

1

## 2 Main Manuscript for

### 3 Unlocking Adults' Implicit Statistical Learning by Cognitive Depletion

4 Eleonore H. M. Smalle<sup>1</sup>, Tatsuya Daikoku<sup>2</sup>, Arnaud Szmalec<sup>1,3,4</sup>, Wouter Duyck<sup>1</sup>, and Riikka  
5 Möttönen<sup>5,6</sup>

6 <sup>1</sup>Department of Experimental Psychology, Ghent University, Belgium; <sup>2</sup>International Research  
7 Center for Neurointelligence (WPI-IRCN), The University of Tokyo, Japan; <sup>3</sup>Psychological  
8 Research Institute, Université catholique de Louvain, Belgium; <sup>4</sup>Institute of Neuroscience,  
9 Université catholique de Louvain, Belgium; <sup>5</sup>School of Psychology, University of Nottingham,  
10 United Kingdom; <sup>6</sup>Cognitive Science, Department of Digital Humanities, University of Helsinki,  
11 Finland


12 \*Eleonore H.M. Smalle is the corresponding author

13 **Email:** [eleonore.smalle@ugent.be](mailto:eleonore.smalle@ugent.be)

14 Eleonore H.M. Smalle  <https://orcid.org/0000-0003-2363-5693>

15 Tatsuya Daikoku  <https://orcid.org/0000-0002-6464-2964>

16 Arnaud Szmalec  <https://orcid.org/0000-0003-3903-3953>

17 Wouter Duyck  <https://orcid.org/0000-0003-2114-6212>

18 Riikka Möttönen  <https://orcid.org/0000-0003-4533-4277>

19

20 **Author Contributions:** *ES:* Conceptualization, Methodology, Software, Validation, Formal  
21 Analysis, Investigation, Resources, Data curation, Writing – Original draft, Review & Editing,  
22 Visualization, Project administration, Funding acquisition. *TD:* Software, Formal Analysis,  
23 Visualization, Funding acquisition. *AS:* Writing – Review & Editing. *WD:* Writing – Review &  
24 Editing, Supervision. *RM:* Conceptualization, Methodology, Software, Validation, Resources,  
25 Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

26 **Competing Interest Statement:** We have no conflict of interest to disclose.

27 **Keywords:** auditory statistical learning; implicit learning; electroencephalography;  
28 cognitive load; transcranial magnetic stimulation

29 **This PDF file includes:** Main text, Figures 1 to 3, Table 1, Supplementary Information

30

31 **Abstract**

32 Human learning is supported by multiple neural mechanisms that mature at different rates  
33 and interact in mostly cooperative, but also sometimes competitive ways. We tested the  
34 hypothesis that the mature cognitive mechanisms constrain implicit statistical learning  
35 mechanisms that contribute to early language acquisition. Specifically, we tested the  
36 prediction that depleting cognitive-control mechanisms in adults enhances their implicit,  
37 auditory word-segmentation abilities. Young adults were exposed to continuous streams of  
38 syllables that repeated into hidden novel words, while watching a silent film. Afterwards,  
39 learning was measured in a forced-choice test that contrasted hidden words versus non-  
40 words. The participants also had to indicate whether they explicitly recalled the word or not, in  
41 order to dissociate explicit versus implicit knowledge. We additionally measured  
42 electroencephalography during exposure to measure neural entrainment to the repeating  
43 words. Engagement of the cognitive mechanisms was manipulated by using two methods. In  
44 Experiment 1 (n = 36), inhibitory continuous theta-burst stimulation (TBS) was applied to the  
45 left dorsolateral prefrontal cortex or a control region. In Experiment 2 (n = 60), participants  
46 performed a dual working-memory task that induced high or low levels of cognitive fatigue. In  
47 both experiments, cognitive depletion enhanced word recognition - especially when  
48 participants reported low confidence in remembering the words, i.e., when their knowledge  
49 was implicit. TBS additionally modulated neural entrainment to the words and syllables. These  
50 findings suggest that cognitive depletion improves the acquisition of linguistic knowledge in adults  
51 by unlocking implicit statistical learning mechanisms and support the hypothesis that adult  
52 language learning is antagonized by higher cognitive mechanisms.

53 **Significance Statement**

54 Statistical learning mechanisms enable extraction of patterns in the environment from infancy to  
55 adulthood. For example, they enable segmentation of continuous speech streams into novel  
56 words. Adults typically become aware of the hidden words even when passively listening to  
57 speech streams. It remains poorly understood how cognitive development and brain maturation  
58 affect implicit statistical learning, i.e., infant-like learning without awareness. Here we show that  
59 the depletion of the cognitive control system by non-invasive brain stimulation or by demanding  
60 cognitive tasks, boosts adults' implicit, but not explicit, word-segmentation abilities. These  
61 findings suggest that the adult cognitive architecture constrains statistical learning mechanisms  
62 that are likely to contribute to early language acquisition, and open new avenues to enhance  
63 language learning abilities in adults.

64  
65 **Introduction**

66  
67 Human learning is thought to be supported by the interactions between two basic memory  
68 systems of the brain, namely declarative and non-declarative memory (1). Declarative memory is  
69 characterized by voluntary, explicit, attention-based processes, such as recall and recognition of  
70 facts/events, and is mediated by medial-temporal lobe and prefrontal cortex structures (2). Non-  
71 declarative memory, also referred to as procedural memory, on the other hand is part of implicit  
72 memory and includes the acquisition of a heterogeneity of skills, habits, and procedures. It is  
73 mediated by basal ganglia, cerebellar and neocortical structures, as well as parts of the prefrontal  
74 cortex (e.g., Broca's area, 3-5).

75 Accumulating evidence supports a competitive relationship between these two memory systems  
76 during human skill learning. Suppression of the declarative memory system by interventions like  
77 repetitive transcranial magnetic stimulation (TMS), distraction tasks, alcohol consumption,  
78 hypnosis, intake of benzodiazepines or cognitive fatigue, can actually enhance performance in  
79 implicit, perceptual-motor learning tasks such as the Serial-Reaction Time Task (6-11), or intuitive  
80 reasoning tasks (12). These findings suggest that higher-level cognitive functions associated with

81 declarative memory and supported by the prefrontal cortex can interfere with behaviour that is  
82 naturally driven by implicit learning processes (13). However, it remains unresolved whether  
83 competing memory systems also affect implicit statistical learning abilities that are critical for the  
84 early, rapid acquisition of language in infants (14). This is an important question as it could  
85 explain why infants and children pick up languages with less effort than adults (cf. 'Why are there  
86 critical periods for language learning?') (15).

87 Language acquisition involves many different memory and learning processes that are dependent  
88 on both procedural and declarative memory (2, 16). The first step for infants acquiring language is  
89 to gain knowledge about the phonological structure in one's spoken language system, the  
90 probabilistic constraints on how speech sounds combine (i.e., phonotactic learning), and the  
91 segments of continuous speech (i.e., word-forms) (17). Word-form learning takes place already in  
92 the first 12 months of life and is an important precursor to vocabulary acquisition (i.e., mapping  
93 form to meaning) and more complex language acquisition (e.g., grammar) later in development  
94 (18). In the present study, we focus on statistical learning mechanisms that contribute to word-  
95 segmentation and thus novel word-form learning in the early stages of language acquisition.

96  
97 Statistical learning is generally known as the ability to pick up on patterns in the environment  
98 through extraction of frequent regularities and distributional properties. The term was first  
99 introduced in the field of cognitive psychology by the work of Saffran, Aslin, and Newport (1996)  
100 (19), who demonstrated that infants of only eight months old can extract word boundaries and  
101 segment novel word-forms from a continuous stream of speech sounds with no other cue than  
102 the transitional probabilities between syllables. Later, this learning was also demonstrated in older  
103 children and adults (20, 21), and across different domains (e.g., music, grammar) or modalities  
104 (e.g., auditory, visual, motor) (22, 23), indicating that statistical learning is a largely continuous  
105 and domain-general learning mechanism for skill acquisition across the human life span.

106  
107 In a typical statistical learning experiment, participants are repeatedly exposed to patterned  
108 stimuli such as consonant strings from an artificial grammar, or recurrent syllable triplets.  
109 Learning is then typically assessed *post-exposure* by using a two-alternative forced-choice  
110 recognition task in which triplets from the exposure stream are pitted against foils. Participants  
111 have to indicate which of the two triplets sounded most familiar, and above-chance accuracy is  
112 taken as indication of learning. Since statistical learning occurs without any instruction or intention  
113 to learn, it is often assumed to result in implicit memory representations (24). This view is also  
114 supported by the evidence that statistical learning occurs in infants and even in sleeping  
115 neonates (25). However, in recent work, Batterink and colleagues demonstrated that even without  
116 intention to learn, adults acquire mainly *explicit* knowledge of the novel word-forms during  
117 statistical learning (26-29). This can be derived from the observation that participants'  
118 performance was above chance when they were confident remembering the triplet, but at chance  
119 when they were unconfident. Knowledge is *implicit* when participants lack awareness of what they  
120 have learned. This means that if participants perform also above chance when they are  
121 unconfident, knowledge is inferred to be implicit (30). In contrast, if they perform at chance level  
122 when confidence is low, no implicit knowledge is gained. Although statistical learning may  
123 produce additional implicit knowledge that cannot be assessed by the recognition and memory  
124 judgement tasks (e.g., 28), Batterink's earlier findings show that adults store the acquired word  
125 knowledge mainly in the explicit memory system.

