

1

2

**Stage 1 Registered Report**

3

4

**Estimating the Effect of Reward on Sleep-  
Dependent Memory Consolidation – A  
Registered Report**

5

6

7

8 David P. Morgan<sup>1,2,3\*</sup>, Juliane Nagel<sup>1,2,3</sup>, N. Cagatay Gürsoy<sup>1,2,3</sup>, Simon Kern<sup>1,2,3</sup>, & Gordon B.  
9 Feld<sup>1,2,3,4\*</sup>

10 <sup>1</sup> Department of Clinical Psychology, Central Institute of Mental Health, Medical Faculty Mannheim,  
11 Heidelberg University, Mannheim, Germany

12 <sup>2</sup> Department of Addiction Behavior and Addiction Medicine, Central Institute of Mental Health,  
13 Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

14 <sup>3</sup> Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty  
15 Mannheim, Heidelberg University, Mannheim, Germany

16 <sup>4</sup> Department of Psychology, Heidelberg University, Heidelberg, Germany

17

18 \* Corresponding authors:

19 David P. Morgan ([david.morgan@zi-mannheim.de](mailto:david.morgan@zi-mannheim.de))

20 Gordon B. Feld ([gordon.feld@zi-mannheim.de](mailto:gordon.feld@zi-mannheim.de))

21

**Abstract**

Rewards play an important role in guiding which memories are formed. Dopamine has been shown to be an important neuromodulator mediating the effect of rewards on memory. In rodents dopaminergic activity during learning has been shown to enhance reactivation of memory traces during sleep, the mechanism driving the benefits of sleep on consolidation. However, evidence that sleep consolidates high reward memories more strongly in humans is mixed and small samples sizes (among other factors) likely drive these inconsistencies. Therefore, we will compare memory for rewarded information between intervals of sleep and wake in a large representative online sample. Participants ( $N = 1750$ ; stratified German sample) will study images associated with high and low rewards and complete a memory test directly afterwards as well as after retention. Our main prediction is that sleep will enhance the retention of high over low reward images compared to wake. In general, we also expect sleep to enhance retention (evident through a reduced decrease in performance compared to wake) and rewards to improve memory. This study will reveal whether sleep facilitates selective consolidation or whether processes at encoding and shortly thereafter suffice. Additionally, it will provide a benchmark effect size to evaluate sleep-based interventions for psychiatric disorders (e.g., addiction). It will also allow us to explore moderators of the effect, such as age and education level.

*Keywords:* Sleep, reward, memory consolidation

*Word count:* 13969

42

43

44 **Introduction**

45 An accumulation of evidence indicates that sleep actively supports the stabilization and  
46 transformation of long-term memory<sup>1-3</sup> and for the most part studies have demonstrated that  
47 sleep compared to wakefulness benefits memory across declarative and procedural tasks<sup>e.g., 4,5-  
48 15, but see 16,17-19</sup>. The preferred explanation for the benefits of sleep on long term memory are  
49 attributed to active systems consolidation, but alternative explanations for the impact of sleep  
50 on memory do exist (e.g., Passive Interference Reduction Hypothesis<sup>20</sup>, Opportunistic  
51 Consolidation<sup>21</sup>). The active systems consolidation hypothesis posits that the associative  
52 connections between elements of new information are encoded by the hippocampus and over  
53 time these connections are redistributed to the neo-cortex via systems consolidation<sup>22</sup>. This  
54 redistribution of information is thought to preferentially occur during sleep, whereby memory  
55 traces that were encoded throughout prior wakefulness are replayed repeatedly and thereby  
56 strengthened, although it should be noted that replay also occurs during wakefulness<sup>23,24</sup>.  
57 During active systems consolidation, sleep specific brain activity and especially the activity of  
58 hallmark oscillations (slow oscillations, hippocampal ripples and sleep spindles) that putatively  
59 coordinate this replay are thought to drive greater memory performance in those tasks<sup>see 25,26-  
60 27,28,29, but also see 30,31</sup>. The limited availability of these reactivation opportunities during sleep<sup>32,33</sup>  
61 suggests the selective consolidation of only relevant information, e.g., rewarded information<sup>2</sup>.  
62 However, it has not yet conclusively been shown that memories associated with a reward are  
63 consolidated more strongly during sleep.

