full-width at half-maximum of 6 mm for the  
356 univariate and of 3 mm for the multivariate analyses.

357 ***Univariate fMRI analyses***

358 Individual subjects and group level data were analyzed using the general lineal model  
359 (GLM) as implemented in SPM 12 in a mass univariate approach. One first level model  
360 was set up for encoding and retrieval. The two runs for each of the three encoding rounds  
361 as well for the two runs of each immediate and delayed retrieval were concatenated where  
362 the run specific constant, the autocorrelation structure and high pass filter were  
363 appropriately adjusted. Regressors were created by convolving the onsets with the  
364 canonical hrf. For each of the three encoding rounds one regressor for SR and NS facts  
365 was created (due to ceiling effects we did not subdivide encoding events into those with a  
366 positive and a negative judgment of learning, respectively memory). For immediate and  
367 delayed retrieval, regressors were created for the high confidence correct, combined for  
368 medium and low confidence correct and for incorrect responses as well as for the final  
369 retrieval of the schema facts. In addition, six movement regressors were added as  
370 nuisance variables.

371 On the second level, encoding related activity was analyzed by the main effect of  
372 condition in an encoding round  $\times$  condition (SR vs NS facts) ANOVA. The analysis of  
373 retrieval activity was restricted to high confidence hits because for those a behavioral  
374 effect of schema-knowledge was observed (Fig. 2 C) which is consistent with previous  
375 literature (Long and Prat, 2002; Bein et al., 2020) and it has been argued high confident  
376 responses are most informative (e.g., Xiao et al., 2016; Lee et al., 2019). Retrieval activity  
377 was analyzed by the main effect of condition and the interaction of condition and delay in a  
378 delay (immediate vs. delayed)  $\times$  condition (SR vs NS facts) ANOVA. To identify areas that  
379 might be involved in semantic memory, retrieval activity during high confidence retrieval of

380 the schema facts (after the delayed retrieval of novel facts) was contrasted against high  
381 confidence immediate retrieval of novel schema-related facts in a paired t-test.

382 In addition, two Psycho-Physiological-Interactions (PPI; Friston et al., 1997) were  
383 conducted based on previous literature (e.g., Sommer, 2017). The first PPI used as seed  
384 the vmPFC cluster (thresholded  $p < 0.001$  uncorrected) that was identified by the analysis  
385 of encoding activity and contrasted its functional coupling during encoding SR and NS  
386 facts for each of the three encoding rounds. On the second level, coupling differences  
387 between conditions were analyzed across encoding rounds in an ANOVA with the factor  
388 round. The second PPI used as seed the vmPFC cluster identified by the retrieval second  
389 level model to be more active during retrieval of SR than control facts immediately and  
390 delayed. This PPI contrasted coupling during immediate as well as delayed retrieval of SR  
391 with control facts. On the second level, it was analyzed where this coupling increases from  
392 immediate to delayed retrieval.

### 393 **Multivariate fMRI analyses**

394 In order to get parameter estimates for each individual trial as input for the multivariate  
395 analyses, for each trial an independent first-level model was created with one regressor  
396 containing only the corresponding trial and one for all other trials in that fMRI run (Mumford  
397 et al., 2012). In addition, six movement regressors were added as nuisance variables as  
398 well a high-pass filter applied and corrected for autocorrelation. The t-maps testing the  
399 beta of the trial of interest in each model against the implicit baseline were used for the  
400 following RSA to reduce the influence of noisy voxels (Dimsdale-Zucker and Ranganath,  
401 2018). In all RSA we employed a whole brain searchlight approach (radius 3 voxels) and  
402 correlated (Pearson's linear rank correlation) the resulting vectors of trial-specific t-values  
403 across conditions of interest. The resulting correlation coefficients were averaged within  
404 condition after Fisher's Z-transformation and saved as value for the center voxel of the  
405 current searchlight.

### 406 *Encoding-encoding similarity*

407 In a first RSA we aimed to find support for a more rapid integration of novel schema-  
408 related facts into the activated associative structure in terms of a greater consistency of  
409 the distributed activity pattern across encoding rounds. Therefore we analyzed the  
410 robustness of fact-specific activity patterns (e.g. *Texana is 2 cm long* or *NIV contains*  
411 *Copper*) between the succeeding encoding rounds, i.e. between round 1 and 2 as well as  
412 between round 2 and 3 (Xue et al., 2010; Bruett et al., 2020). In particular the activity

413 pattern during encoding of a specific fact was correlated with encoding the same fact in the  
414 succeeding round (similarity) as well as with encoding of all the other facts in that round  
415 (dissimilarity), separately for novel schema-related and control facts. On the second level  
416 we contrasted the similarity-dissimilarity difference maps in a 2 (round 1/2 vs. round 2/3) x  
417 2 (SR vs NS) ANOVA.

#### 418 *Encoding-operation similarity analyses*

419 In a second RSA we aimed to find support for more effective organization processing  
420 during encoding of novel schema-related facts. Therefore, we aimed to identify brain  
421 regions in which prior knowledge affects the encoding operation irrespective of the specific  
422 to be encoded fact. To this end we correlated on the one hand activity during encoding  
423 each novel schema-related fact with all other schema-related facts in a round, but  
424 excluded the 5 other trials with facts related to the same exemplar (e.g. 'Texana') and on  
425 the other hand activity during encoding of each novel schema-related fact with all control  
426 trials resulting for each round in a SR-SR and a SR-NS correlation per subject. For the  
427 latter we excluded also all facts related to one control exemplars to avoid different  
428 numbers of correlations contributing to SR-SR and SR-NS similarities. On the second level  
429 we contrasted the resulting similarity maps in a 3 (round 1 vs. 2 vs. 3) x 2 (SR-SR vs SR-  
430 NS) ANOVA.

431 To test whether prior knowledge also results in more similar activity patterns related  
432 to the encoding operation across rounds, we repeated the above analysis but correlated  
433 activity during encoding across succeeding round, e.g. each schema-related trial in round  
434 1 with all other schema-related trials in round 2 except with the 6 trials related to the same  
435 exemplar (e.g. 'Texana'). On the second level we ran a 2 (round 1/2 vs. round 2/3) x 2  
436 (SR-SR vs SR-NS) ANOVA.

437 In addition, with the same goal, we conducted two similar, complementary analyses  
438 where we contrasted within and between rounds the correlation of activity during all novel  
439 schema-related trials (but again excluded the trials related to the same exemplar) with the  
440 correlation of activity during all control trials. This approach resulted at the second level in  
441 a 3 (round 1 vs. 2 vs. 3) x 2 (SR-SR vs. NS-NS) and a 2 (rounds 1/2 vs. rounds 2/3) x 2  
442 (SR-SR vs NS-NS) ANOVA.

#### 443 *Encoding-retrieval similarity*

444 As an alternative approach to find support for differences in consolidation due to prior  
445 knowledge we computed the similarity of activity patterns between encoding round 1 and



446 immediate as well as delayed retrieval for the novel schema-related and control facts. On  
447 the second level we run a 2 (encoding-immediate vs. –delayed retrieval) x 2 (SR-SR vs.  
448 NS-NS) ANOVA.

449 *Encoding - schema knowledge retrieval similarity*

450 To more directly test which areas are involved in the assimilation of the novel schema-  
451 related facts into the overlearned schemata, we correlated in another RSA activity during  
452 encoding novel schema-related facts with activity during retrieval of the schema-  
453 knowledge (where we excluded again trials related to the same exemplar) and contrasted  
454 it with the correlation of control facts with the retrieval of schema-knowledge. This results  
455 at the second level in a 3 (round 1 vs. 2 vs. 3) x 2 (SR-Schema vs NS-Schema) ANOVA.  
456 However, it should be noted that this analysis is to some extent confounded by the  
457 shortcoming that the schema-knowledge was retrieved only after the delayed retrieval of  
458 the novel schema-related and control facts. At this time point, the novel schema-related  
459 facts were at least partly already assimilated into the schema-knowledge which may have  
460 led to changes in the representation of that initial schema itself, i.e. accommodation.