126 We and others have proposed that cognitive development and maturation of the prefrontal areas  
127 negatively affect language acquisition, such as word-form or grammar learning (31-35). For  
128 instance, we showed that children outperform adults on the Hebb repetition learning paradigm  
129 (32-33), a memory paradigm in which participants are asked to immediately recall syllable  
130 sequences that consist of hidden repeated word-forms. Interestingly, in a follow-up study, we  
131 found that cognitive depletion by Transcranial Magnetic Stimulation (TMS) to the left dorso-lateral  
132 prefrontal cortex (DLPFC), an area closely related to declarative memory and cognitive control,

133 enhanced Hebb performance in adult participants (34). This suggests that late-developing  
134 prefrontal cognitive mechanisms can induce changes in efficiently acquiring sequential language  
135 information from the environment, a finding that is largely in line with previously reported evidence  
136 in skill learning (13). Recently, we corroborated this idea further by showing enhanced  
137 phonotactic constraint learning in adults under cognitive fatigue (35). Based on these findings, we  
138 hypothesize that the higher cognitive control system could reduce access to implicit memory  
139 processes in adults, thereby making them less efficient in language acquisition relative to infants  
140 and children. This idea is in line with the well-known less-is-more hypothesis that attributes  
141 developmental changes in language acquisition, such as phonology and grammar, to  
142 maturational changes in attention and memory capacities (36-38). In our previous work,  
143 participants were explicitly asked to memorize (34) or produce (35) syllable sequences and thus  
144 exposure to the novel language was not passive, or “infant-like”. Moreover, we did not separate  
145 implicit and explicit memory representations. Thus it remains unresolved how higher-order  
146 cognitive functions affect acquisition of implicit linguistic knowledge during passive listening to  
147 continuous speech, using statistical learning mechanisms that support infant language acquisition  
148 (23, 39).

149 The aim of the current study was to directly address this question using the auditory statistical  
150 learning paradigm. In particular, we aimed to determine whether a temporary depletion of the  
151 higher cognitive control system, using two different interventions, can unlock adults’ implicit  
152 statistical learning processes that serve infant word segmentation. To investigate this, we  
153 exposed young adults to continuous streams of syllables with, unknown to them, repeating three-  
154 syllable pseudo-words, while watching a silent film. In the first experiment, inhibitory continuous  
155 theta-burst stimulation was used to induce a long-lasting disruption in left DLPFC or a control site,  
156 prior to exposure, similar to the method used in Smalle et al., 2017 (34). In the second  
157 experiment, participants first performed an effortful dual working-memory task under high or low  
158 cognitive load conditions, which induces cognitive fatigue that hampers subsequent cognitive  
159 performance (7, 35, 40), or did not perform a cognitive load task prior to the language exposure  
160 (control or no load condition). Our primary measure of statistical learning was the offline  
161 recognition of the hidden words, which was assessed 15 minutes after exposure. This was  
162 combined with a memory judgement procedure, which measured how confident the participants  
163 were that they remembered the hidden words. This task dissociates explicit versus implicit  
164 memory representations (e.g., 27-29, 41). In both experiments, electroencephalography (EEG)  
165 was also measured during the 20-min language exposure in order to investigate an online  
166 perceptual component, as second independent measure of statistical learning. Research has  
167 shown that the steady state response of the brain shows a decrease at the frequency of individual  
168 syllables and an increase at the rhythm of 3-syllable words while listening to continuous sound  
169 streams that consist of repeating 3-syllable structures. This shift in neural entrainment indicates  
170 online statistical learning of novel words as a function of auditory exposure (29). Overall, we  
171 predicted that TMS-induced disruption of the DLPFC (in Experiment 1) and cognitive fatigue (in  
172 Experiment 2) would enhance statistical language learning and especially strengthen implicit  
173 memory representations for the hidden novel words.

174

## 175 Results

176

### 177 *The effect of cognitive depletion on recognition of the hidden words*

178 **Experiment 1.** Participants performed above chance on the forced choice recognition task,  
179 indicating statistical learning, in both groups (i.e., DLPFC: mean = 68.8, SE = 3.5,  $t_{17} = 5.4$ ,  $p <$   
180  $.001$ ,  $d = 1.3$ ; Vertex: mean = 57.3, SE = 3.4,  $t_{17} = 2.2$ ,  $p < .05$ ,  $d = .5$ ). TMS-induced disruption of  
181 the DLPFC improved recognition accuracy of the hidden words [the effect of TMS:  $\beta = 0.28$ , SE =  
182  $0.096$ ,  $Z = 2.87$ ;  $X^2(1) = 8.25$ ,  $p = .004$ ,  $d = .6$ , Figure 1]. For the unconfident responses (64% of  
183 all trials; 31% in TMS disrupted group, 33% in control group), accuracy was significantly above

184 chance in the disrupted group [mean = 63.9, SE = 3.7,  $t_{17} = 3.8$ ,  $p < .001$ ,  $d = .9$ ], but not in the  
185 control group [mean = 51.7, SE = 3.8,  $t < 1$ ,  $p = .33$ ,  $d = .1$ ]. The control and disrupted groups  
186 differed significantly from each other [the main effect of TMS:  $\beta = 0.23$ , SE = 0.108,  $Z = 2.16$ ,  
187  $\chi^2(1) = 4.68$ ,  $p = .030$ ,  $d = .5$ ]. For the confident responses (36% of all trials; 19% in TMS  
188 disrupted group, 17% in control group), all participants showed above-chance accuracy [DLPFC:  
189 mean = 75.9, SE = 6.0,  $t_{17} = 4.3$ ,  $p < .001$ ,  $d = 1.0$ ; Vertex: mean = 66.3, SE = 7.8,  $t_{16} = 2.08$ ,  $p <$   
190  $.05$ ,  $d = .5$ ]. There was no significant difference between groups [the main effect of TMS:  $\beta =$   
191  $0.35$ , SE = 0.214,  $Z = 1.62$ ,  $\chi^2(1) = 2.63$ ,  $p = .11$ ].

192 **Experiment 2.** Participants performed above chance on the forced choice recognition task,  
193 indicating statistical learning, in all groups [High load: mean = 64.7, SE = 2.7,  $t_{19} = 5.4$ ,  $p < .001$ ,  $d =$   
194  $1.2$ ; Low load: mean = 65.6, SE = 3.9,  $t_{19} = 4.0$ ,  $p < .001$ ,  $d = .9$ ; No load: mean = 55.3, SE =  
195  $2.7$ ,  $t_{19} = 1.9$ ,  $p < .05$ ,  $d = .4$ ]. Cognitive load improved recognition accuracy of the hidden words  
196 [Effect of cognitive load: High vs Control:  $\beta = 0.42$ , SE = 0.19,  $Z = 2.23$ ,  $p = .026$ ,  $d = .4$ ; Low vs.  
197 Control:  $\beta = 0.46$ , SE = 0.19,  $Z = 2.47$ ,  $p = .013$ ,  $d = .5$ ;  $\chi^2(2) = 7.50$ ,  $p = .024$ , Figure 1]. For the  
198 unconfident responses (64% of all trials; 21% in high, 20% in low and 24% in no load group),  
199 accuracy was above chance in the high cognitive load group [mean = 61.8, SE = 3.8,  $t_{19} = 3.08$ ,  $p$   
200  $< .01$ ,  $d = .7$ ], and in the low cognitive load group [mean = 57.8, SE = 4.7,  $t_{19} = 1.64$ ,  $p = .05$ ,  $d =$   
201  $.4$ ], but not in the no cognitive load group [mean = 51.3, SE = 4.5,  $t < 1$ ,  $p = .4$ ,  $d = .07$ ]. The  
202 cognitive load enhanced accuracy relatively to the control group [High vs Control:  $\beta = 0.52$ , SE =  
203  $0.20$ ,  $Z = 2.57$ ,  $p = .01$ ,  $d = .5$ ; Low vs. Control:  $\beta = 0.37$ , SE = 0.20,  $Z = 1.81$ ,  $p = .070$ ,  $d = .4$ ;  
204 the main effect of cognitive load:  $\chi^2(2) = 7.16$ ,  $p = .028$ ]. For the confident responses (36% of all  
205 trials; 12% in high, 14% in low and 10% in no load group), all groups showed above-chance  
206 accuracy [High cognitive load: mean = 66.0, SE = 6.5,  $t_{19} = 2.5$ ,  $p = .012$ ,  $d = .6$ ; Low cognitive  
207 load: mean = 75.1, SE = 6.3,  $t_{19} = 4.0$ ,  $p < .001$ ,  $d = .9$ ; no cognitive load: mean = 68.7, SE = 5.7,  
208  $t_{19} = 3.3$ ,  $p < .01$ ,  $d = .7$ ]. No significant differences were found between the groups [High vs  
209 Control:  $\beta = -0.012$ , SE = 0.45,  $Z = -0.028$ ,  $p = .98$ ; Low vs. Control:  $\beta = 0.43$ , SE = 0.46,  $Z = 0.$   
210  $94$ ,  $p = .35$ ; the main effect of cognitive load:  $\chi^2(2) = 1.24$ ,  $p = .54$ ].

## 211 ***The effect of cognitive depletion on neural entrainment during exposure***

212 **Experiment 1.** During exposure, neural entrainment to the underlying word patterns was  
213 enhanced during structured exposure relative to random exposure: ITC increased at the word  
214 frequency ( $p = .01$ , Cohen's  $d = 1.1$ ) but decreased at the syllable frequency ( $p = .003$ ,  $d = 1.4$ )  
215 relative to exposure to the random syllable stream [Frequency x Exposure:  $F(1, 31) = 16.4$ ,  $p <$   
216  $.001$ , Figure 2]. Disrupting the left DLPFC with TMS increased the neural word-learning index  
217 [Exposure:  $F(1, 30) = 31.6$ ,  $p < .001$ , TMS x Exposure:  $F(1, 30) = 5.6$ ,  $p = .025$ , random:  $p = .6$ ,  
218  $d = .2$ , structured:  $p = .004$ ,  $d = 1.1$ , Figure 3].