64 Reward plays an important role in memory<sup>34-42, for a review see 43</sup>. In the declarative domain,  
65 its role has been demonstrated in humans using the motivated learning task. In that task, stimuli  
66 associated with a high or low reward are presented to participants and corresponding rewards  
67 are paid out for subsequent successful retrieval<sup>34</sup>. Researchers have consistently shown that

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

68 memory for items associated with higher reward is greater than for those associated with lower  
69 rewards [in humans](#)<sup>34-39</sup>. Often such studies include a period of sleep, implicating the role of  
70 sleep in consolidation of reward associated memories<sup>35,41,42, 44</sup> and consolidation of highly  
71 rewarded information has been linked to sleep spindle activity<sup>38,42</sup>. This link to spindle activity  
72 during sleep suggests that sleep and reward fundamentally interact to consolidate  
73 motivationally relevant information indicating that reward plays a crucial role even long after  
74 encoding has taken place. However, the precise mechanism and time-frame by which sleep  
75 benefits reward memories remains ambiguous.

76 At encoding, dopamine modulates memory performance by recruiting reward areas in  
77 a ventral-striatum-ventral-tegmental-area-hippocampus feedback loop.<sup>45</sup> Using the Motivated  
78 Learning [Task \(in humans\)](#) a landmark study demonstrated that high reward cues activated the  
79 nucleus accumbens (located in the ventral striatum), the ventral tegmental area and the  
80 hippocampus during encoding<sup>34</sup>. Hippocampus activity was functionally coupled with activity  
81 in the ventral striatum and this predicted subsequent memory performance for high reward  
82 items. Behaviorally, this effect manifested as greater memory performance for high vs. low  
83 rewards at high levels of confidence. Regarding sleep, there is no consensus whether sleep  
84 enhances rewarded memories through additional dopaminergic neuromodulation during  
85 reactivation<sup>41,45,46</sup> or rather dopamine sets a tag during learning that leads to enhanced  
86 reactivation without additional dopaminergic neuromodulation<sup>40</sup>. [Before answering this, it is](#)  
87 [first necessary to establish behaviorally whether or not sleep preferentially consolidates highly](#)  
88 [rewarded memories over lowly rewarded memories. Only then can the underlying neuronal](#)  
89 [mechanisms be characterized.](#)

90 Independent of the putative underlying neurophysiological mechanisms, [in humans](#),  
91 evidence is inconclusive, overall, regarding sleep's role for rewarded memories. Several

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

92 studies did not find that rewards enhance sleep-dependent memory consolidation.<sup>10,47,48</sup> In one  
93 experiment, participants were asked to learn object locations associated with high or low  
94 reward and were tested on those locations after a nap or a period of wakeful rest.<sup>47</sup> No  
95 difference in the magnitude of memory for high and low rewards was found between the  
96 napping and wakeful conditions. The absence of this effect is not uncommon and even extends  
97 to comparisons of a full night of sleep with typical daytime wakefulness and across recognition  
98 memory and verbal free recall tasks<sup>10,48</sup>. Nevertheless, sleep was still found to benefit memory  
99 overall.

100 This conflicts with another study using a procedural finger sequence tapping task.<sup>49</sup> In  
101 that study sleep preferentially consolidated highly rewarded sequences relative to a period of  
102 wakefulness. That finding was corroborated by another a study using a recognition memory  
103 task where a retention interval including a nap yielded greater memory for highly rewarded  
104 items compared to lowly rewarded items and this difference was not present in an equivalent  
105 wake condition<sup>50</sup>. However, in the latter experiment, there was no significant interaction  
106 between those groups, which despite the authors' conclusions would be necessary to conclude  
107 that high vs low reward items are preferentially consolidated during sleep<sup>51</sup>. One study found  
108 that the benefits of sleep on reward compared to wake may only unfold after much longer  
109 periods, which could allow further consolidation processes to take place<sup>44</sup>.

110 Mutually exclusive theoretical conclusions from these studies can be drawn by ignoring  
111 the respective evidence that is not in their favor. Either sleep selectively consolidates  
112 information associated with high rewards<sup>2</sup> or reward related processes during encoding  
113 together with sleep-independent consolidation processes initiated shortly after learning are  
114 sufficient to enhance reward memory.<sup>21</sup> **A third possibility is that consolidation does not affect**  
115 **reward related differences in memory performance and the difference are only due to encoding**

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

116 processes. Like for other memories it is evident that sleep is involved in the consolidation of  
117 rewarded memories per se, yet it is unclear whether sleep specifically enhances differences in  
118 memory performance based on reward amplitude (e.g., high vs. low rewarded information).