461 *Multivariate–Univariate Dependence Analysis*

462 Pattern similarity can be caused not only by distributed patterns of activity but also by  
463 consistent (de-) activation of voxels in a univariate fashion. Therefore, we conducted a  
464 multivariate-univariate dependence analysis (MUD) previously suggested by Aly and Turk-  
465 Browne (2015) for significant RSA-peak voxels that were in brain regions which also show  
466 univariate effects. In particular, we first multiplied in each voxel of the spheres the  
467 normalized values of the corresponding trials, e.g. encoding of a specific fact in round 1  
468 and 2. These products indicate how much a voxel contributes to the correlation, i.e. to  
469 pattern similarity. These products and the mean activity of each voxel were averaged  
470 across trials and then correlated across voxels. The magnitude of the correlation indicates  
471 how much univariate effects contribute to the RSA result.

472 Importantly however, a MUD correlation does not necessarily imply that the voxels  
473 in a sphere show a univariate effect consistently in the same direction but only that voxels  
474 activate or deactivate consistently for trials in a condition and that the same voxels  
475 contribute to the observed similarity. Therefore, we computed in addition the correlation of  
476 the differences in multivariate similarity between conditions in the RSA-peak voxels with  
477 the difference in univariate mean activity in the searchlight-sphere around the RSA-peak  
478 voxels across participants (Wagner et al., 2016). The differences to the MUD are that the

479 similarity difference in the RSA-peak voxel (i.e. the difference between the correlations  
480 across all voxels in the surrounding searchlight sphere) and the difference in mean activity  
481 in all voxels of the searchlight sphere are used. This approach is diagnostic of the extent to  
482 which the observed similarity difference might be driven by differences in mean activity  
483 between conditions.

#### 484 **Correction for multiple comparisons**

485 Results of all fMRI analyses were considered significant at  $p < 0.05$ , family-wise-error  
486 (FWE) corrected for multiple comparisons across the entire scan volume or within the a  
487 priori defined anatomical regions of interest (ROIs). ROIs for the bilateral hippocampus,  
488 bilateral precuneus/posterior cingulate and angular gyrus were computed from the  
489 Harvard-Oxford cortical and subcortical structural atlases. A vmPFC ROI was manually  
490 traced on the mean T1 image based on previously published post-mortem data (Mackey  
491 and Petrides, 2014) using ITK-SNAP 3.6.0 (Yushkevich et al., 2006). The vIPFC ROI we  
492 functionally defined as a sphere with radius 10 mm centered around the previously  
493 observed peak voxel ( $xyz = [-40\ 4\ 28]$ ; Sommer, 2017).

#### 494 **Results**

##### 495 ***Behavioral results***

496 Encoding and retrieval performance was analyzed in mixed effects models using the R  
497 base package and the *lme4* as well as the *lmerTest* packages for computing type III  
498 ANOVAs with Satterthwaite's method for the approximation to degrees of freedom (Bates  
499 et al., 2015; Kuznetsova et al., 2017). Model fit, such as normality of model residuals, was  
500 verified using the *check\_distribution* function in R package *performance* (Lüdtke et al.,  
501 2020). Post-hoc Tukey HSD tests were performed using the *lemans* package (Lenth,  
502 2016) if a paired comparison was of relevance for the interpretation of the results.

##### 503 ***Encoding***

504 Confirmative (i.e. 'yes') responses in the judgement of learning (round 1) and judgment of  
505 memory (rounds 2 to 3) for schema-related and control facts were analyzed in mixed  
506 effects models with the fixed effects condition (schema vs. control) and encoding round (1  
507 vs. 2 vs. 3) as well as subject as a random factor. This model was compared with a similar  
508 one but with a subject specific slope across rounds as additional random factor. Because  
509 the model fit was similar ( $X^2(5) = 0.6$ ,  $p = .99$ ) we used the less complex model. The  
510 effects of condition and round on judgments of learning, respectively memory (Fig. 2 A)  
511 reached significance ( $F(1,135.00) = 376.35$ ,  $p < 0.001$ ;  $F(2,135) = 14.67$ ,  $p < 0.001$ ) but

512 not their interaction ( $F(2,135.00) = 0.49$ ,  $p = 0.614$ ). Participant's judgment of memory for  
513 the control facts increased outside of the scanner ( $F(3,78) = 36.48$ ,  $p < 0.001$ ) where post-  
514 hoc Tukey HSD tests show that there was no increase between round 3 and 4 ( $p = 0.714$ )  
515 but only between round 1 and 2 as well 2 and 3 ( $p < 0.001$ ;  $p = 0.007$ ) suggesting  
516 participants performance reached an asymptote. A direct comparison with round 3 of  
517 learning schema-related facts and round 7 of learning control facts showed less subjective  
518 memory ( $t(26) = 2.48$ ,  $p < 0.020$ ) for control facts.

519 The effects of condition and round on reaction times (Fig. 2 B) also reached  
520 significance ( $F(1,81) = 16.813$ ,  $p < 0.001$ ;  $F(2,27) = 144.766$ ,  $p < 0.001$ ) as did their  
521 interaction ( $F(2,81) = 10.005$ ,  $p < 0.001$ ). Restricting the analyses to confirmatory (i.e.  
522 'yes') responses resulted in a similar pattern, i.e. significant main effects ( $F(1,81) = 9.532$ ,  
523  $p = .003$ ;  $F(2,27) = 186.87$ ,  $p < .001$ ) but no interaction ( $F(2,81) = 0.542$ ,  $p = .584$ ).  
524 Reaction times did not further decrease significantly outside of the scanner during learning  
525 round 4 to 7 for control facts ( $F(3,78) = 2.12$ ,  $p = 0.104$ ). Reaction times in the last round of  
526 control facts encoding were faster than in the third and last round of schema-related facts  
527 ( $t(26) = 2.61$ ,  $p = 0.015$ ).

#### 528 *Immediate and delayed retrieval*

529 Hits during immediate and delayed retrieval were analyzed in mixed models with the  
530 fixed effects condition (schema vs. control) and confidence (3 levels) as well as subjects  
531 as a random factor with confidence as random slope (Fig. 2 C). This model had a  
532 significantly better model fit than a less complex model without the random slope  
533 (immediate retrieval  $X^2(5) = 174.1$ ,  $p < .001$ ; delayed retrieval  $X^2(5) = 147.8$ ,  $p < .001$ ).  
534 Immediate retrieval showed only a significant effect of confidence level (Fig 2 B;  $F(2,27) =$   
535  $264.33$ ,  $p < 0.001$ ), but not of condition, and no interaction ( $F(1,108) = 0.0037$ ,  $p = 0.951$ ;  
536  $F(2,108) = 0.076$ ,  $p = 0.923$ ). In other words, the additional 4 encoding rounds outside of  
537 the scanner for control facts resulted as intended in a similar immediate hit rate for  
538 schema-related and control facts. Therefore, potential difference in delayed memory  
539 performance cannot be attributed to differences in the initial memory strength. For delayed  
540 retrieval, there was again a significant effect of confidence level ( $F(2,27) = 24.05$ ,  $p <$   
541  $0.001$ ), and also a significant interaction ( $F(2,81) = 7.13$ ,  $p = 0.001$ ) which was driven by  
542 more high confidence hits for schema-related than control facts (post hoc Tukey HSD  $p =$   
543  $0.014$ ), but no effect of condition ( $F(1,81) = 0.32$ ,  $p = 0.585$ ). In a mixed model including  
544 both retrieval tests with the additional factor and random slope delay (immediate vs.

545 delayed) the interaction of condition, delay and confidence reached only a trend towards  
546 significance ( $F(2,269.99) = 2.63, p = 0.074$ ) besides the significant effects of delay,  
547 confidence and their interaction. (Loftus and Masson, 1994)

548

549 The proportion of incorrect responses across the confidence levels was also  
550 analyzed in order to test whether prior-knowledge might result in relative higher proportion  
551 of high confidence incorrect responses. Neither during immediate nor delayed retrieval  
552 was the interaction of schema and confidence significant (immediate retrieval: effect of  
553 schema  $F(1,168) < 0.01, p = .971$ ; effect of confidence  $F(2,168) = 324.26, p < .001$ ;  
554 interaction  $F(2,168) = < .01, p = .991$ ; delayed retrieval: effect of schema  $F(1,168) = 0.03$ .  
555  $p = .872$ ; effect of confidence  $F(2,168) = 31.03, p < .001$ ; interaction  $F(2,168) = 1.28, p$   
556  $= .28$ ).