219 **Experiment 2.** During exposure, neural entrainment to the underlying word patterns was  
220 enhanced during structured exposure relative to random exposure: ITC increased at word  
221 frequency ( $p < .001$ ,  $d = 1.4$ ) but decreased at syllable frequency ( $p = .08$ , Cohen's  $d = .5$ ) relative  
222 to exposure to a random syllable sequence [Frequency x Exposure:  $F(1, 57) = 19.3$ ,  $p < .001$ ,  
223 Figure 2]. Cognitive load did not affect the neural word learning index [Exposure:  $F(1, 55) = 40.4$ ,  
224  $p < .001$ , Load x Exposure:  $F(2, 55) = 1.04$ ,  $p = .36$ , Figure 3].

225

## 226 **Discussion**

227

228 Overall, our findings provide evidence for a competitive interaction between higher cognitive  
229 control functions and implicit statistical learning mechanisms that contribute to word-segmentation  
230 in the early stages of language acquisition. More specifically, we depleted the cognitive control  
231 mechanisms in young adults by applying TMS to the left DLPFC (Experiment 1) and by inducing  
232 cognitive fatigue (Experiment 2). These interventions with long-lasting effects were applied prior  
233 to exposure to a continuous stream of speech sounds that consisted of tri-syllabic word patterns.

234 There were no acoustic cues of the word boundaries in the streams. Both TMS-induced  
235 disruption of the left DLPFC and cognitive fatigue enhanced recognition accuracy for the hidden  
236 words, indicating enhanced statistical learning. Intriguingly, these cognitive manipulations  
237 specifically enhanced recognition accuracy when the participants had low confidence in  
238 remembering the hidden words, indicating enhanced implicit statistical learning. In fact, the  
239 control adults showed chance-level accuracy for the words stored in implicit memory in  
240 agreement with earlier studies in adults (27-29), whereas the cognitively depleted adults showed  
241 above chance-level recognition accuracy. The results therefore suggest that cognitive depletion  
242 unlocks implicit word-segmentation abilities in adults.

243

244 All groups showed above-chance recognition accuracy when confidence in remembering the  
245 hidden words was high, indicating explicit statistical learning. Cognitive manipulations had no  
246 effect on explicit statistical learning. This is important because the cognitive manipulations could  
247 have reduced explicit or declarative learning mechanisms. Our findings suggest that the  
248 enhancement of implicit statistical learning did not occur at the expense of explicit memory  
249 formation and thus does not completely support the idea of a direct competition between implicit  
250 and explicit learning. These observations are more in line with the idea that implicit statistical  
251 learning remains available across development (39), but that the adult brain imposes a  
252 “bottleneck” which prioritizes access to the explicit memory system (42). Disrupting the higher  
253 cognitive control mechanisms abolishes this “bottleneck” and, as a consequence, improves  
254 adults’ capacity to simultaneously store linguistic knowledge in both implicit and explicit memory  
255 systems. This results in overall enhanced word recognition in the disrupted adults. Further  
256 research is needed to test this model and investigate how facilitating the higher-order cognitive  
257 system affects acquisition of implicit versus explicit linguistic knowledge.

258

259 Our experimental paradigm included a break after the exposure to allow participants to recover  
260 from the cognitive manipulations before performing the memory tests in Experiment 1 and 2. It is  
261 possible that this delay strengthened implicit memory consolidation rather than episodic recall so  
262 that participants were less likely to explicitly remember the words from the stream. In fact, in both  
263 experiments, participants were confident in only 36% of the total trials. The memory judgement  
264 (i.e., the proportion of confident versus unconfident responses) was however unaffected by TMS  
265 or cognitive fatigue. Cognitive manipulations specifically affected accuracy of the unconfident, but  
266 not the confident, responses after a 15-min delay. Further research is needed to investigate  
267 whether this effect is dependent on a short consolidation period and whether a longer  
268 consolidation period would further strengthen this effect.

269

270 We additionally measured neural entrainment during statistical learning in order to investigate  
271 perceptual binding of neighbouring syllables into words during passive listening to the structured  
272 syllable streams, as a second, independent marker of learning. As expected, based on recent  
273 findings by Batterink and Paller (2017), neural entrainment decreased at the frequency of  
274 syllables while it increased at the level of the words as a function of the structured exposure (29).  
275 This indicates successful perceptual binding across all participants. Interestingly, however, TMS-  
276 induced disruption of the left DLPFC enhanced this perceptual binding (measured with a word-  
277 learning index) whereas cognitive fatigue had no effect on it, although both TMS and cognitive  
278 fatigue enhanced memory for the words as measured in the post-exposure recognition task.  
279 These findings are in line with the view that a perceptual binding (also called, ‘processing-based’)  
280 component of statistical learning (as measured here with online EEG) is dissociable from a  
281 memory storage or retrieval component of statistical learning (as measured here with the offline  
282 recognition task) (26, 43, 51). This is further supported by the absence of correlation between the  
283 online and offline measures (see Supplementary Information). Exploratory oscillatory power  
284 analyses (see Supplementary Information) showed that TMS-induced disruption of the DLPFC  
285 marginally enhanced the overall power of theta (4-8 Hz) and alpha oscillations (8-12 Hz), and  
286 significantly enhanced alpha oscillations in the parietal-occipital area specifically. Cognitive  
287 fatigue, in contrast, *decreased* the overall power of delta (1-4 Hz) oscillations. This suggests that

288 the cognitive manipulations had different neural consequences. Theta and lower alpha power are  
289 strong in early childhood and it decreases during cognitive development (e.g., 44-45). Power of  
290 lower frequency oscillations, such as delta and theta, continue to decrease during adulthood (46).  
291 Thus, the pattern of results invites a tentative interpretation that disruption of the left DLPFC  
292 made the young adults to function in the same way as young children do and therefore enhanced  
293 both perceptual binding and implicit memory components of statistical learning. The cognitive  
294 fatigue manipulation, however, possibly mimicked cognitive decline throughout adulthood.  
295 Statistical language learning and implicit motor learning have been shown to be well-preserved in  
296 older adults (20, 47-50).

297

298 The current findings are largely in agreement with previous studies showing that TMS-induced  
299 disruption of the DLPFC and/or cognitive fatigue enhances sequence learning in the motor  
300 system (6-7) and the language system (34-35). To our knowledge, the present study is the first  
301 study to demonstrate the effects of cognitive depletion on implicit versus explicit outcomes of  
302 statistical language learning in the absence of explicit instructions (i.e., passive listening during  
303 exposure) and replicating these effects by using two different methods: TMS-induced disruption  
304 and cognitive fatigue. A previous study found that only high cognitive load induced feelings of  
305 cognitive fatigue and improved motor sequence learning (7). We found that both low and high  
306 load tasks increased feelings of cognitive fatigue in our participant sample, and consequently  
307 statistical language learning was enhanced in both groups (relative to a control condition where  
308 no cognitive load task was performed).

309

310 Statistical language learning relies on multiple brain regions and pathways that work in parallel,  
311 such as hippocampus (51), inferior frontal cortex (52), striatum (53) and auditory-motor pathways  
312 (54). It remains however poorly understood how these brain regions and pathways contribute to  
313 acquisition of implicit vs. explicit memory representations for linguistic knowledge from structured  
314 sound sequences. Another important goal for future research is to determine how the prefrontal  
315 cognitive control mechanisms affect competition and cooperation between various language  
316 learning mechanisms.

317

318 This experimental study provides causal evidence for a hypothesis that the cognitive control  
319 system constrains implicit language learning abilities in adults. Our findings show that depletion of  
320 the mature cognitive system can enhance implicit, statistical learning mechanisms that are used  
321 in early language acquisition. This finding is an important step in science of human language  
322 development as it could help us to understand maturational constraints and inter-individual  
323 differences (e.g., language-related difficulties) in language learning. Importantly, cognitive  
324 depletion could be a key for unlocking infant-like implicit learning mechanisms and, as a result,  
325 enhance foreign language learning in adults.

326

327

## 328 **Materials and Methods**

329

### 330 ***Participants***

331 We decided to test 20 participants in each group based on the large effect sizes in a previous  
332 study by Batterink and colleagues (28), who used similar dependent measures, and in our  
333 previous study in which we used an identical TMS paradigm as in the current study (34). In our  
334 previous study, we obtained a t-test effect size  $d$  of 0.88 for the difference in Hebb learning  
335 performance between a DLPFC-disrupted group ( $n = 14$ , mean correct recall on the last block of  
336 Hebb trials = 89.3%,  $SD = 16.04\%$ ) and a control group ( $n = 14$ , mean correct recall = 72.5%,  $SD$   
337 = 21.9%). After post-collection exclusion of four non-fluent English speaking participants, we  
338 report the data of 36 participants in Experiment 1 who were randomly assigned to either the left  
339 DLPFC stimulation ( $n = 18$ , age =  $25.3_M \pm 4.8_{SD}$ , 9 females) or the control stimulation to Vertex ( $n$   
340 = 18, age =  $23.4_M \pm 5.0_{SD}$ , 12 females). In three participants (two in the TMS group and one in the

341 control group), EEG was not recorded due to technical failures. Data of these participants were  
342 however still included for the behavioural analyses. In Experiment 2, we report the data of 60  
343 participants who were randomly assigned to a high load ( $n = 20$ , age =  $22.7_M \pm 3.1_{SD}$ , 12  
344 females), low load ( $n = 20$ , age =  $21.8_M \pm 3.9_{SD}$ , 16 females), or no load (i.e., control,  $n = 20$ , age  
345 =  $18.8_M \pm .81_{SD}$ , 16 females) condition. We have no EEG recording from two participants (one in  
346 the control group and one in the low load group) due to technical failure, but their behavioral data  
347 were included in the analyses. All included participants were right-handed and native (or non-  
348 native but fluent) English speakers. None of the participants had a history of language (learning)  
349 impairments or neurological problems. Participants from all groups were matched on various  
350 cognitive control abilities (see Table 1). Experiments were undertaken with written informed  
351 consent and blind to the purpose of the study. Participants received financial compensation at the  
352 end of the experiment (£10/hour). The study was approved by the Research Ethics Committee of  
353 the School of Psychology at the University of Nottingham (Reference: F1003).