119 On the one hand, one could attempt to explain the divergent findings by evaluating the  
120 large number of differences in experimental designs (e.g., recognition vs recall, images vs  
121 words, napping vs 12-hours of sleep etc.). Here, one would conclude that the enhancement of  
122 sleep's beneficial effect by reward is sensitive to a host of moderators, as has been discussed  
123 for other inconsistencies in the field. For instance, mode of retrieval (e.g., free recall  
124 vs. recognition), mode of learning (e.g., implicit vs. explicit), material learned (e.g., declarative  
125 vs. procedural) and the timing of sleep (e.g., delay between learning and sleep onset) are all  
126 thought to moderate the sleep effect<sup>52</sup>. Such views have recently been reiterated in an  
127 assessment of the robustness of the sleep effect on memory<sup>53</sup>. However, this explanation leads  
128 to the unsatisfactory conclusion that the enhancement of sleep's beneficial effect on memory  
129 by reward is sensitive to moderators that were not systematically controlled in many of these  
130 studies. On the other hand, there exists a striking similarity between all of these experiments:  
131 low statistical power (maximum n = 20 per group)<sup>e.g., 10, 30, 40-42,44,47,49,50</sup>. Small samples have  
132 been shown to reduce generalizability, increase false negatives as well as false positives and  
133 can overestimate effect sizes<sup>54,55</sup>, which may be the source of divergence. An argument that  
134 was sympathetically pointed out in a recent systematic review of the sleep and reward memory  
135 literature<sup>56</sup>.

136 Our study will address this divergence by performing a large-scale investigation of the  
137 influence of rewards on sleep-dependent memory consolidation in the general population and  
138 asks the question: do rewards affect the magnitude of sleep-dependent memory consolidation?  
139 It is highly relevant to understand the impact of sleep on rewarded information since it guides

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

140 (mal-) adaptive behavior such as unhealthy eating, smoking or alcohol consumption. Reward  
141 related learning mechanisms and other dopamine related plastic changes in the brain have been  
142 proposed to play a crucial role for establishing addictive behavior<sup>57</sup>. However, it remains  
143 unclear whether sleep-dependent consolidation of drug taking experiences occurs. Showing  
144 that sleep has a unique and sizable role for preferentially consolidating rewarded memory in  
145 the general population may fuel systematic investigations and targeted sleep interventions to  
146 better understand and treat, e.g., substance abuse and anxiety disorders. One such intervention  
147 may make use of the targeted memory reactivation procedure<sup>58</sup>, where cues are used to  
148 reactivate memories during sleep. In some scenarios cueing during sleep has been shown to  
149 extinguish conditioned fear responses<sup>59</sup> and therefore extinguishing addictive behavior during  
150 sleep by using appropriate cues may be promising.