557 Reaction times for hits were analyzed in similar mixed models (Fig. 2 D). For  
558 immediate retrieval both main effects as well as the interaction reached significance  
559 ( $F(1,101.34) = 55.61, p < 0.001$ ;  $F(2,38.19) = 40.30, p < 0.001$ ;  $F(2,101.04) = 31.68, p <$   
560  $0.001$ ) indicating overall slower retrieval of schema-related facts, where there was no  
561 difference for high confidence but for lower confidence hits. Also, for delayed retrieval both  
562 main effects and the interaction reached significance ( $F(1,100.96) = 8.49, p = 0.0041$ ;  
563  $F(2,47.95) = 47.95, p < 0.001$ ;  $F(2,100.90) = 3.68, p = 0.029$ ) suggesting a similar pattern,  
564 i.e. no difference for high confidence hits. In a mixed model including both retrieval tests  
565 with the additional factor and random slope delay, the effects of condition ( $F(1,225.86) =$   
566  $43.13, p < 0.001$ ), confidence ( $F(2,37.08) = 50.45, p < 0.001$ ), and delay ( $F(1,28.02) =$   
567  $59.27, p < 0.001$ ), as well as the condition  $\times$  delay ( $F(2,226.13) = 20.17, p < 0.001$ ) and the  
568 confidence  $\times$  delay interaction ( $F(2,225.33) = 4.35, p = 0.014$ ) reached significance  
569 indicating overall slower retrieval and a smaller difference between conditions after the  
570 delay.

### 571 ***Univariate fMRI results***

#### 572 *Encoding*

573 At first, we present activity related to retrieval of the overlearned schema-knowledge  
574 although it was assessed last (i.e. after delayed retrieval of novel schema-related and  
575 control facts) because its consolidation is the basis for the assimilation of novel schema-  
576 related facts. We contrasted retrieval of schema-knowledge with immediate retrieval of  
577 novel schema-related facts because both refer to schema-exemplars and were similarly

578 fast. Activity during retrieval of the over-learned schema-knowledge was greater in the  
579 vIPFC and other areas (Fig. 3 A, Table 1). Activity in the other three conditions, i.e.  
580 immediate retrieval of control facts as well as delayed retrieval of schema-related and  
581 control facts, is plotted in transparent bars because these conditions did not contribute to  
582 the statistical test that identified this area. We present it to show that activity did not differ  
583 between the immediate and delayed retrieval of novel-schema-related and control facts.  
584 This plot shows that activity during retrieval of the schema-knowledge was also greater  
585 than during immediate retrieval of control facts and delayed retrieval of novel schema-  
586 related and control facts.

587 During encoding the main effect of SR > NS across rounds was significant in the  
588 vmPFC, in a cluster comprising the ventral precuneus/retrosplenial cortex (vPC/RSC) as  
589 well as in the superior parietal cortex (Fig. 3B) implicating greater activity during encoding  
590 of novel schema-related facts. The interaction of condition and round showed that in the  
591 vmPFC and vPC/RSC the difference was greater in the first than the last round (Table 1 for  
592 full list of results). Moreover, the vmPFC was more strongly coupled with the hippocampus,  
593 fusiform, supramarginal and inferior frontal gyri, dorsal precuneus and superior parietal  
594 cortex during encoding of schema-related than control facts (Fig. 3 C). The interaction with  
595 round showed that this difference in coupling in the hippocampus and dorsal precuneus  
596 was largest in round 1.

#### 597 *Retrieval*

598 The vmPFC and the vPC/RSC were also more active during retrieval of novel schema-  
599 related than control facts irrespective of delay (Fig. 4 A). The PPI using the vmPFC cluster  
600 as seed revealed stronger coupling differences with the precuneus/posterior cingulate  
601 between retrieval of SR and NS facts after the delay (Fig. 4 B). A region of the vmPFC, in  
602 particular the subgenual ACC, showed an increase in activity only during delayed retrieval  
603 of schema-related facts whereas the hippocampus revealed a decrease in activity from  
604 early to delayed retrieval only of control facts (Fig. 4 C). There was no such interaction  
605 effect in the vIPFC ROI (largest  $Z = 1.43$ ,  $p = 0.661$  at  $[-42\ 0\ 18]$ ) contrary to our previous  
606 study (Sommer, 2017).

#### 607 **Multivariate fMRI results**

##### 608 *Encoding-encoding similarity*

609 In Fig 5 A right panel we present the results of the encoding-encoding similarity analysis  
610 irrespective of prior knowledge (i.e. the main effect similar > greater > dissimilar) in order to

611 visualize its sensitivity because contrasting schema-related and control facts using this  
612 measure revealed only relatively subtle differences. In particular, pattern robustness in  
613 terms of encoding-encoding similarity between rounds was greater for novel schema-  
614 related than control facts in the right inferior frontal gyrus (IFG, [48 12 21],  $Z = 3.90$ , Fig. 5  
615 A left panel). Importantly though, as the IFG was not an a priori defined ROI the peak  
616 would not survive correction for multiple comparisons. We decided to still report and  
617 visualize it for exploratory reasons because the IFG has been observed before to be more  
618 active and stronger coupled with the hippocampus during schema retrieval (Bein et al.,  
619 2014; van Buuren et al., 2014; Wagner et al., 2015).

620 The interaction of similarity and encoding round showed that encoding-encoding  
621 similarity between the first two rounds was greater for schema related items in early visual  
622 cortex ([0 -90 3]  $Z = 3.93$ , Fig. 5 B left panel) and the amygdala ([27 -3 -18],  $Z = 4.10$ ) and  
623 between round 2 and 3 in the precuneus ([9 -72 39]  $Z = 4.78$ , Fig 5B right panel) where  
624 only the latter peak reached significance corrected for multiple comparisons. Because the  
625 precuneus also showed univariate effects (Fig 3 A) we conducted a multivariate-univariate  
626 dependence analysis (MUD; see methods; Aly and Turk-Browne, 2015) which revealed no  
627 correlation between the observed multivariate effect and the univariate effect ( $r = -.098$ ,  $p$   
628  $= .122$ ).

#### 629 *Encoding operation similarity analysis*

630 The RSA analysis comparing correlations between encoding of schema related facts with  
631 schema related facts (SR-SR) and the correlation between encoding schema related facts  
632 and novel facts (SR-NS) within each round revealed the vmPFC ([3 51 -9],  $Z = 5.66$ ) and  
633 vPC/RSC ([-15 60 21],  $Z = 6.72$ ) as well as the midcingulate gyrus ([0 -15 48],  $Z = 5.23$ ),  
634 left middle frontal gyrus ([-21 15 42],  $Z = 4.82$ ) and the right central operculum ([48 -18 21],  
635  $Z = 4.57$ ) as a main effect across the three rounds. Notably, when multiple facts related to  
636 the same exemplar they were excluded. Because the first two areas also showed  
637 univariate effects we conducted MUD analyses for the peak voxels which revealed no  
638 correlation for the vmPFC ( $r = .028$ ,  $p = .694$ ) but did show a weak correlation for the  
639 vPV/RSC ( $r = .254$ ,  $p = .006$ ). Therefore, we correlated in addition the differences in  
640 multivariate similarity in the RSA-peak voxels and the differences in univariate mean  
641 activity in the searchlight-spheres around the RSA-peak voxels between conditions across  
642 participants which revealed no relationship ( $r = .301$ ,  $p = .115$ ;  $r = .189$ ,  $p = .337$ )

643 suggesting no major contribution of univariate activity differences between conditions to  
644 the multivariate results.

645 The interaction between the within-round similarity and round (Fig. 6 A) showed  
646 greater similarity in the vmPFC ([-3 48 -6], 6.37), vPC/RSC ([-9 -60 30],  $Z = 5.45$ ), bilateral  
647 angular gyrus ([-48 -60 24],  $Z = 4.56$ ; [51 -51 27],  $Z = 4.07$ ), OFC ([-27 36 -12],  $Z = 7.52$ )  
648 and left hippocampus ([-33 -24 -12],  $Z = 3.75$ ) in the first compared to the third round. For  
649 the first two peaks we again computed a MUD analysis to disentangle uni- and multivariate  
650 contributions. This showed subtle but significant correlations between uni- and multivariate  
651 effects (vmPFC:  $r = .113$ ,  $p = .038$ ; vPC/RSC:  $r = .249$ ,  $p < .001$ ). Therefore, we correlated  
652 again in addition the multivariate difference in the peak-voxels and the mean univariate  
653 differences in the corresponding searchlight sphere across participants which revealed  
654 only a trend towards significance for the vPC/RSC ( $r = -.225$ ,  $p = .202$ ;  $r = .323$ ,  $p = .094$ ).