### 354 ***Experimental Design***

355 During a pre-test, individual cognitive control abilities were assessed (Table 1). Participants in  
356 Experiment 2 were additionally pre-tested on their maximal processing speed capacity, i.e., the  
357 shortest processing time interval to simultaneously perform two working memory tasks while  
358 maintaining an accuracy of at least 85%. This assessment was necessary for the experimental  
359 cognitive load manipulation in Experiment 2 (see Methods later). For both experiments, the main  
360 experiment took place on a separate day during which the participants in both experiments were  
361 exposed to a 20-minute auditory syllable stream while EEG was recorded. After a 15-minute  
362 break, the participants completed a post-exposure recognition test that examined implicit and  
363 explicit memory of the hidden words. The main experimental design is presented in Figure 4.

### 364 ***Transcranial Magnetic Stimulation (Experiment 1)***

365 TMS was delivered using a 70-mm diameter figure-eight coil attached to DuoMAG XT stimulator  
366 (by Deymed, Brainbox Ltd, United Kingdom). We first localized the left DLPFC in each participant  
367 using the BeamF3 algorithm (55, 56). We then identified the left motor cortex as the spot eliciting  
368 reliable twitches in the resting contralateral hand. The active motor threshold (aMT) was defined  
369 as the lowest intensity at which TMS elicited at least five out of ten visible muscle twitches, whilst  
370 the subject sustained a light contraction of their pinch. After defining the participant's aMT, the  
371 coil was placed over the left DLPFC or a control area (2 cm posterior to vertex), similarly to our  
372 previous TMS study (34). The control area was assumed not to play a role in statistical learning  
373 or cognitive control (57). The location of the coil (i.e., DLPFC or Vertex) was decided randomly  
374 based on the number of participants entering the experiment. The coil was placed tangentially to  
375 the scalp with the handle pointing posterior at a  $45^\circ$  angle with respect to the anterior-posterior  
376 axis for DLPFC and at  $0^\circ$  for the control site. The intensity of the stimulation was set at 80% of  
377 each participant's aMT, i.e. at 43.7% ( $SD = 6.3$ ) for the DLPFC group and at 49.3% ( $SD = 8.3$ ) for  
378 the control group. Similar to Smalle et al (2017), a modified cTBS protocol was used in which 600  
379 pulses were delivered in a continuous train of 200 bursts. Each burst consisted of 3 pulses at 30  
380 Hz, repeated at 6 Hz. The total stimulation duration was 30 seconds. This modified cTBS protocol  
381 is known to inhibit cortical excitability for at least 30 minutes after stimulation over the primary  
382 motor area (58). Importantly, cTBS to the DLPFC does not impair metacognition or conscious  
383 perception processes (59).

### 384 ***Cognitive fatigue (Experiment 2)***

385 Cognitive fatigue was induced with the TloadDback task (40). The script of the TloadDback task  
386 is freely available on Open Science Framework ([osf.io/ay6er](https://osf.io/ay6er)). The task was run in  
387 Matlab2016b/Psychtoolbox on a Dell laptop (refresh rate 60Hz). The letters were centrally  
388 presented in Arial font size 120 on a 15.6inch screen. For each participant, the shortest time  
389 needed for accurately processing two ongoing task demands, namely n-back letter detection and



390 parity number decision, was defined during a pre-test on a first assessment day (see  
391 Experimental Design). During the TloadDback task, digits (1, 2, 3, 4, 6, 7, 8 and 9) and letters (A,  
392 C, T, L, N, E, U and P) were presented in alternation on the screen. Participants were instructed  
393 to press the space bar with their left hand every time the displayed letter was the same as the  
394 last-seen letter, and to indicate with their right hand whether the subsequently displayed digit was  
395 odd (pressing “1” on the numeric keypad) or even (pressing “2”). Different levels of cognitive load  
396 were created by presenting the two tasks at different paces based on our participant’s pre-tested  
397 maximum processing speed capacities (no a priori group differences, see Table 1). This is  
398 defined as the fastest stimulus time duration (STD) allowing an accuracy performance of at least  
399 85%. Under High Cognitive Load (HCL) conditions, the task was performed for 16 minutes at the  
400 subject’s max. STD while under Low Cognitive Load (LCL) conditions, the presentation rate was  
401 made 1/3 slower (i.e.  $STD = \max. STD + \frac{1}{2} \max. STD$ ). This results in different cognitive  
402 demands, with higher sustained attentional-control requirements for the former condition (despite  
403 the same level of task complexity), eventually leading to a higher state of ‘cognitive depletion or  
404 mental fatigue’ (7, 40). As expected, the low cognitive load participants showed higher dual-task  
405 performance than the high cognitive load participants, who performed around the minimal 85%  
406 accuracy level defined during the pretest: i.e.,  $93.3_M \pm 4.3_{SD}$  vs.  $81.5_M \pm 13.1_{SD}$ ,  $t_{38} = 23.1$ ,  $p < .001$ ,  
407 respectively. Participants under the no load condition immediately started with the main  
408 experiment that is visualized in Figure 4. A simple numeric self-report rating scale assessing  
409 fatigue (1: I feel no mental fatigue, to 10: I feel the worst possible mental fatigue) was presented  
410 immediately before the TloadDback task and immediately after (i.e., before exposure), as a quick  
411 manipulation check for the induction of cognitive fatigue. Participants who performed no cognitive  
412 load task reported lower subjective feelings of fatigue prior to exposure (i.e.,  $4.2_M \pm 2.01_{SD}$ ) than  
413 participants who performed the cognitive load task (i.e.,  $5.3_M \pm 1.9_{SD}$ ;  $p = 0.001$ ). However,  
414 unexpectedly, there were no reliable differences between the high and low load participants (i.e.,  
415  $5.6_M \pm 2.1_{SD}$  vs.  $5.0_M \pm 2.3_{SD}$ ,  $p = 0.21$ ). There were no differences in baseline subjective reports  
416 for cognitive fatigue across all groups, tested at the start of the experiment (i.e., all  $p$ 's  $> .23$ ).

#### 417 **Exposure (Experiment 1 and 2)**

418 Twelve unique syllables of a consonant-vowel structure were selected and structured into four  
419 novel word-forms, i.e. /tu:paɪrou/, /goula:bu:/, /bi:da:ku:/, and /pa:di:ta:/). The individual syllables  
420 within each word-form occurred at a first, second and third position across participants so that in  
421 each stimulation group the subjects 1-7 were exposed to the word-forms as listed above, while  
422 subjects 8-14 received the word-forms /paɪrou:tu:/, /la:bu:gou/, /da:ku:bi:/, /di:ta:pa:/ and subjects  
423 15-21 the word-forms /rou:tu:paɪ/, /bu:goula:/, /ku:bi:da:/, /ta:pa:di:/. This was done to minimize  
424 any stimulus-driven effects that could be caused by position-preferences for syllables within a  
425 word. Across all language lists, the words were matched on average English phonotactic  
426 probability ( $p_s > .40$ ). The twelve syllables were recorded using an online artificial speech  
427 synthesizer of a female British English voice. The audio files were edited to have a duration of  
428 250 ms and saved with a sampling rate of 44100 Hz using Audacity software.

429 Participants were informed that they would hear a continuous stream of speech sounds and were  
430 asked to listen carefully to the sounds. No information was given about the hidden structures, nor  
431 about a post-exposure test on segments of the heard syllable sequences, hence language  
432 exposure was implicit. Exposure always started with a random stream, in which all twelve  
433 syllables were concatenated in a pseudorandom order without any higher-order structure; the  
434 only constraint was that syllables did not repeat and that no anagrams of the novel words  
435 appeared. In this stream, 900 syllables were presented (each syllable was repeated 75 times).  
436 After the random stream, the 20-minute structured stream started in which the speech sounds  
437 were grouped into four repeating tri-syllabic words (Figure 4). Here, the transitional probability  
438 between neighbouring syllables within words was 100% and 33% between words. For instance,  
439 for subject 1-3, /tu:/ in the stream is always followed by /paɪ/ while /rou/ could be equally followed  
440 by /gou/, /bi:/ or /pa:./.