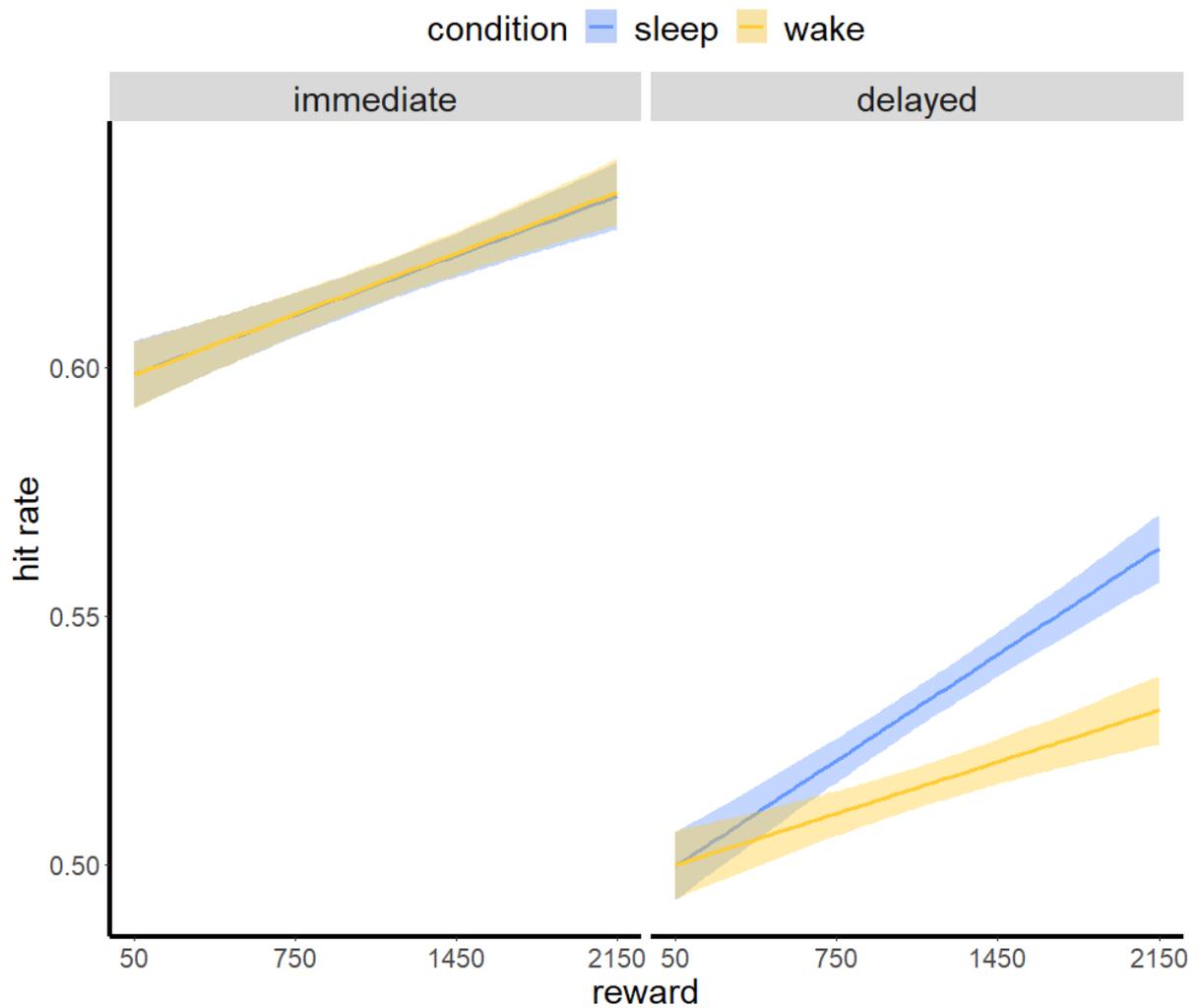
151 We will implement the AM:PM-PM:AM design in an online testing environment to  
152 collect a large sample of representative participants effectively, a strategy that has been  
153 successfully used by us before in a previous large-scale registered report in sleep and memory  
154 research<sup>17</sup>. In the AM:PM PM:AM design, participants undergo a wake condition, where the  
155 learning phase occurs in the morning (AM) and the test phase occurs in the evening (PM) on  
156 the same day. Participants also undergo a sleep condition, where the learning phase occurs in  
157 the evening (PM) and the test phase occurs the following morning (AM). In recent years  
158 researchers investigating the impact of sleep on memory have begun using web-based  
159 alternatives by performing online sleep experiments<sup>60,61</sup>. It should be noted that generally such  
160 experiments do not appear to limit the capacity to detect the impact of sleep on memory.  
161 Reward memory will be measured using a paradigm adapted from earlier studies<sup>34,40,41</sup> and  
162 recently validated in our laboratory to yield positive effects of reward on memory performance  
163 (see supplementary material: [https://cloud.zi-  
mannheim.de/index.php/s/jDnY35CM4WMdQCg](https://cloud.zi-<br/>mannheim.de/index.php/s/jDnY35CM4WMdQCg)), where participants ( $N = 1750$ ) will study

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

165 images associated with high to low rewards and will retain them across sleep and wakefulness.  
166 This paradigm uses a recognition task to measure memory performance and although  
167 recognition tasks have been shown to be somewhat less sensitive to the effect of sleep on  
168 memory than free or cued recall procedures<sup>62</sup> our power analysis indicates that we have  
169 sufficient power.

170 We predict (see Figure 1 and Table 1), **H1) that sleep will yield greater retention**  
171 **compared to an equivalent period of wake (although we expect a general decline in**  
172 **performance across retention); H2) that items associated with high rewards will be better**  
173 **retained compared to those associated with low rewards; H3) the magnitude of the decline**  
174 **of high reward memories will be less in the sleep condition compared to the wake**  
175 **condition.** In addition, to these three main hypotheses our study will include several control  
176 variables to investigate known confounding factors (i.e., vigilance, sleepiness, general retrieval  
177 performance, memory strength and task difficulty) as well as variables that will allow us to  
178 explore moderating factors (i.e., age, education status, morningness-eveningness, mental  
179 health, shift work, travel and medication). Of note, our study will not be able to show how sleep  
180 parameters mechanistically affect reward memory, as the sleep vs. wake design cannot reveal  
181 such relationships (irrespective of whether sleep deprivation or as in our case day wakefulness  
182 is being used). However, our study will enable the planning of much more resource intensive  
183 mechanistic studies that manipulate sleep (e.g., by drugs) by delivering an effects size estimate  
184 with much less uncertainty than previously.

185



186 *Figure 1.* Visualization of the first simulated run of our predictions produced by our data  
187 generating model, with a sample size of  $N = 1750$ . The estimated memory performance for  
188 each reward category is represented by the thick lines and shaded areas represent standard  
189 error estimated using linear models. Note that the standard error is small due to the large  
190 sample size.

REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

191 Table 1. Design table.

Sampling plan:		As per our data simulation described on page 29 a maximum of $N = 1750$ participants will be collected to test the hypotheses described below.			
Question	Hypothesis	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation given different outcomes	Theory that could be shown wrong by the outcomes
Does sleep yield greater memory performance compared to an equivalent period of wake?	<b>H1:</b> In the sleep condition, hit rate for the Motivated Learning Task at delayed recognition will be greater when compared to the wake condition.	Linear mixed model: The interaction between retention x time point will be significant.  <i>Planned comparisons:</i> the main effect of retention at delayed recognition will be significant.  Equivalence: Our smallest effect size of interest for this comparison is in the same range as the reward effect or the reward X retention X timepoint interaction (both unstandardised estimates of .015; see Table 5). This is why we assume equivalence if the 90% confidence interval around the parameter estimated from the actual data does not include .015.	In our model the effect of retention is 0 because sleep only exerts its influence at delayed recognition for high rewards only. However, it is still plausible that an overall difference between sleep and wake emerges at delayed recognition (e.g., this depends on how motivating the low rewards are). For our smallest effect size of interest we chose a value that is comparable with our smallest effect size of interest for the reward effect and the reward X retention X timepoint interaction.	Significance: If the main effect of retention at delayed recognition is significant and the mean memory performance for the sleep condition is greater compared to the wake condition, sleep positively affects memory in comparison to an equivalent period of wake. This would confirm <b>H1</b> .  Equivalence: If the difference between sleep and wake conditions is found to be statistically equivalent this would disconfirm <b>H1</b> .  If the difference between sleep and wake conditions is found to be equivalent, H3 described below will still be tested as alternative interpretations may exist as described in our supplemental information, which can demonstrate the preferential impact of sleep on consolidation of high vs. low rewards.	If there is no difference in memory performance between the sleep and wake groups at delayed recognition this could demonstrate that the theory that sleep benefits memory generally over and above a period of wake is incorrect.  Alternatively, if there is a difference in memory performance between the sleep and wake groups at delayed testing and the wake condition yields better memory performance than the sleep condition then this would point towards periods of wakefulness being more beneficial for memory retention as compared to sleep."
Does information associated with high rewards yield greater memory performance compared to low rewards?	<b>H2:</b> The hit rate for high rewards will be greater than the hit rate for low rewards at delayed recognition	Linear mixed model: The interaction between reward x time point will be significant.  <i>Planned comparisons:</i> the main effect of reward at immediate and delayed testing will be significant  Equivalence: If the 90% confidence interval around the parameter estimated from the actual data does not include the parameter we set in our simulations (i.e. .015; see Table 5) we assume equivalence. This is because our simulations are based on our minimum effect sizes of interest.	The power analysis derived from our data simulation indicated that we have 95% power to detect at least an unstandardized effect size of .015 with an alpha of $p < .020$ . This effect size was selected on the basis of our pilot data for the motivated learning task presented in the supplemental material.	Significance: If the main effect of reward at immediate and delayed recognition is significant and the mean memory performance for the high rewards is greater compared to low rewards, high rewards have a greater impact on positively affecting memory in comparison to low rewards. This would confirm <b>H2</b> .  Equivalence: If the difference between high rewards and low rewards is found to be statistically equivalent this would disconfirm <b>H2</b> .	If there is no difference in memory performance between low rewards and high rewards this could demonstrate that the theory that high rewards are beneficial for memory is incorrect.
Does sleep yield greater recognition memory performance for high vs. low reward items?	<b>H3:</b> The magnitude of the positive effect of reward on the hit rate will be greater in the sleep condition compared to the wake condition at delayed recognition.	Linear mixed model: The interaction between retention x time point x reward will be significant.  <i>Planned comparisons:</i> The interaction between retention x reward at delayed recognition will be significant.  Equivalence: If the 90% confidence interval around the parameter estimated from the actual data does not include the parameter we set in our simulations (i.e. .015; see Table 5) we assume equivalence. This is because our simulations are based on our minimum effect sizes of interest.	The power analysis derived from our data simulation indicated that we have 95% power to detect at least an unstandardized effect size of .015 with an alpha of $p < .020$ . This effect size was selected due to the unknown nature of the size of the interaction between retention and reward and resource constraints.	Significance: if the interaction between retention x time point x reward and the interaction between retention and reward at delayed recognition is significant such that the magnitude of the benefit of reward on memory is greater in the sleep condition compared to the wake condition at delayed testing this would confirm <b>H3</b> .  Equivalence: If the difference in the magnitude of the effect of reward on memory between sleep and wake conditions at delayed recognition is found to be statistically equivalent this would disconfirm <b>H2</b> .	If there is not a greater positive effect of reward on memory performance in the sleep condition compared to the wake condition at delayed recognition then the theory that sleep preferentially consolidates reward information could be wrong. This would indicate that rewards exert their influence on memory predominantly during encoding.