655 The across round encoding operation similarity analysis (Fig. 6 B) revealed as main  
656 effect (i.e. across rounds) the vmPFC (-9 51 0],  $Z = 3.50$ ) and the vPC/RSC ([-12 -57 15],  
657  $Z = 4.87$ ; [15 -54 18],  $Z = 4.28$ ) but no interaction with round. The MUD analyses showed a  
658 subtle but significant correlation only in the vmPFC ( $r = .244$ ,  $p < .001$ ;  $r = -.056$ ,  $p = .495$ ;  $r$   
659  $= .028$ ,  $p = .686$ ) where the additional follow-up correlation of RSA-peak differences in  
660 similarity and mean activity in the corresponding searchlight-spheres suggested that the  
661 univariate contributed  $r = -.197$ ,  $p = .316$ ;  $r = .280$ ,  $p = .150$ ;  $r = .442$ ,  $p = .019$ ) only to the  
662 second peak in the vPC/RSC.

663 In addition we computed a complementary within-round encoding operation  
664 similarity analysis in which we contrasted the correlation between the novel schema-  
665 related facts (SR-SR) with the correlation of the control facts (NS-NS), excluding the trials  
666 related to the same exemplar. This also revealed greater similarity across rounds in the  
667 vmPFC ([-12 57 -3],  $Z = 5.79$ ) and precuneus/RSC ([-15 -60 21],  $Z = 4.59$ ; [-6 -48 12],  $Z =$   
668  $6.13$ ). Because both areas also showed univariate effects we conducted a multivariate-  
669 univariate dependence analyses (Aly and Turk-Browne, 2015) for the three peak voxels.  
670 The univariate contribution to the observed similarity differences correlated towards a  
671 trend in the vmPFC ( $r = .105$ ,  $p = .085$ ), and in the first peak in the vPC/RSC ( $r = .003$ ,  $p$   
672  $= .096$ ;  $r = -.037$ ,  $p = .574$ ). Thus, in the vmPFC and parts of the vPC/RSC univariate  
673 effects might have contributed modestly to the similarity.

674 The interaction of round and schema-related vs. control showed that the difference  
675 in similarity was greater in the first round in vmPFC (Fig. 6 A; [-9 57 0],  $Z = 5.98$ ), OFC ([-

676 27 36 -12],  $Z = 7.25$ ), bilateral angular gyrus (Fig 6 A; [-48 -66 21],  $Z = 4.18$ ; [51 -51 27],  $Z$   
677 = 4.13), and the vPC/RSC ([-3 -54 18],  $Z = 4.10$ ). The multi-univariate dependence  
678 analyses for these peaks revealed a marginal trend towards a correlation only for the  
679 precuneus ( $r = .092$ ,  $p = .089$ ). Therefore, pattern similarity in the precuneus appears to  
680 be, at least to some degree, related to unidirectional changes in voxel activity. Across  
681 rounds similarity of encoding operation was also greater for schema than control trials in  
682 the vmPFC (Fig. 6 B; [-9 54 0],  $Z = 4.37$ ) and vPC/RSC ([-6 -48 12],  $Z = 6.30$ ; [6 -45 15],  $Z$   
683 = 5.20). The univariate-multivariate dependence analyses revealed no significant  
684 correlation ( $r = .083$ ,  $p = .130$ ;  $r = .050$ ;  $p = .380$ ,  $r = .008$ ,  $p = .859$ ).

#### 685 *Encoding–retrieval similarity*

686 The contrast of encoding-retrieval similarities showed higher similarity during encoding and  
687 retrieval of novel schema-related facts at both delays in the middle temporal ([54 -57 3],  $Z$   
688 = 5.31) and left angular gyrus ([-45 -57 42],  $Z = 4.23$ ). The interaction revealed an increase  
689 in the parahippocampus ([27 -36 -15],  $Z = 4.08$ ).

#### 690 *Encoding - schema knowledge retrieval similarity*

691 We observed a main effect of greater similarity between encoding novel schema-related  
692 than control facts with the retrieval of schema-knowledge in the vmPFC ([-9 51 0],  $Z =$   
693 5.36, Fig. 6 C), the vPC/RSC(-3 -54 18],  $Z = 4.51$ ; [-15 -60 21],  $Z = 3.78$ ) and the posterior  
694 cingulate ([0 .30 36],  $Z = 5.34$ ). The MUD analyses for the regions where we observed  
695 univariate effects showed subtle but significant correlations for the vmPFC and one peak  
696 in the vPC/RSC ( $r = .215$ ,  $p = .001$ ;  $r = .150$ ;  $p = .008$ ;  $r = .053$   $p = .400$ ). However, the  
697 individual differences in multivariate similarity and univariate mean activity did not correlate  
698 across participants ( $r = -.103$ ,  $p = .602$ ;  $r = .065$ ,  $p = .743$ ;  $r = .046$ ,  $p = .815$ ) suggesting  
699 that the similarity difference was not caused by a consistently greater activity in the voxel  
700 of the sphere.

701 Finally, an interaction with round in terms of greater difference in round 1 than round  
702 3 was observed in the posterior cingulate ([-3 -30 36],  $Z = 4.64$ ), the left middle frontal  
703 gyrus ([-27 24 45],  $Z = 5.52$ ), right middle temporal gyrus ([57 -21 -18],  $Z = 5.09$ ), as well  
704 as trends towards significance in the bilateral angular gyrus ([-42 -60 21],  $Z = 4.55$ ; [51 -33  
705 33],  $Z = 4.46$ ).



706 **Discussion**

707 The over-learning procedure in the current study likely resulted in semanticized, cortical  
708 representations of the schema-knowledge. Learning of novel schema-related but  
709 expectation-neutral facts strongly benefited from this prior knowledge. On the neural level,  
710 we observed enhanced vmPFC-hippocampal coupling when information can be  
711 assimilated in prior-knowledge. Not only was mean activity greater in vmPFC and the  
712 vPC/RSC but also the distributed activity patterns in these areas showed greater similarity,  
713 i.e. consistency for schema-based encoding operations with the angular gyrus within and  
714 between rounds. Consolidation of the assimilated facts was also enhanced, as reflected by  
715 slightly higher confidence retrieval and an increase in vmPFC activity and vmPFC-  
716 precuneus coupling.

717 ***Acquisition of schema-knowledge***

718 The repetitive reactivation of the hierarchical schema-knowledge in various contexts (in the  
719 institute, at home, on computer screens, as written handouts) and retrieval formats (essay-  
720 like texts, free recall, multiple-choice questions, pictures) across 7 weeks was expected to  
721 result in semanticized and consolidated associative knowledge structures (Sekeres et al.,  
722 2018; Ferreira et al., 2019). The close to ceiling performance already early in learning (Fig  
723 1 B) and, even more so, in the schema-knowledge memory test 14 days after the last  
724 training session (Fig. 2 C) where response times were very fast (Fig. 2 D) show that the  
725 knowledge was highly overlearned and likely semanticized. The strong involvement of the  
726 vIPFC in retrieval of schema-knowledge further supports its proposed semanticization  
727 because this area has been implicated in semantic memory and in retrieval of  
728 semanticized memory (Binder and Desai, 2011; Sommer, 2017). Taken together, the  
729 acquired schema-knowledge very likely fulfilled the previously schema-defining criteria of  
730 being an associative network structure, based on multiple episodes, lacking unit detail,  
731 being adaptable as well as being cortical, i.e. semantic, representations (van Kesteren et  
732 al., 2012; Ghosh and Gilboa, 2014).

733 ***Encoding of novel schema-related facts***

734 The substantially higher judgments of memory for novel schema-related facts together with  
735 the more learning rounds for control facts that were necessary to reach similar  
736 performance in the immediate memory test illustrate the power of the prior-knowledge  
737 effect in the current paradigm (Fig. 2 A). Due to the invented schemata we can rule out the  
738 possibility that congruency with schema-driven expectations or differences in  
739 interest/motivation to learn the novel facts had a systematic effect on this mnemonic

740 advantage (Chua et al., 2012; Witherby and Carpenter, 2021). The possibility that the  
741 much more rapid learning of novel schema-related facts was not driven by their  
742 assimilation but rather based on better item-level memory, i.e. familiarity based  
743 recognition, cannot be ruled out. However, the only difference between retrieval of novel  
744 schema-related and control facts was observed in terms of more high confidence  
745 responses which are unlikely driven by familiarity and schema-knowledge is known to  
746 impact predominantly recollection-based recognition (Long and Prat, 2002; Brandt et al.,  
747 2005). Taken together, the observed effect is likely driven by more effective organizational  
748 processing which results in assimilation of the novel facts into the pre-activated associative  
749 network (Ericsson and Kintsch, 1995) but schema effects on item-memory cannot be ruled  
750 out.