441 In total, 1200 repeating word-forms (each word was repeated 300 times) and 3600 syllables (plus  
442 the syllables from the random block) were presented. Stimulus onset asynchrony was 320 ms in  
443 Experiment 1 and 310 ms in Experiment 2 (this 10-ms difference between the experiments was  
444 unintended). The speech stream was presented using Presentation® software (Version 18.0,  
445 Neurobehavioral Systems, Inc., Berkeley, CA, [www.neurobs.com](http://www.neurobs.com)). The syllables were presented  
446 at a comfortable listening level for each participant through inserted earphones attached to a Dell  
447 desktop computer. During exposure, participants watched an episode of Planet Earth in silent  
448 mode (i.e., without subtitles or sound). Every ~5 minutes, a short break (10s) was inserted which  
449 reminded the participants to attend the sounds through a visual instruction on the screen.

#### 450 ***EEG recording and analyses (Experiment 1 and 2)***

451 We recorded an electroencephalogram (EEG) with 27 cap-mounted electrodes (Fp1, Fp2, Fz, F3,  
452 F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, T3, T4, CP1, CP2, CP5, CP6, Pz, P3, P4, T5, T6,  
453 O1, O2) using a TMS-compatible EEG system (TruScan Research by Deymed, Brainbox Ltd,  
454 United Kingdom). The nose electrode was used as reference, and an electrode attached to the  
455 forehead was used as a ground during recordings. Horizontal and vertical electro-oculograms  
456 (EOGs) were recorded bipolarly with electrodes placed at the outer canthi of the eyes, and above  
457 and below the right eye. Electrode impedance was kept below 10 k $\Omega$ . Signals were filtered online  
458 with a 0.1-1000 Hz band-pass filter and recorded at a sampling rate of 3000 Hz.

459 The EEG data analyses were performed using EEGLAB which is an open-source toolbox  
460 operated in the MATLAB environment (60). The continuous raw data files were re-referenced to  
461 the algebraic mean of the left and right mastoid electrodes, down-sampled to 500 Hz and filtered  
462 at 30-Hz low-pass filter. Bad channels were identified and interpolated when necessary. Mean  
463 number of interpolated channels per participant was 0.18 in Experiment 1, while there were no  
464 interpolated channels in Experiment 2. Periods of EEG signal containing strong muscle artifacts,  
465 electrode drifts, or technical artifacts were removed from further analysis. In Experiment 1 (TMS),  
466 one participant (from the control group) was rejected due to very noisy EEG activities, while no  
467 extra participant was excluded in Experiment 2. Independent component analysis (ICA) was used  
468 for linear decomposition of the continuous data to remove the contributions of artifact sources  
469 (slow drifts, eye blink/movement, and muscle artifacts) on the scalp sensors. The data were  
470 segmented into 5,000-ms epochs (-2000 to 3000 ms relative to the onset of each word in word  
471 condition or every third syllable in the random condition). Amplitude changes exceeding  $\pm 70 \mu\text{V}$   
472 (including the EOG channels) were removed from further analysis (less than 20% of the trials).  
473 Selective response averaging was conducted separately for each block.

474 We quantified neural entrainment at the syllabic and word frequencies by measuring inter-trial  
475 coherence (ITC) within each condition (word vs. syllable / random vs. structured). ITC is known  
476 as phase-locking value and is a measure of event-related phase locking. The higher the ITC  
477 value is, the higher phase coincidence across epochs is. That is, the ITC values range from 0 to  
478 1, from purely non-phase-locked activity to strictly phase-locked activity. ITC was calculated from  
479 0.5 to 5 Hz in 0.1 Hz steps using a continuous Morlet wavelet transformation in which the number  
480 of cycles is increased linearly with frequency from 1 cycle length to obtain better frequency  
481 resolution at higher frequencies. This approach can optimize the trade-off between temporal  
482 resolution at lower frequencies and frequency resolution at high frequencies (60). Then, ITC  
483 values in each epoch were averaged.

484 The word presentation frequency was 1.0Hz and 1.1Hz and the syllable presentation frequency  
485 was 3.1 and 3.2 Hz in Experiments 1 and 2, respectively. If participants become more sensitive  
486 to the underlying word structure in the continuous speech stream, we should observe a higher  
487 ITC at the word-frequency but a lower ITC at the syllable frequency during exposure to the  
488 structured sequence relative to exposure to the random sequence. In other words, if participants  
489 learn the novel words, they would show a preferential shift in the entrainment of neural  
490 oscillations to underlying words, relative to individual syllables. This can also be indexed by a

491 simple formula, also referred to as the Word Learning Index ( $WLI = ITC_{\text{word frequency}} / ITC_{\text{syllable}}$   
492  $_{\text{frequency}}$ ) (29). The WLI was computed across 6 centro-frontal midline electrodes where ITC at the  
493 word and syllable frequencies showed the strongest values (i.e., FC1, FC2, F3, F4, FC5, Fz;  
494 supplementary documentation on ITC across electrodes is available in our open science  
495 repository, <https://osf.io/dequ9/>).

#### 496 ***Forced Choice Recognition Task (Experiment 1 and 2)***

497 After a break of ~15 minutes during which participants could remove the EEG cap and wash their  
498 hair, participants completed the two-alternative forced-choice recognition task. For each trial, a  
499 fixation cross appeared while the auditory presentation of a target 3-syllable string (hidden word)  
500 and a foil 3-syllable string (non-word) were presented, separated by an inter-stimulus interval of  
501 1500ms. The nonword foils were created from the same list of twelve unique syllables that were  
502 structured into word-forms. The only restriction was that the syllables within the foils never  
503 followed each other in the speech stream, not even across word boundaries. All words and  
504 nonword foils were matched on average English phonotactic probability ( $p_s > .70$ ). The task was  
505 1) to indicate which of the two strings sounded more familiar, and 2) to judge on their recall  
506 decision ('I recalled from exposure' versus 'It sounds familiar but I have no clear memory' or 'I  
507 guessed'). Strings that were 'recalled' are referred to as confident responses and the strings that  
508 were 'familiar without memory' or 'guessed' are referred to as unconfident responses. The next  
509 trial started 1500 ms after the participant entered his or her response. The syllable strings were  
510 presented at the same rate as during the exposure. Each of the four targets and four foils were  
511 paired exhaustively to a total of 16 trials. In half of the trials, a target was followed by a foil while  
512 in the other half a foil was followed by a target. The order of presentation was counterbalanced  
513 across participants.

#### 514 ***Statistical Analyses (Experiment 1 and 2)***

515 To investigate the effect of cognitive depletion on language learning, linear mixed effect analyses  
516 were performed on the EEG data (i.e. ITC values and WLI indexes) and hierarchical logistic  
517 regression analyses were performed on the behavioural data (i.e., recognition accuracy). These  
518 analyses were performed using the lme4 package (61) and the afex package (62) in R (R  
519 Development Core Team, 2011). We always strived for models including maximal random effects  
520 structure justified by the design (61, 62). In case of convergences issues (e.g., singular fits), we  
521 refitted the maximal model by first removing correlations among random slopes, after which the  
522 highest order random slopes with the least estimated variance were removed (62). The  $p$  values  
523 were derived using Kenward-Roger approximations for degrees of freedom with the anova  
524 function in the afex package (63). Effects-coding was used for all fixed factors, except for the  
525 Group factor in Experiment 2 where dummy coding was used with No Cognitive Load as  
526 reference level. Bonferroni correction was used for all planned tests. Cohen's  $d$  effect sizes on  
527 the model's estimates are calculated with the eff\_size function from the emmeans package (64).  
528 To allow across-study comparison and facilitate secondary analyses (e.g., power calculations and  
529 meta-analyses), we also provide t-test effect sizes in the Supplementary Information. We  
530 additionally performed one-sample t-tests to test for above-chance performance for the  
531 unconfident and confident responses in the recognition task. One control participant in  
532 Experiment 1 reported low confidence in all trials and so did not have confident responses. The  
533 stimulus materials and data-files, including scripts for analysis are available on an open science  
534 repository: <https://osf.io/dequ9/>.

535

#### 536 **Acknowledgments**

537

538 We thank Cristina Pancotto, Afrina Sallehuddin and Ayaka Tsuchiya for help with data collection.  
539 ES was funded by a grant from the Research Foundation – Flanders (1211421N). TD was funded  
540 by a travel grant from the Suntory Foundation. The research costs were covered by start-up funds

541 from the School of Psychology (University of Nottingham) and the TMS and EEG equipment by  
542 Medical Research Council fellowship (G1000566) to RM. RM was also supported by Profi5 (Mind  
543 and Matter) funding from the Academy of Finland to the University of Helsinki.  
544  
545  
546  
547