192 **Note:** The factor retention refers to the retention manipulation and contains the two levels sleep and wake. Also, since we are using a declarative task, we cannot generalize our inferences to the procedural domain and declarative  
 193 memory is meant whenever we write memory in this table

## 194 **Methods**

195

### 196 **Participants**

197 The size of our sample is guided by resource constraints as well as a data simulation based on  
198 the data shown in Figure 1 and previous literature<sup>34,40,44,49,50</sup>. Our predictions indicated that  
199 1750 participants suffice to detect a very small effect size and a broad range of much larger  
200 effect sizes to achieve  $1 - \beta = 0.95$ . Of note, the simulation uses a data generating linear mixed  
201 model with specific input parameters shown in our analysis plan. Proportions of our  
202 representative sample stratified across sex (male and female), age (ages 20-29 to 50-59 years),  
203 highest professional qualification and highest school level qualification were calculated based  
204 on the German 2011 Census (See Figure 2). Sampling of strata will be ended individually as  
205 soon as they are full.

206 Participants will take part in this experiment online and will be recruited using targeted  
207 online advertisements on popular social media websites (e.g., Facebook, twitter) and media  
208 outlets (e.g., news websites). We will use Meta Advertisements, an advertisement service using  
209 Facebook and Instagram to target strata that we identify as currently under sampled. We will  
210 also use our contacts writing for national news outlets to further boost the visibility of the study.  
211 We will additionally implement a “refer a friend” strategy where participants can refer one or  
212 more friends. If at least one friend then goes on to complete the procedure the referrer will  
213 automatically receive a 5€ Amazon voucher. Participants will not receive compensation for  
214 their participation, but will have the chance to win a voucher dependent on their performance  
215 in the task. The voucher values will be 500 x 7.50€, 150 x 15€, 125 x 20€ and 100 x 25€ (adding  
216 up to 11000 € in vouchers). The average amount of the vouchers is thus 6.28 €, which is  
217 approximately the average bonus that we paid out in our pilot study which was the basis of our  
218 power calculation (see supplementary material; <https://cloud.zi->

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

219 [mannheim.de/index.php/s/jDnY35CM4WMdQCg](http://mannheim.de/index.php/s/jDnY35CM4WMdQCg)). Participants will be informed that each  
220 point (gem) earned in the reward task will improve their chances of winning a high value  
221 voucher, but that this also depends on the performance of the other participants. To minimize  
222 attrition participants who do not complete the whole experiment will not have the opportunity  
223 to win a voucher. The German Psychological Society (DGPs) ethics committee approved this  
224 experiment. Written informed consent will be obtained from participants prior to participating  
225 in this experiment. Since we propose to collect a large stratified sample across multiple  
226 sessions, we estimate that data collection will be completed within 12 months.

227 Our inclusion and exclusion criteria are presented below in Table 2. Participants who  
228 meet the exclusion criteria will not be included in the data analysis and will be resampled  
229 until our desired sample size is achieved. We have chosen not to exclude participants with  
230 mental health conditions which can impact participants' memory consolidation. This is  
231 because based on previous experience conducting large-scale online sleep experiments, such  
232 exclusion criteria can cause severe limitations on the recruitment process, since mental health  
233 issues are quite wide spread (i.e., one in three women and about one in four men aged 18–79  
234 in Germany meets diagnostic criteria of at least one mental disorder during the past  
235 12 months<sup>63</sup>). Additionally, a main goal of this research is to yield a demographically diverse  
236 (representative) sample, which can be used to derive an effect size estimate of the impact of  
237 sleep on reward memory, to be used in therapeutic settings. Therefore, the effect size must be  
238 as generalizable as possible beyond the samples typically used in sleep and memory  
239 experiments which are largely performed with highly educated young students. Such samples  
240 create a translational gap between basic science and clinical research which limits the  
241 generalization of our findings to samples with mental health conditions.