751         On the neural level, the much more efficient encoding of novel schema-related facts  
752 was paralleled by higher coupling of the vmPFC and hippocampus in addition to higher  
753 mean activity in the vmPFC and the vPC/RSC where in the hippocampus no activity  
754 difference was observed (Fig. 3 B/C). These results conceptually replicate our previous  
755 findings using very different but also experimental schemata, i.e. associative structures of  
756 random object-location associations, that also did not result in expectations about the  
757 nature of novel facts, and also another study employing arbitrary expectation-neutral  
758 information (Liu et al., 2016; Sommer, 2017). Together, these studies therefore support a  
759 model that predicts stronger vmPFC-hippocampal coupling when novel expectation-neutral  
760 information can be related to prior-knowledge (Gilboa and Marlatte, 2017). On the first  
761 sight this interpretation stands in contrast to the prominent SLIMM model that proposes  
762 that the vmPFC inhibits hippocampal encoding when it detects information congruent to  
763 our expectations, i.e. reduced vmPFC-hippocampus coupling and less hippocampal  
764 activity when schema-congruent information is encoded (van Kesteren et al., 2012; Greve  
765 et al., 2019).

766         However, two distinct prior-knowledge effects might exist that are mediated by  
767 qualitatively different encoding operations: When schemata allow specific expectations  
768 based on previous encounters with the to be encoded information (i.e. palm-trees at a  
769 beach) there is reduced vmPFC-hippocampus coupling when congruent information is  
770 encoded (van Kesteren et al., 2012). This type of prior-knowledge effects results in  
771 enhanced gist memory prone to schema-based distortions. However, when expectation-  
772 neutral novel information can be assimilated into a schema, e.g. a novel fact in our  
773 academic discipline, that is further developed by this (i.e. accommodation) vmPFC-

774 hippocampal coupling is increased which results in enhanced and more accurate memory  
775 formation (Long and Prat, 2002; Brandt et al., 2005; Tse et al., 2007).

776         The multivariate analyses show that more efficient encoding of novel schema-  
777 related facts seems not to be predominantly reflected in less encoding variability perhaps  
778 due to the rapidity of integration (Fig. 5). The greater item-specific similarity in the IFG for  
779 novel schema-related facts – although not predicted and hence not significant after  
780 correction for multiple comparisons – would be consistent with this structure's previously  
781 described role in representing dissimilarities during knowledge integration (Schlichting et  
782 al., 2015). However, prior-knowledge substantially enhanced the consistency of the  
783 encoding operations in the vmPFC, vPC/RSC and angular gyrus suggesting more effective  
784 organizational processing (Fig. 6). In the first two areas we observed higher mean activity  
785 replicating earlier findings with expectation-neutral (Maguire et al., 1999; Tse et al., 2011;  
786 van Kesteren et al., 2014; Liu et al., 2016; Sommer, 2017) but also schema-congruent  
787 facts (van Kesteren et al., 2013; Bonasia et al., 2018). The multivariate-univariate  
788 dependence analyses suggested that the two effects might reflect different processes or  
789 distinct aspects of the same process. Activity in the vPC/RSC has been associated with  
790 retrieval, in particular with retrieval during encoding of novel information (Huijbers et al.,  
791 2012; Sestieri et al., 2017; van Kesteren et al., 2020) - consistent with activation of  
792 schema-knowledge during encoding. The greater consistency of the distributed activity  
793 patterns across all schema-trials suggest that irrespective of the current to be encoded fact  
794 (e.g. 'Texana is 2 cm long') the superordinate schema-knowledge is activated.

795         The greater similarity of activity patterns during encoding of novel schema-related  
796 than control facts with the retrieval of the overlearned schema-knowledge supports this  
797 interpretation. However, because the schema-knowledge was only retrieved after the  
798 potential assimilation of the novel information might have been already complete, this  
799 similarity might reflect both, effects of assimilation and accommodation. This interpretation  
800 would also apply to the greater consistency of encoding operations in the angular gyrus.  
801 Interestingly, the angular gyrus has been - also by using multivariate analyses of activity  
802 patterns - implicated in combining different schema components when it is applied to novel  
803 related information (Wagner et al., 2015). The vmPFC has been suggested to bind  
804 together co-activated vPC/RSC and angular gyrus representations to form a superordinate  
805 knowledge template and to maintain the active schema when novel information is  
806 processed (Gilboa and Marlatte, 2017). This interpretation of our multi- and univariate  
807 results would be consistent with the more effective organizational processing of the novel

808 related information proposed by cognitive psychologists (Ericsson and Kintsch, 1995), i.e.  
809 the integration in the associative network which leads to referential connections and the  
810 association with appropriate retrieval cues.

811         The difference between encoding of schema-related and control facts was larger  
812 early in learning in several analyses. It is possible that overlearning the names of the  
813 control exemplars resulted in an arbitrary associative structure of these meaningless  
814 terms. After participants associated facts with the elements of this structure in the first  
815 round, it became to some extent schema-like. Alternatively, after the first round the novel  
816 schema-related facts may have already been assimilated to a large degree, which would  
817 also reduce the difference to encoding of the control facts.

### 818 **Consolidation**

819 The phenomenon that prior-knowledge leads also to more efficient consolidation has been  
820 described only relatively recently (compared to the long history of schema-effects in  
821 memory) in the aforementioned animal studies using novel expectation-neutral schema-  
822 related information (Tse et al., 2007, 2011). We confirmed this effects in humans in our  
823 previous study where we translated the animal experiment to an fMRI design (Sommer,  
824 2017). In the current study, using very different experimental schemata, we conceptually  
825 replicated this effect. The impact on consolidation was rather subtle and was specific to  
826 high confidence responses which is however consistent with previous reports on prior-  
827 knowledge effects in recognition memory (Long and Prat, 2002; Brandt et al., 2005). The  
828 subtly of the prior-knowledge effect on consolidation might be also a side effect of the  
829 seven learning rounds needed for the control information to reach similar immediate  
830 memory performance because repetitive encoding and retrieval might result itself in faster  
831 systems consolidation – although via different mechanisms (Brodts et al., 2016, 2018;  
832 Himmer et al., 2019). In the current study this effect might be even stronger due to the  
833 over-learned and consolidated names of the control exemplars. Using the same paradigm,  
834 we previously showed that sleep spindle density as a proxy for nightly replay and systems  
835 consolidation predicted the individual benefit of prior-knowledge on novel related facts  
836 (Hennies et al., 2016). This finding supports the interpretation that the reduced forgetting  
837 we observed in the current study is caused by more efficient systems consolidation.

838         Prior-knowledge resulted in an increase in vmPFC-precuneus coupling and of  
839 vmPFC and hippocampus activity from immediate to retrieval of schema-related facts two  
840 weeks later. In our previous study we did not observe the vmPFC but vIPFC, i.e. a  
841 semantic memory area, to contribute more strongly to the retrieval of assimilated

842 information after two weeks (Sommer, 2017). This difference might be caused by  
843 differences in the experimental designs, for instance that the schemata in our previous  
844 study were much simpler, no meaningful hierarchical associative structures and probably  
845 semanticized to a larger degree by even more intense over-learning. However, the vmPFC  
846 (and RSC) has been implicated before in the retrieval of assimilated schema-related facts  
847 (Tse et al., 2011) and the parallel relative increase in hippocampal activity would be  
848 consistent with the incidental retrieval of associated schema knowledge during high  
849 confidence recognition (Schultz et al., 2022). This parallel involvement of cortical, i.e.  
850 vmPFC and RSC, and hippocampal retrieval would be consistent with the Trace  
851 Transformation theory that proposes that both episodic and cortical traces can exist in  
852 parallel (Sekeres et al., 2018).