548 **References**

- 549 1. R. A. Poldrack *et al.*, Interactive memory systems in the human brain. *Nature* **414**(6863), 546-  
550 550 (2001).
- 551 2. M. T. Ullman, Contributions of memory circuits to language: the declarative/procedural model.  
552 *Cognition* **92**(1), 231-270 (2004).
- 553 3. M. H. De Vries *et al.*, Electrical stimulation of Broca's area enhances implicit learning of an  
554 artificial grammar. *J. Cogn. Neurosci.* **22**(11), 2427-2436 (2010).
- 555 4. J. Uddén *et al.*, The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain*  
556 *Res.* **1224**, 69-78 (2008).
- 557 5. M. T. Ullman, Is Broca's Area Part of a Basal Ganglia Thalamocortical Circuit? *Cortex*, **42**(4),  
558 480-485 (2006).
- 559 6. G. G. Ambrus *et al.*, When less is more: Enhanced statistical learning of non-adjacent  
560 dependencies after disruption of bilateral DLPFC. *J. Mem. Lang.* **114**, 104144 (2020).
- 561 7. G. Borragán, H. Slama, A. Destrebecqz, P. Peigneux, Cognitive Fatigue Facilitates  
562 Procedural Sequence Learning. *Front. Hum.*  
563 *Neurosci.* <https://doi.org/10.3389/fnhum.2016.00086> (2016).
- 564 8. K. Foerde, B. J. Knowlton, R. A. Poldrack, Modulation of competing memory systems by  
565 distraction. *Proc. Natl. Acad. Sci. U.S.A.* **103**(31), 11778-11783 (2006).
- 566 9. J. M. Galea, N. B. Albert, T. Ditye, R. C. Miall, Disruption of the dorsolateral prefrontal cortex  
567 facilitates the consolidation of procedural skills. *J. Cogn. Neurosci.* **22**(6), 1158-1164 (2010).
- 568 10. D. Nemeth, K. Janacek, B. Polner, Z. A. Kovacs, Boosting human learning by hypnosis.  
569 *Cereb. Cortex.* **23**(4), 801-805 (2013).
- 570 11. M. Virag *et al.*, Competition between frontal lobe functions and implicit sequence learning:  
571 evidence from the long-term effects of alcohol. *Exp. Brain Res.* **233**(7), 2081-2089 (2015).
- 572 12. M. J. Frank, R. C. O'Reilly, T. Curran, When memory fails, intuition reigns: midazolam  
573 enhances implicit inference in humans. *Psychol. Sci.* **17**(8), 700-707 (2006).
- 574 13. S. L. Thompson-Schill, M. Ramscar, E. G. Chrysikou; Cognition without control: When a little  
575 frontal lobe goes a long way. *Curr. Dir. Psychol. Sci.* **18**(5), 259-263 (2009).
- 576 14. A. R. Romberg, J. R. Saffran, Statistical learning and language acquisition. *Wiley Interdiscip.*  
577 *Rev. Cogn. Sci.* **1**(6), 906-914 (2010).
- 578 15. D. Kennedy, C. Norman, What Don't We Know? *Sci.* **309**(5731), 75-75 (2005).
- 579 16. M. T. Ullman, "The declarative/procedural model: a neurobiological model of language  
580 learning, knowledge and use" in *The Neurobiology of Language*, G. Hickok, S. A. Small, Eds.  
581 (Elsevier, 2016), pp. 953-968.
- 582 17. K. Graf Estes, J. L. Evans, M. W. Alibali, J. R. Saffran, Can Infants Map Meaning to Newly  
583 Segmented Words?: Statistical Segmentation and Word Learning. *Psychol. Sci.* **18**(3), 254-  
584 260 (2007).
- 585 18. M. G. Gaskell, A. W. Ellis, Word learning and lexical development across the lifespan. *Philos.*  
586 *Trans. R. Soc. B.* **364**, 3607-3615 (2009).
- 587 19. J. R. Saffran, R. N. Aslin, E. L. Newport, Statistical learning by 8-month-old infants. *Sci.*  
588 **274**(5294), 1926-1928 (1996).
- 589 20. S. D. Palmer, J. Hutson, S. L. Mattys, Statistical learning for speech segmentation: Age-  
590 related changes and underlying mechanisms. *Psychol. Aging.* **33**(7), 1035-1044 (2018).
- 591 21. J. R. Saffran, E. L. Newport, R. N. Aslin, R. A. Tunick, S. Barrueco, Incidental language  
592 learning: Listening (and learning) out of the corner of your ear. *Psychol. Sci.* **8**(2), 101-105  
593 (1997).
- 594 22. R. Frost, B. C. Armstrong, M. H. Christiansen, Statistical learning research: A critical review  
595 and possible new directions. *Psychol. Bull.* **145**(12), 1128-1153 (2019).
- 596 23. C. M. Conway, How does the brain learn environmental structure? Ten core principles for  
597 understanding the neurocognitive mechanisms of statistical learning. *Neurosci. Biobehav.*  
598 *Rev.* **112**, 279-299 (2020).
- 599 24. P. Perruchet, S. Pacton, Implicit learning and statistical learning: One phenomenon, two  
600 approaches. *Trends Cogni. Sci.* **10**(5), 233-238 (2006).

- 601 25. T. Teinonen, V. Fellman, R. Näätänen, P. Alku, M. Huotilainen, Statistical language learning  
602 in neonates revealed by event-related brain potentials. *BMC Neurosci.* **10**, 21-21 (2009).
- 603 26. L. J. Batterink, K. A. Paller, P. J. Reber, Understanding the Neural Bases of Implicit and  
604 Statistical Learning. *Top. Cogn. Sci.* **11**(3), 482-503 (2019).
- 605 27. L. J. Batterink, P. J. Reber, H. J. Neville, K. A. Paller, Implicit and explicit contributions to  
606 statistical learning. *J. Mem. Lang.* **83**, 62-78 (2015).
- 607 28. L. J. Batterink, P. J. Reber, K. A. Paller, Functional differences between statistical learning  
608 with and without explicit training. *Learn. Mem.* **22**(11), 544-556 (2015).
- 609 29. L. J. Batterink, K. A. Paller, Online neural monitoring of statistical learning. *Cortex* **90**, 31-45  
610 (2017).
- 611 30. Z. Dienes, D. Berry, D., Implicit learning: Below the subjective threshold. *Psychon. Bull.*  
612 *Rev.* **4**(1), 3-23 (1997).
- 613 31. A. S. Finn, T. Lee, A. Kraus, C. L. Hudson Kam, When It Hurts (and Helps) to Try: The Role  
614 of Effort in Language Learning. *PLoS One* **9**(7), e101806 (2014).
- 615 32. E. H. Smalle *et al.*, Can chunk size differences explain developmental changes in lexical  
616 learning? *Front. Psychol.* **6**, 1925 (2016).
- 617 33. E. H. Smalle, M. P. Page, W. Duyck, M. G. Edwards, A. Szmalec, Children retain implicitly  
618 learned phonological sequences better than adults: A longitudinal study. *Dev. Sci.* **21**(5),  
619 e12634 (2018).
- 620 34. E. H. Smalle, M. Panouilleres, A. Szmalec, R. Möttönen, Language learning in the adult  
621 brain: disrupting the dorsolateral prefrontal cortex facilitates word-form learning. *Sci. Rep.*  
622 **7**(1), 13966 (2017).
- 623 35. E. H. Smalle, M. Muylle, W. Duyck, A. Szmalec, Less is more: Depleting cognitive resources  
624 enhances language learning abilities in adults. *J. Exp. Psychol. Gen.* Advance online  
625 publication. <https://doi.org/10.1037/xge0001058> (2021).
- 626 36. E. L. Newport, Maturation constraints on language learning. *Cogn. Sci.* **14**(1), 11-28 (1990).
- 627 37. J. L. Elman, Learning and development in neural networks: the importance of starting  
628 small. *Cognition* **48**, 71-99 (1993).
- 629 38. E. L. Newport, D. Bavelier, H. J. Neville, "Critical thinking about critical periods: Perspectives  
630 on a critical period for language acquisition" in *Language, brain, and cognitive development:  
631 Essays in honor of Jacques Mehler*, E. Dupoux, Ed. (The MIT Press, 2001), pp. 481-502.
- 632 39. E. D. Thiessen, S. Girard, L. C. Erickson, Statistical learning and the critical period: how a  
633 continuous learning mechanism can give rise to discontinuous learning. *Wiley Interdiscip.*  
634 *Rev. Cogn. Sci.* **7**(4), 276-288 (2016).
- 635 40. G. Borragan, H. Slama, M. Bartolomei, P. Peigneux, Cognitive fatigue: A Time-based  
636 Resource-sharing account. *Cortex* **89**, 71-84 (2017).
- 637 41. J. Bertels, A. Franco, A. Destrebecqz, How implicit is visual statistical learning? *J. Exp.*  
638 *Psychol. Learn. Mem. Cogn.* **38**, 1425-1431 (2012).
- 639 42. J. Breton, E. M. Robertson, Flipping the switch: mechanisms that regulate memory  
640 consolidation. *Trends Cogn. Sci.* **18**(12), 629-634 (2014).
- 641 43. M. H. Christiansen, Implicit Statistical Learning: A Tale of Two Literatures. *Top. Cogn. Sci.*  
642 **11**(3), 468-481 (2019).
- 643 44. C. Benninger, P. Matthis, D. Scheffner, EEG development of healthy boys and girls. Results  
644 of a longitudinal study. *Electroencephalogr Clin. Neurophysiol.* **57**(1), 1-12 (1984).
- 645 45. P. J. Marshall, Y. Bar-Haim, N. A. Fox, Development of the EEG from 5 months to 4 years of  
646 age. *Clin. Neurophysiol.* **113**(8), 1199-208 (2002).
- 647 46. A. Hashemi *et al.*, Characterizing population EEG dynamics throughout adulthood. *eNeuro*  
648 **3**(6), 1-13 (2018).
- 649 47. J. H. Howard Jr., D. V. Howard, Aging mind and brain: is implicit learning spared in healthy  
650 aging? *Front. Psychol.* **4**, 817 (2013).
- 651 48. D. Juhasz, D. Nemeth, K. Janacsek, Is there more room to improve? The lifespan trajectory  
652 of procedural learning and its relationship to the between- and within-group differences in  
653 average response times. *PLoS One* **14**(7), e0215116 (2019).