242

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

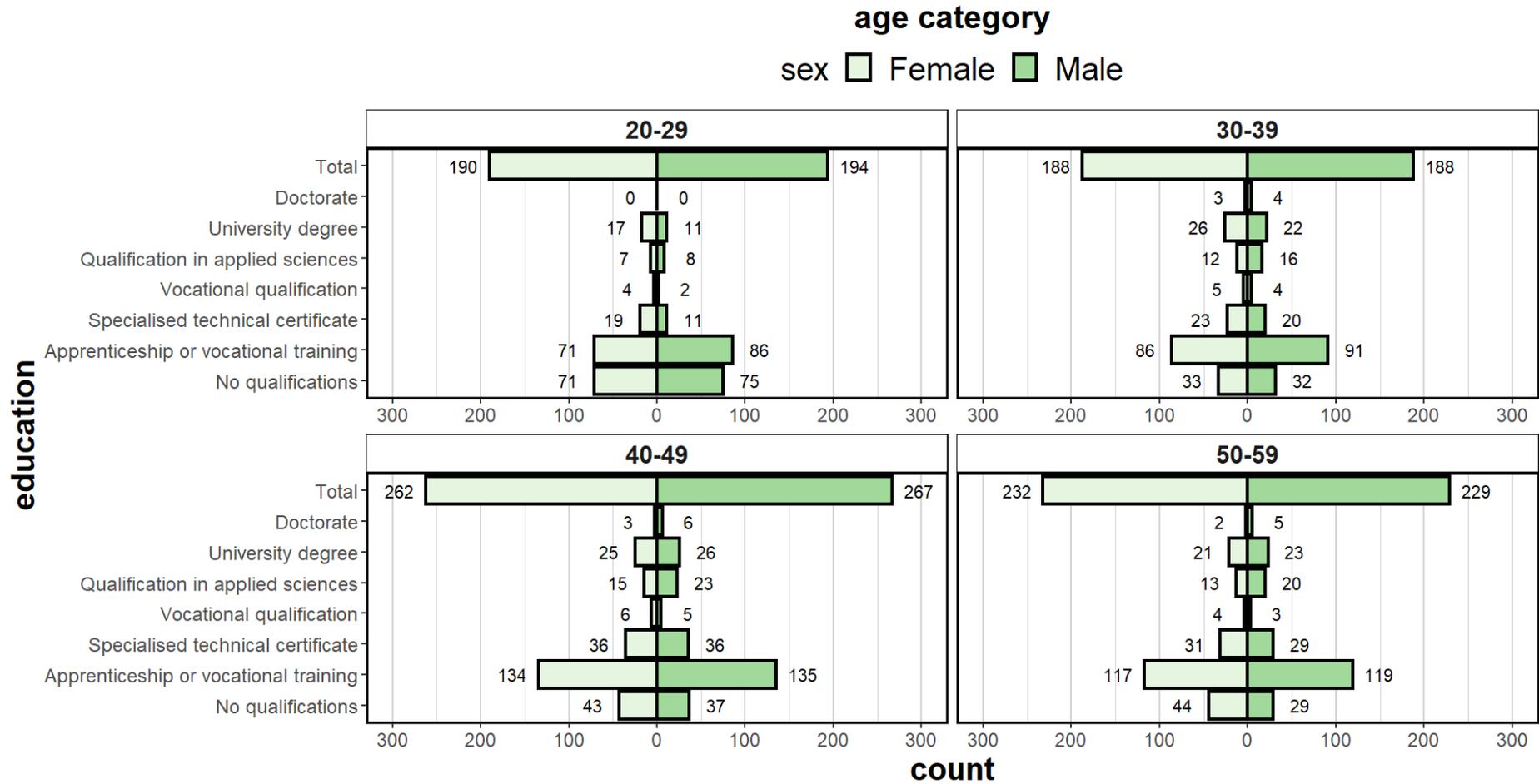
243 Table 2. Inclusion criteria necessary to participate in the experiment and exclusion criteria to be  
 244 included in the data analysis.

Inclusion Criteria	Exclusion criteria
Stratification: Sex, Male or Female Aged 20-59 Highest professional qualification Highest school leaving qualification Resident in Germany	Napping between study and test in the wake condition  Sleeping less than 6 hours in the sleep condition  Consumption of alcohol between study and test within sessions 1 or 2  Participants who respond too slowly on the Flankers task on 3 consecutive trials  Participants who respond too slowly on 3 consecutive trials for the recognition memory test  Participants who fail the validation questions on any occasion after their second attempt  A $d'$ score $\pm$ 3 SD away from the mean within each age category collapsed across timepoint (immediate vs. delayed), retention (sleep vs. wake), rewards and durations

245

246 To ensure completion of the sample, we will implement the following contingencies  
 247 incrementally: 1) If after 7 months of data collection we have not achieved at least 50%  
 248 of our desired sample size we will collapse the strata of the “highest professional  
 249 qualification” and “highest school-leaving qualification” categories into three groups  
 250 respectively; 2) If after 9 months of data collection we have not achieved at least 50% of our  
 251 desired sample we will remove the aforementioned education strata; and 3) Finally, if after 11  
 252 months of data collection we have not achieved at least 50% of our desired sample we will  
 253 open up data collection to the UK and USA (English versions of all materials already exist in  
 254 the lab). In each scenario the stratification will be adjusted.