### 853 **Conclusions**

854 The increased vmPFC-hippocampal coupling during the highly efficient encoding – likely  
855 due to their assimilation - of novel schema-related expectation-neutral facts suggests a  
856 prior-knowledge effect which is distinct from situations where the prior-knowledge allows  
857 expectations. Together, our uni- and multivariate results support cognitive and  
858 neuroscientific models about the processes underlying the putative assimilation, i.e. that a  
859 vmPFC, vPC/RSC, angular-network results in the activation of schema knowledge  
860 enabling more effective organizational processing of novel related facts. Moreover, the  
861 results confirm that assimilation of novel related information also results in more effective  
862 consolidation which is reflected, for not fully semanticized information, in vmPFC activity.  
863

864 **References**

- 865 Alba JW, Hasher L (1983) Is Memory Schematic? *Psychol Bull* 93:203–231.
- 866 Aly M, Turk-Browne NB (2015) Attention Stabilizes Representations in the Human Hippocampus.  
867 *Cereb Cortex*:bhv041.
- 868 Bartlett FC (1932) *Remembering: An experimental and social study*. Cambridge, UK: Cambridge  
869 University.
- 870 Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4.  
871 *J Stat Softw* 67:1–48.
- 872 Bein O, Reggev N, Maril A (2014) Prior knowledge influences on hippocampus and medial  
873 prefrontal cortex interactions in subsequent memory. *Neuropsychologia* 64:320–330.
- 874 Bein O, Reggev N, Maril A (2020) Prior knowledge promotes hippocampal separation but cortical  
875 assimilation in the left inferior frontal gyrus. *Nat Commun* 11:4590.
- 876 Binder JR, Desai RH (2011) The neurobiology of semantic memory. *Trends Cogn Sci* 15:527–536.
- 877 Bonasia K, Sekeres MJ, Gilboa A, Grady CL, Winocur G, Moscovitch M (2018) Prior knowledge  
878 modulates the neural substrates of encoding and retrieving naturalistic events at short and  
879 long delays. *Neurobiol Learn Mem* 153:26–39.
- 880 Brandt KR, Cooper LM, Dewhurst SA (2005) Expertise and recollective experience: recognition  
881 memory for familiar and unfamiliar academic subjects. *Appl Cogn Psychol* 19:1113–1125.
- 882 Brod G, Lindenberger U, Wagner AD, Shing YL (2016) Knowledge Acquisition during Exam  
883 Preparation Improves Memory and Modulates Memory Formation. *J Neurosci* 36:8103–  
884 8111.
- 885 Brodt S, Gais S, Beck J, Erb M, Scheffler K, Schönauer M (2018) Fast track to the neocortex: A  
886 memory engram in the posterior parietal cortex. *Science* 362:1045–1048.
- 887 Brodt S, Pöhlchen D, Flanagan VL, Glasauer S, Gais S, Schönauer M (2016) Rapid and independent  
888 memory formation in the parietal cortex. *Proc Natl Acad Sci* 113:13251–13256.
- 889 Bruett H, Calloway RC, Tokowicz N, Coutanche MN (2020) Neural pattern similarity across  
890 concept exemplars predicts memory after a long delay. *NeuroImage* 219:117030.
- 891 Carpenter SK, Haynes CL, Corral D, Yeung KL (2018) Hypercorrection of high-confidence errors  
892 in the classroom. *Memory* 26:1379–1384.
- 893 Chua EF, Hannula DE, Ranganath C (2012) Distinguishing highly confident accurate and inaccurate  
894 memory: Insights about relevant and irrelevant influences on memory confidence. *Memory*  
895 20:48–62.
- 896 Dimsdale-Zucker HR, Ranganath C (2018) Representational Similarity Analyses. In: *Handbook of*  
897 *Behavioral Neuroscience*, pp 509–525. Elsevier. Available at:  
898 <https://linkinghub.elsevier.com/retrieve/pii/B9780128120286000276> [Accessed September  
899 8, 2021].

- 900 Ericsson KA, Kintsch W (1995) Long-term working memory. *Psychol Rev* 102:211–245.
- 901 Ferreira CS, Charest I, Wimber M (2019) Retrieval aids the creation of a generalised memory trace  
902 and strengthens episode-unique information. *NeuroImage* 201:115996.
- 903 Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and  
904 modulatory interactions in neuroimaging. *NeuroImage* 6:218–229.
- 905 Ghosh VE, Gilboa A (2014) What is a memory schema? A historical perspective on current  
906 neuroscience literature. *Neuropsychologia* 53:104–114.
- 907 Gilboa A, Marlatte H (2017) Neurobiology of Schemas and Schema-Mediated Memory. *Trends*  
908 *Cogn Sci* 21:618–631.
- 909 Greve A, Cooper E, Tibon R, Henson RN (2019) Knowledge is power: Prior knowledge aids  
910 memory for both congruent and incongruent events, but in different ways. *J Exp Psychol*  
911 *Gen* 148:325–341.
- 912 Hennies N, Lambon Ralph MA, Kempkes M, Cousins JN, Lewis PA (2016) Sleep Spindle Density  
913 Predicts the Effect of Prior Knowledge on Memory Consolidation. *J Neurosci* 36:3799–  
914 3810.
- 915 Himmer L, Schönauer M, Heib DPJ, Schabus M, Gais S (2019) Rehearsal initiates systems memory  
916 consolidation, sleep makes it last. *Sci Adv* 5:eaav1695.
- 917 Huijbers W, Vannini P, Sperling RA, C.M. P, Cabeza R, Daselaar SM (2012) Explaining the  
918 encoding/retrieval flip: Memory-related deactivations and activations in the posteromedial  
919 cortex. *Neuropsychologia* 50:3764–3774.
- 920 Kuznetsova A, Brockhoff PB, Christensen RHB (2017) **lmerTest** Package: Tests in Linear Mixed  
921 Effects Models. *J Stat Softw* 82 Available at: <http://www.jstatsoft.org/v82/i13/> [Accessed  
922 March 2, 2021].
- 923 Lee H, Samide R, Richter FR, Kuhl BA (2019) Decomposing Parietal Memory Reactivation to  
924 Predict Consequences of Remembering. *Cereb Cortex* 29:3305–3318.
- 925 Lenth RV (2016) Least-Squares Means: The R Package **lsmeans**. *J Stat Softw* 69 Available at:  
926 <http://www.jstatsoft.org/v69/i01/> [Accessed March 2, 2021].
- 927 Liu Z-X, Grady C, Moscovitch M (2016) Effects of Prior-Knowledge on Brain Activation and  
928 Connectivity During Associative Memory Encoding. *Cereb Cortex*:bhw047.
- 929 Loftus GR, Masson MEJ (1994) Using confidence intervals in within-subject designs. *Psychon Bull*  
930 *Rev* 1:476–490.
- 931 Long DL, Prat CS (2002) Memory for Star Trek: the role of prior knowledge in recognition  
932 revisited. *J Exp Psychol Learn Mem Cogn* 28:1073–1082.
- 933 Lüdecke D, Makowski D, Indrajeet Patil, Waggoner P (2020) easystats/performance: performance  
934 0.4.7. Zenodo. Available at: <https://zenodo.org/record/3952174> [Accessed March 2, 2021].
- 935 Mackey S, Petrides M (2014) Architecture and morphology of the human ventromedial prefrontal  
936 cortex. *Eur J Neurosci* 40:2777–2796.