- 654 49. M. Verneau, J. van der Kamp, G. J. Savelsbergh, M. P. de Looze, Age and time effects on  
655 implicit and explicit learning. *Exp. Aging Res.* **40**(4), 477-511 (2014).
- 656 50. M. Muylle, E. H. Smalle, R. Hartsuiker, Rapid phonotactic constraint learning in ageing:  
657 evidence from speech errors. *Lang. Cogn. Neurosci.* Advance Online Publication  
658 <https://doi.org/10.1080/23273798.2021.1897149> (2021).
- 659 51. S. Henin *et al.*, Learning hierarchical sequence representations across human cortex and  
660 hippocampus. *Sci. Adv.* **7**(8), eabc4530 (2021).
- 661 52. K. McNealy, J. C. Mazziotta, M. Dapretto, Cracking the language code: neural mechanisms  
662 underlying speech parsing. *J. Neurosci.* **26**(29), 7629-7639 (2006).
- 663 53. R. De Diego-Balaguer *et al.*, Striatal degeneration impairs language learning: evidence from  
664 Huntington's disease. *Brain* **131**(11), 2870-2881 (2008).
- 665 54. M. F. Assaneo *et al.*, Spontaneous synchronization to speech reveals neural mechanisms  
666 facilitating language learning. *Nat. Neurosci.* **22**(4), 627 (2019).
- 667 55. W. Beam, J. J. Borckardt, S. T. Reeves, M. S. George, An efficient and accurate new method  
668 for locating the F3 position for prefrontal TMS applications. *Brain Stimul.* **2**, 50-54 (2009).
- 669 56. A. Mir-Moghtadaei *et al.*, Concordance between BeamF3 and MRI-neuronavigated target  
670 sites for repetitive transcranial magnetic stimulation of the left dorsolateral prefrontal  
671 cortex. *Brain Stimul.* **8**, 965-973 (2015).
- 672 57. J. Jung, A. Bungert, R. Bowtell, S. R. Jackson. Vertex stimulation as a control site for  
673 transcranial magnetic stimulation: a concurrent TMS/fMRI study. *Brain Stimul.* **9**(1), 58-64  
674 (2016)
- 675 58. M. R. Goldsworthy, J. B. Pitcher, M. C. Ridding, A comparison of two different continuous  
676 theta burst stimulation paradigms applied to the human primary motor cortex. *Clin.*  
677 *Neurophysiol.* **123**, 2256-2263 (2012).
- 678 59. D. Bor, D. J. Schwartzman, A. B. Barrett, A. K. Seth, Theta-burst transcranial magnetic  
679 stimulation to the the prefrontal or parietal cortex does not impair metacognitive visual  
680 awareness. *PLoS One* **12**(2): e0171793 (2017).
- 681 60. A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG  
682 dynamics including independent component analysis. *J. Neurosci. Methods* **134**(1), 9-21  
683 (2004).
- 684 61. D. Bates, R. Kliegl, S. Vasishth, H. Baayen, Parsimonious mixed models. arXiv [preprint]  
685 (2015). <https://arxiv.org/abs/1506.04967>
- 686 62. H. Singmann, D. Kellen, "An introduction to mixed models for experimental psychology" in  
687 *New Methods in Cognitive Psychology*, D. H. Spieler, E. Schumacher, Eds. (Psychology  
688 Press, 2019, pp. 4-31 (2019).
- 689 63. H. Singmann *et al.*, *Afex: Analysis of factorial experiments*. R package version 0.13-145  
690 (2015).
- 691 64. R. V. Lenth *et al.* *Estimated Marginal Means, aka Least-Squares Means*. R package version  
692 1.6.2-1 (2021).
- 693

694

695 **Figures and Tables**

696

697 **Figure 1. Behavioural results reflecting statistical word-segmentation.** The upper panel  
698 shows total recognition accuracy (%) for the hidden words as a function of the cognitive disruption  
699 in Experiment 1 (left) and Experiment 2 (right). The lower panel shows recognition accuracy for  
700 the hidden words when participants indicated low and high confidence in remembering the hidden  
701 word, referred to as unconfident versus confident responses, respectively. Above-chance  
702 performance on confident responses reflects knowledge stored in explicit memory. Above-chance  
703 performance on unconfident responses reflects knowledge stored in implicit memory. Error bars  
704 denote standard errors of the mean. Asterisks denote significance for one-tailed t-tests: \* $p < 0.05$ ,  
705 \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

706

707 **Figure 2. EEG results reflecting online statistical learning: Inter-Trial Coherence (ITC)** (mean of 6  
708 centro-frontal midline electrodes, FC1, FC2, F3, F4, FC5, and Fz) during exposure to random and  
709 structured sequences in Experiment 1 (TMS) and Experiment 2 (Cognitive Load). Topographical  
710 plot shows distribution of ITC across the scalp, as a function of exposure and frequency. ITC  
711 values were used to calculate the word learning index (WLI) (i.e., ITC word / ITC syllable), see  
712 Figure 3.

713 **Figure 3. EEG results reflecting online statistical learning:** Changes in word learning index  
714 (WLI) as a function of exposure (Random vs. Structured sequences) with and without cognitive  
715 disruption (Experiment 1: TMS; Experiment 2: Cognitive load). Error bars denote standard errors  
716 of the mean.

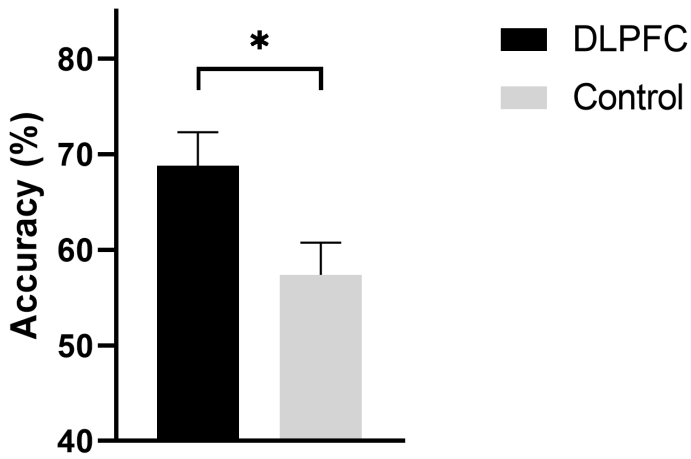
717 **Figure 4. Main procedure in Experiments 1 and 2:** In Experiment 1, participants received cTBS  
718 either over DLPFC or vertex prior to exposure. In Experiment 2, participants performed a dual  
719 working-memory task under high or low cognitive load conditions or did not perform a task prior to  
720 exposure. Exposure started with a random stream for 5 minutes, after which a structured stream  
721 was presented for 20 minutes. A short break (10 s) was inserted every ~5 minutes. EEG was  
722 recorded throughout. Participants watched a silent nature documentary during the exposure. After  
723 a 15-minute break, participants completed a post-exposure recognition test that examined implicit  
724 and explicit memory of the hidden words through memory judgement. During the exposure and  
725 the break participants were not aware that the behavioural test would include segments of the  
726 syllable sequences.

727 **Table 1.** Participant characteristics: *pre-tested* individual cognitive-control abilities of the different  
728 groups

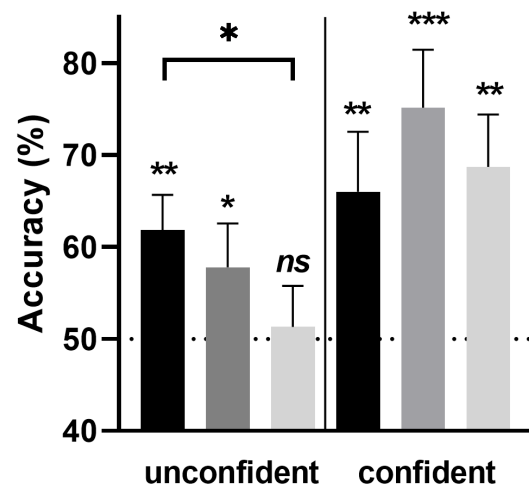
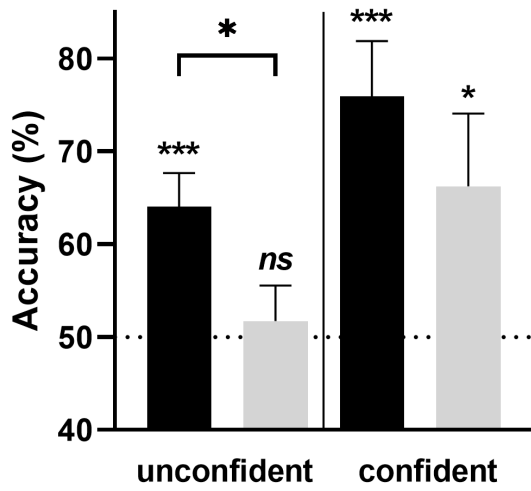
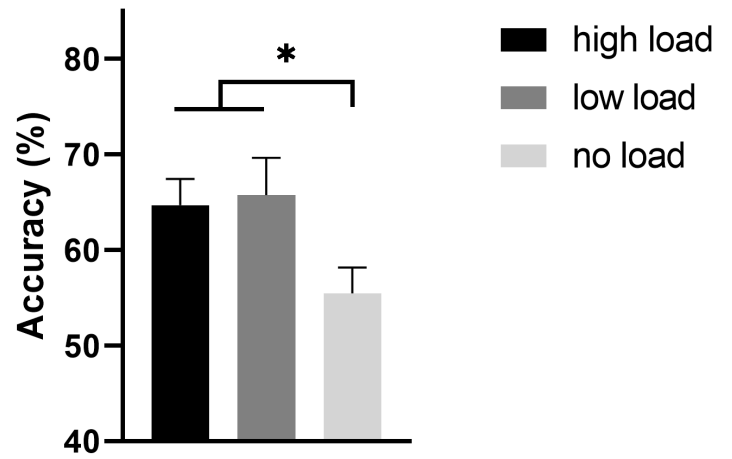
729



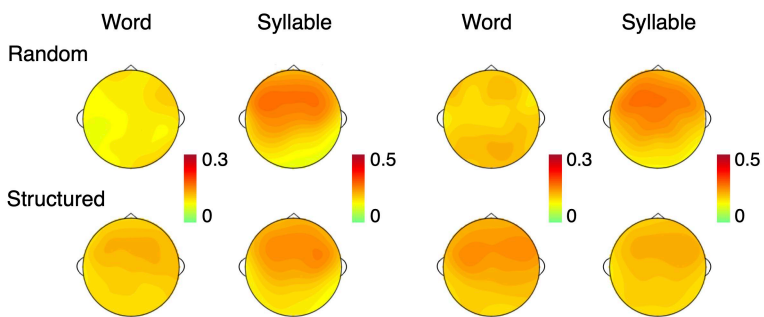
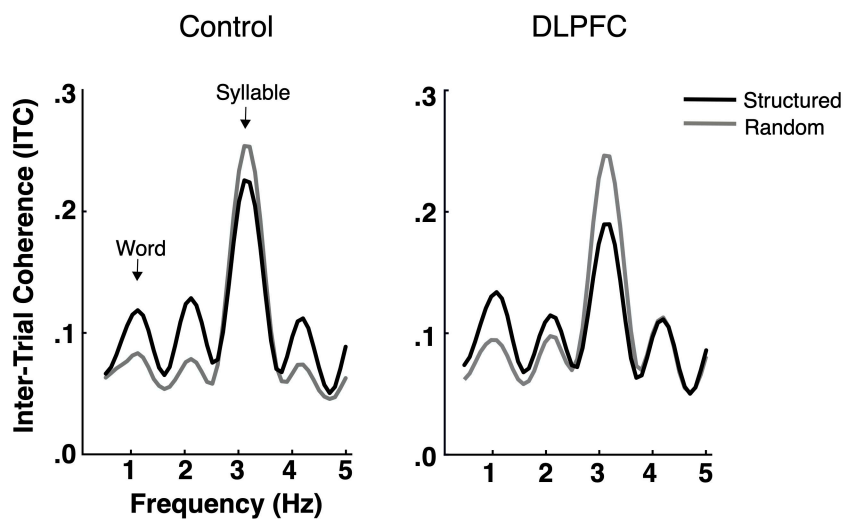
### Experiment 1



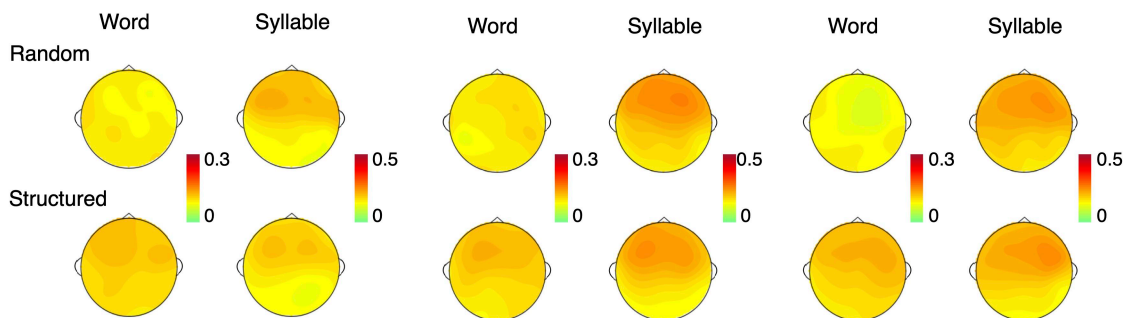
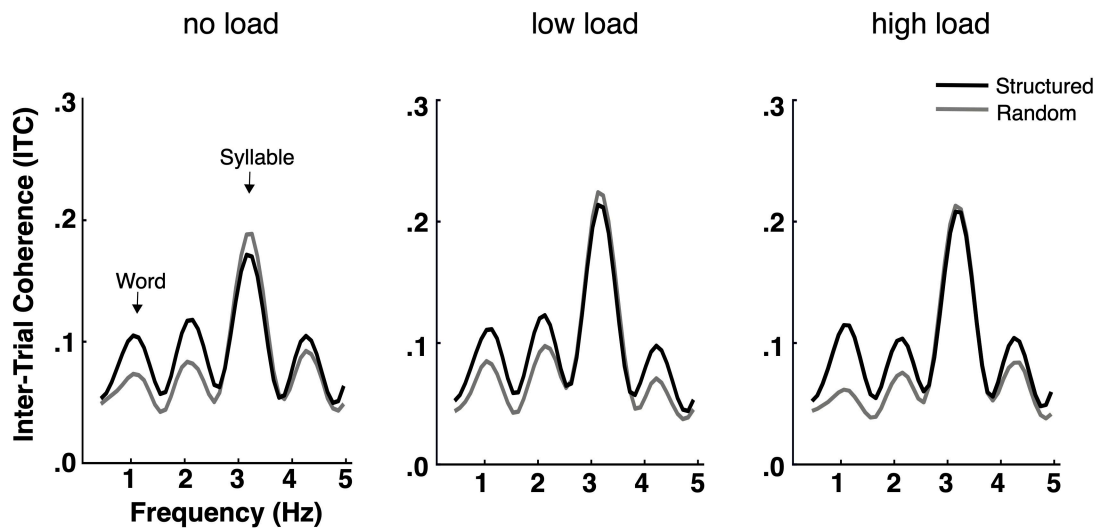
### Experiment 2



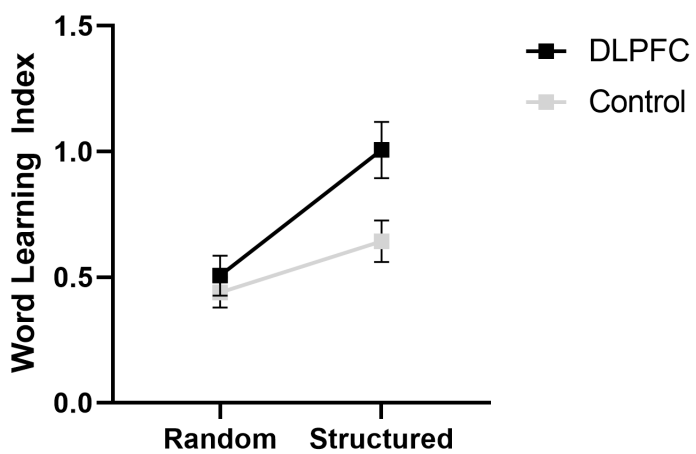
# Experiment 1



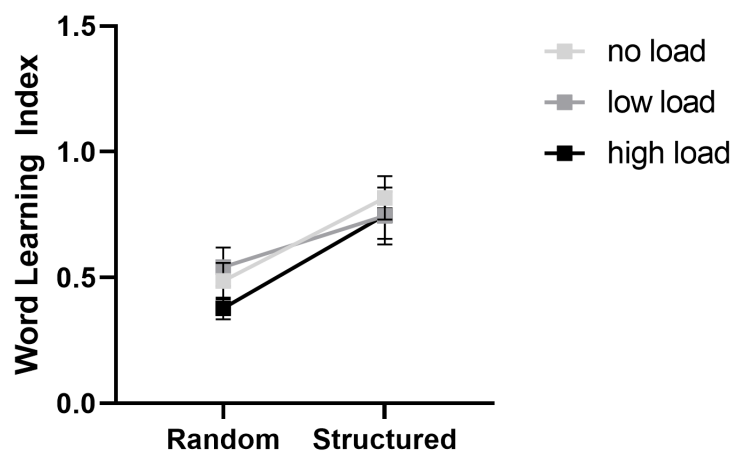
# Experiment 2



### Experiment 1



### Experiment 2



**Experiment 1:  
TMS over DLPFC or Control site**

**Experiment 2:  
High, Low or No Cognitive load task**

**Exposure (EEG)**

10-s break

betarogodikugo ...

tupirogolabubedaku

paditagolabubedaku

tupiropaditagolabu

paditatupirobedaku

random (5 min)

structured streams (4 x 5 min)

**15-min break**

**Recognition task and memory judgement**

tupiro

or

godatu

Remember/Familiar/Guess?

(word)

(non-word)

time

	CONTROL	DEPLETED	
<b>EXPERIMENT 1</b>	<b>N = 18</b>	<b>N = 18</b>	
<b>DIGIT SPAN (FORWARD + BACKWARD)</b>	18 (11-25)	19 (11-28)	
<b>WCST (N PERSEVERATION ERRORS)</b>	7 (5-17)	7 (5-11)	
<b>EXPERIMENT 2</b>	<b>N<sub>NCL</sub> = 20</b>	<b>N<sub>LCL</sub> = 20</b>	<b>N<sub>HCL</sub> = 20</b>
<b>DIGIT SPAN (FORWARD + BACKWARD)</b>	18 (10-22)	19 (12-24)	20 (12-29)
<b>WCST (N PERSEVERATION ERRORS)</b>	8 (5-19)	8 (5-17)	7 (5-11)
<b>STD</b>	.91 (.46-1.3)	.87 (.50-1.4)	.91 (.46-1.3)

WCST = Wisconsin Card Sorting Test for cognitive reasoning (Berg, 1948); STD = Stimulus Time Duration, i.e., maximum processing speed capacity on the dual working-memory task (Borrigan et al., 2017). All *ns* according to independent sample t-tests ( $ps > .05$ ).