- 937 Maguire EA, Frith CD, Morris RG (1999) The functional neuroanatomy of comprehension and  
938 memory: the importance of prior knowledge. *Brain* 122 (Pt 10):1839–1850.
- 939 Mumford JA, Turner BO, Ashby FG, Poldrack RA (2012) Deconvolving BOLD activation in event-  
940 related designs for multivoxel pattern classification analyses. *NeuroImage* 59:2636–2643.
- 941 Rojahn K, Pettigrew TF (1992) Memory for schema-relevant information: A meta-analytic  
942 resolution. *Br J Soc Psychol* 31:81–109.
- 943 Schlichting ML, Mumford JA, Preston AR (2015) Learning-related representational changes reveal  
944 dissociable integration and separation signatures in the hippocampus and prefrontal cortex.  
945 *Nat Commun* 6:8151.
- 946 Schultz H, Sommer T, Peters J (2022) Category-sensitive incidental reinstatement in medial  
947 temporal lobe subregions during word recognition. *Learn Mem* 29:126–135.
- 948 Sekeres MJ, Winocur G, Moscovitch M (2018) The hippocampus and related neocortical structures  
949 in memory transformation. *Neurosci Lett* 680:39–53.
- 950 Sestieri C, Shulman GL, Corbetta M (2017) The contribution of the human posterior parietal cortex  
951 to episodic memory. *Nat Rev Neurosci* 18:183–192.
- 952 Sommer T (2017) The Emergence of Knowledge and How it Supports the Memory for Novel  
953 Related Information. *Cereb Cortex N Y N 1991* 27:1906–1921.
- 954 Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions  
955 in fMRI. *Proc Natl Acad Sci U A* 98:12760–12766.
- 956 Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, Morris RG (2007)  
957 Schemas and memory consolidation. *Science* 316:76–82.
- 958 Tse D, Takeuchi T, Kakeyama M, Kajii Y, Okuno H, Tohyama C, Bito H, Morris RG (2011)  
959 Schema-dependent gene activation and memory encoding in neocortex. *Science* 333:891–  
960 895.
- 961 van Buuren M, Kroes MCW, Wagner IC, Genzel L, Morris RGM, Fernandez G (2014) Initial  
962 Investigation of the Effects of an Experimentally Learned Schema on Spatial Associative  
963 Memory in Humans. *J Neurosci* 34:16662–16670.
- 964 van Kesteren MT, Ruiters DJ, Fernandez G, Henson RN (2012) How schema and novelty augment  
965 memory formation. *Trends Neurosci* 35:211–219.
- 966 van Kesteren MTR, Beul SF, Takashima A, Henson RN, Ruiters DJ, Fernández G (2013) Differential  
967 roles for medial prefrontal and medial temporal cortices in schema-dependent encoding:  
968 From congruent to incongruent. *Neuropsychologia* 51:2352–2359.
- 969 van Kesteren MTR, Rignanesi P, Gianferrara PG, Krabbendam L, Meeter M (2020) Congruency  
970 and reactivation aid memory integration through reinstatement of prior knowledge. *Sci Rep*  
971 10:4776.
- 972 van Kesteren MTR, Rijpkema M, Ruiters DJ, Morris RGM, Fernández G (2014) Building on prior  
973 knowledge: schema-dependent encoding processes relate to academic performance. *J Cogn*  
974 *Neurosci* 26:2250–2261.



- 975 Wagner IC, van Buuren M, Bovy L, Fernández G (2016) Parallel Engagement of Regions  
976 Associated with Encoding and Later Retrieval Forms Durable Memories. *J Neurosci*  
977 36:7985–7995.
- 978 Wagner IC, van Buuren M, Kroes MC, Gutteling TP, van der Linden M, Morris RG, Fernández G  
979 (2015) Schematic memory components converge within angular gyrus during retrieval.  
980 *eLife* 4:e09668.
- 981 Witherby AE, Carpenter SK (2021) The rich-get-richer effect: Prior knowledge predicts new  
982 learning of domain-relevant information. *J Exp Psychol Learn Mem Cogn* Available at:  
983 <http://doi.apa.org/getdoi.cfm?doi=10.1037/xlm0000996> [Accessed March 19, 2021].
- 984 Xiao X, Dong Q, Chen C, Xue G (2016) Neural pattern similarity underlies the mnemonic  
985 advantages for living words. *Cortex* 79:99–111.
- 986 Xue G, Dong Q, Chen C, Lu Z, Mumford JA, Poldrack RA (2010) Greater neural pattern similarity  
987 across repetitions is associated with better memory. *Science* 330:97–101.
- 988 Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G (2006) User-guided 3D  
989 active contour segmentation of anatomical structures: significantly improved efficiency and  
990 reliability. *NeuroImage* 31:1116–1128.
- 991  
992

993 Table 1: univariate fMRI results

contrast	area	xyz coordinate peak voxel	Z-value peak voxel
<b>encoding</b>			
SR > NS main effect	vmPFC	0 54 -15	4.34
	ventral precuneus/RSC	-9 -57 15	6.36
		12 -57 18	4.67
	superior parietal cortex	-36 -78 39	5.34
		48 -72 33	4.79*
	post cingulate cortex	-3 -36 33	4.81
	frontal pole	-30 51 21	5.00
NS > SR main effect		36 54 18	5.32
	anterior cingulate cortex	6 36 18	5.69
	occipital pole	9 -81 -3	6.94
	lateral occipital cortex	-45 -78 0	4.85
		33 -90 -9	4.78
SR > NS increase across rounds	supramarginal gyrus	57 -39 24	4.95
	inferior frontal gyrus/insula	-51 -6 0	4.82
		54 12 -3	4.88
	anterior cingulate cortex	-3 33 15	4.93
SR > NS decrease across rounds	vmPFC	-3 51 -32	4.70
	ventral precuneus/RSC	-12 -57 15	4.09
	dorsal medial PFC	-3 36 39	6.07
	inferior frontal gyrus	-45 39 -12	5.92
	dorsolateral PFC	-42 15 42	4.88
PPI SR > NS seed vmPFC main effect	hippocampus	-21 -24 -15	3.99
	fusiform gyrus	-30 -57 -15	5.33
		33 -54 -12	6.00
	inferior frontal gyrus	48 36 3	5.63
	superior parietal cortex	15 -66 51	5.04
		33 -42 42	4.91
	dorsal precuneus	-9 -39 57	5.03
		-9 -69 51	4.89
PPI SR > NS seed vmPFC decrease across rounds	supramarginal gyrus	-57 -45 27	4.98
	hippocampus	-27 -21 -21	3.63
		15 -12 -21	5.09
	dorsal precuneus	9 -63 51	6.69
<b>retrieval</b>			
SR > NS	vmPFC	3 30 -21	4.57
	ventral precuneus/RSC	-9 -57 15	6.49
		9 -54 15	5.68
NS > SR	anterior cingulate cortex	-6 33 21	5.06
	insula	-39 15 -9	4.97
		33 18 -15	4.46*
	frontal pole	24 57 -3	4.66
increase SR > NS	vmPFC	-6 18 -9	3.76

	hippocampus	27 -24 -15	3.74
PPI increase SR > NR seed vmPFC	ventral precuneus	15 -48 33	4.23
<b>schema-knowledge</b>			
	ventro-lateral PFC	-42 9 30	3.99
	anterior cingulate cortex	9 18 39	5.31
schema retrieval > immediate novel SR retrieval	ventral precuneus	21 -66 3	5.04
	insula	30 24 -9	5.09
		-30 24 -6	4.96
	ventral striatum	-18 9 -6	4.84
		18 25 -6	5.58

994

995 Peal coordinates in MNI space. Correction for multiple comparisons was done on the  
 996 whole-brain level or within pre-defined anatomical regions of interest, specifically the  
 997 hippocampus, precuneus/retrosplenial cortex (RSC) and ventromedial prefrontal cortex  
 998 (vmPFC). \* trend towards significance  
 999

1000 **Figure 1.** Schema and timeline of the experiment. A) Hierarchical structure of one of the two schemata  
1001 (arthropods) and exemplar names of the other schema (cells) which served as control in this example.  
1002 Schema and control were randomized across participants. See Hennies et al., 2016 for a figure with the  
1003 hierarchical structure of the 'cell'-schema. B) Acquisition of schema knowledge and familiarization with the  
1004 control names over 7 weeks with 1 learning session per week in the institute and homework in-between.  
1005 Participants achieved high performance in the multiple-choice questions (mc questions) and the picture  
1006 naming task of their schema. In the scanner, participants encoded 3 times (encoding round 1 to 3) 72 novel  
1007 facts related to the exemplars of their schema and 72 facts related to their control exemplars. In encoding  
1008 round 4 to 7 outside of the scanner, they only repeated the control facts to ensure equal immediate memory  
1009 for schema-related and control facts In the first encoding round, participants judged whether they will  
1010 remember the novel facts in rounds 2 to 7 whether they did remember it. Encoding was followed by  
1011 immediate retrieval of all learned novel facts in the scanner. 2 weeks later all facts were retrieved again in  
1012 the scanner followed by retrieval of 24 of the overlearned schema-knowledge facts During retrieval two  
1013 equally plausible response alternatives were presented (targets and lures were randomized across  
1014 participants) and participants indicated their confidence on a 3-point scale (hc – high, mc – medium, lc - low  
1015 confidence).  
1016

1017 **Figure 2.** Behavioral results. A) Encoding rounds 1 to 3 for novel schema-related and control facts took  
1018 place in the MR scanner, rounds 4 to 7 only for control facts outside of the scanner. In the first round partici-  
1019 pants rated whether they will remember the fact (judgment of learning) and in round 2 to 7 whether they did  
1020 remember the fact (judgement of memory). B) Response times for the judgment of learning (round 1), re-  
1021 spectively judgment of memory (round 2 to 7) during encoding. C) Proportion of high (hc), medium (mc) and  
1022 low confident (lc) hits (relative to all responses in that delay x schema condition) during immediate and de-  
1023 layed retrieval for the schema-related and control facts as well as for the a subset of facts of the schema  
1024 knowledge only during delayed retrieval. D) Retrieval times for high, medium and low confident hits during  
1025 retrieval. Error bars are standard errors of the mean around the mean, corrected for interindividual differ-  
1026 ences (Loftus and Mason 1994).  
1027

1028 **Figure 3.** Activity related to retrieval of the over-learned schema knowledge and to the encoding of novel  
1029 schema-related and control facts. A) Activity during retrieval of the overlearned schema-knowledge (after the  
1030 delayed retrieval of the novel schema-related and control facts) was greater in the ventrolateral PFC (and  
1031 other areas) compared with immediate retrieval of novel schema-related facts (red bar). Activity during  
1032 retrieval of schema-knowledge was statistically contrasted against immediate retrieval of schema-related  
1033 facts because response times were similar in both conditions (Fig. 2 D). Activity in the other three conditions  
1034 (immediate retrieval of control facts as well as delayed retrieval of schema-related and control facts) in this  
1035 voxel is plotted in transparent bars because it was not statistically tested against retrieval of schema-  
1036 knowledge. B) During encoding of novel schema-related (SR, red bars) than control (NS, blue bars) facts  
1037 activity was greater in the vmPFC and the vPC/RSC. C) Coupling differences between encoding schema-  
1038 related and control facts. The vmPFC was more strongly coupled with the hippocampus and fusiform gyrus  
1039 during encoding of schema-related (SR) than control (NS) facts .in the three rounds. Error bars are standard  
1040 errors of the mean around the mean, corrected for interindividual differences (Loftus and Mason 1994).  
1041 Visualization threshold  $p < 0.001$ , uncorrected.  
1042

1043 **Figure 4.** Activity differences during retrieval. A) During immediate and delayed retrieval of schema-related  
1044 (SR, red bars) facts activity was greater in the vmPFC and the vPC/RSC. B) The Difference in coupling of  
1045 the vmPFC with the precuneus during retrieval of schema-related (SR) and control (NS) facts increased from  
1046 immediate to delayed retrieval. C) The vmPFC (subgenual ACC) and the hippocampus showed a larger  
1047 activity increase from immediate to delayed retrieval of schema-related (SR, red bars) than control (NS, blue  
1048 bars) facts. Error bars are standard errors of the mean around the mean, corrected for interindividual  
1049 differences (Loftus and Mason 1994). Visualization threshold  $p < 0.001$ , uncorrected.  
1050

1051 **Figure 5.** Encoding-encoding similarity (pattern robustness). A) Encoding-encoding similarity between  
1052 succeeding rounds was greater for novel schema-related (SR, red bars) than control facts (NS, blue bars) in  
1053 the right IFG (left panel). The overall sensitivity of this approach is visualized in the right panel of A in terms  
1054 of the main effect. B) Encoding-encoding similarity between the first two rounds was greater in early visual  
1055 cortex (left panel) and between round 2 and 3 in the precuneus (left panel) for novel schema-related (SR)  
1056 than control facts (NS). Note that the IFG and cuneus cluster are not significant corrected for multiple  
1057 comparisons and are reported for exploratory reasons. Error bars are standard errors of the mean around  
1058 the mean, corrected for interindividual differences (Loftus and Mason 1994). Visualization threshold  $p <$   
1059 0.001, uncorrected.  
1060

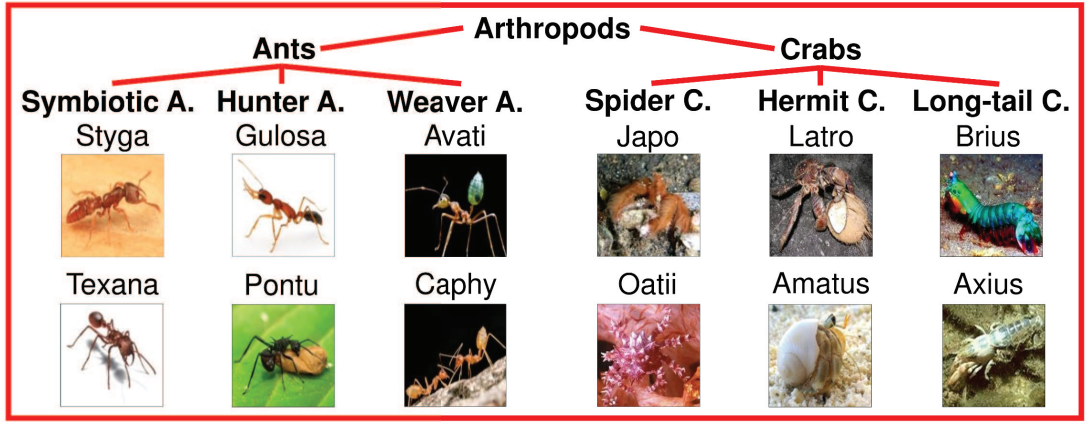


1061 **Figure 6.** Encoding operation similarity. A) Encoding-operation similarity within rounds was early in learning  
1062 greater between encoding of novel schema-related (SR-SR, red bars) than between schema-related and  
1063 control (SR-NS, blue bars) facts in the vPC/RSC, vmPFC, and the bilateral angular gyrus. B) Encoding-  
1064 operation similarity between rounds during encoding of novel-related (SR-SR) facts was also greater  
1065 compared to SR-NS facts in the vPC/RSC and vmPFC. C) Operation similarity between encoding of  
1066 schema-related facts and retrieval of schema knowledge was also greater compared to the encoding of  
1067 control facts in the vPC/RSC and vmPFC. Error bars are standard errors of the mean around the mean,  
1068 corrected for interindividual differences (Loftus and Mason 1994). Visualization threshold  $p < 0.001$ ,  
1069 uncorrected.

A

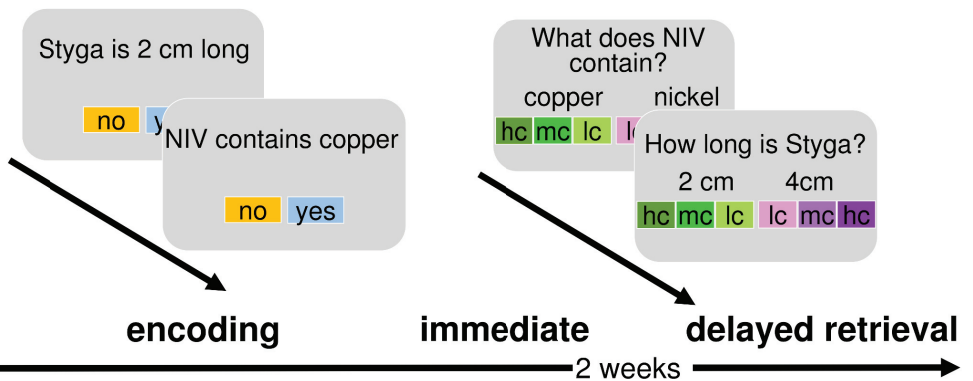
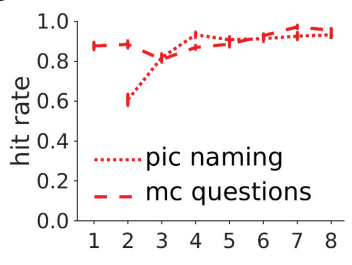
hierarchical schema-knowledge

control names



Kantaso	Maksa
NIV	Syti
Podo	Lih
Tuma	Mitoko
Erribo	SER
Rakkula	Lys22

B



acquisition of hierarchical schema-knowledge names of control exemplars

fMRI: 3 encoding rounds novel schema-related & control facts

outside of MR-scanner: 4 encoding rounds control facts

fMRI: immediate & delayed retrieval novel schema-related & control facts

fMRI: delayed retrieval schema-knowledge