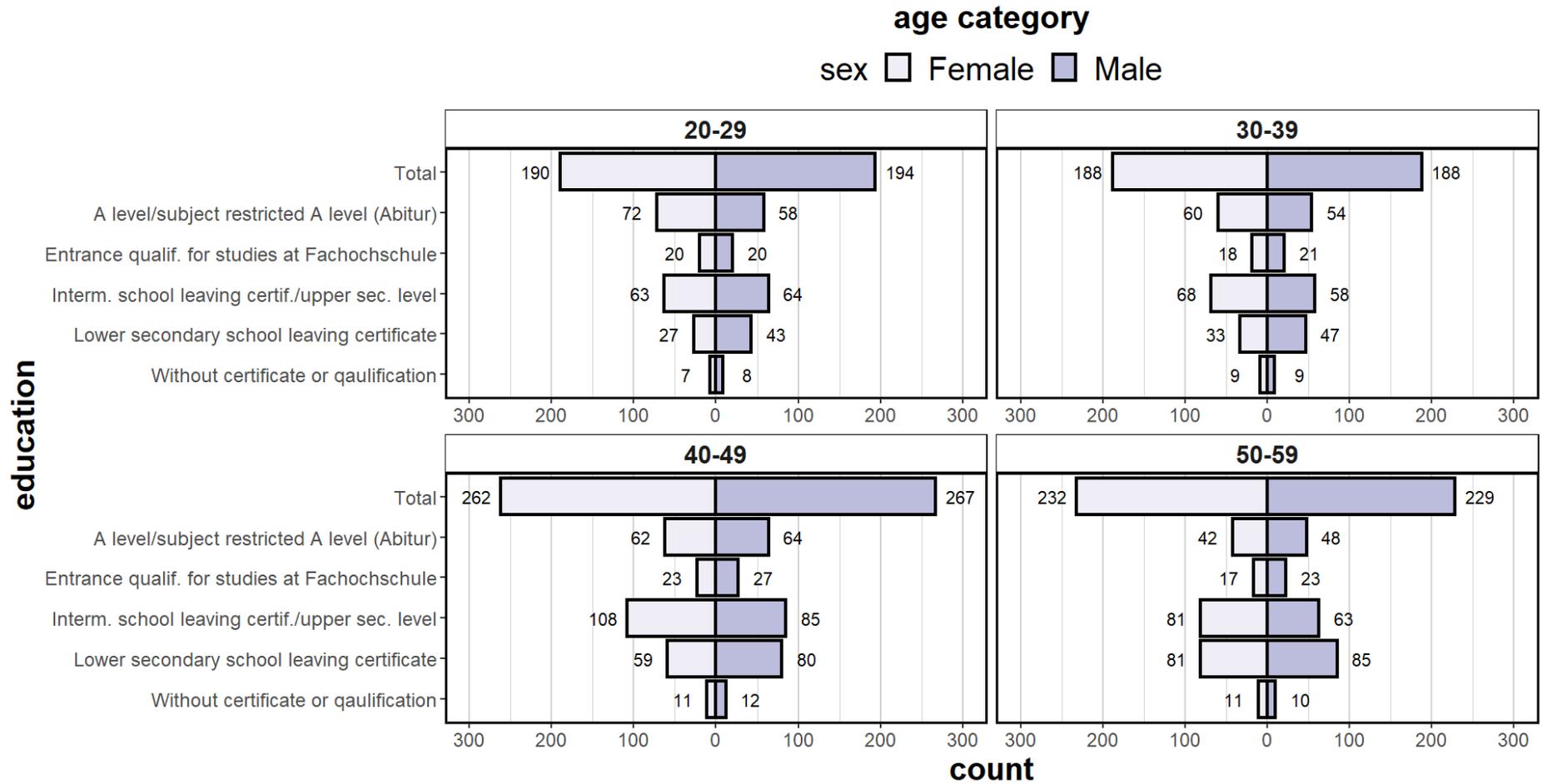
REWARD AND SLEEP-BASED MEMORY CONSOLIDATION



255

256 *Figure 2.* Demographic profile of the proposed stratified sample. For simplicity the strata are presented by highest professional qualification and  
 257 highest school level qualification. Nevertheless the final strata will consist of each stratified combination of the relevant categories, see  
 258 Supplemental Material. A) y-axes indicate the highest professional qualification attained and x-axes indicate the number of participants required  
 259 within each sex, age and professional qualification category combination. The precise number of participants that are required to yield a  
 260 representative sample are labeled against each bar for each sex, age and education category.

REWARD AND SLEEP-BASED MEMORY CONSOLIDATION



261 B) y-axes indicate the highest school level qualification attained and x-axes indicate the number of participants required within each sex, age and  
 262 school level qualification category combination. Again, the precise number of participants that are required to yield a representative sample are  
 263 labeled against each bar for each sex, age and education category.<sup>1</sup>  
 264

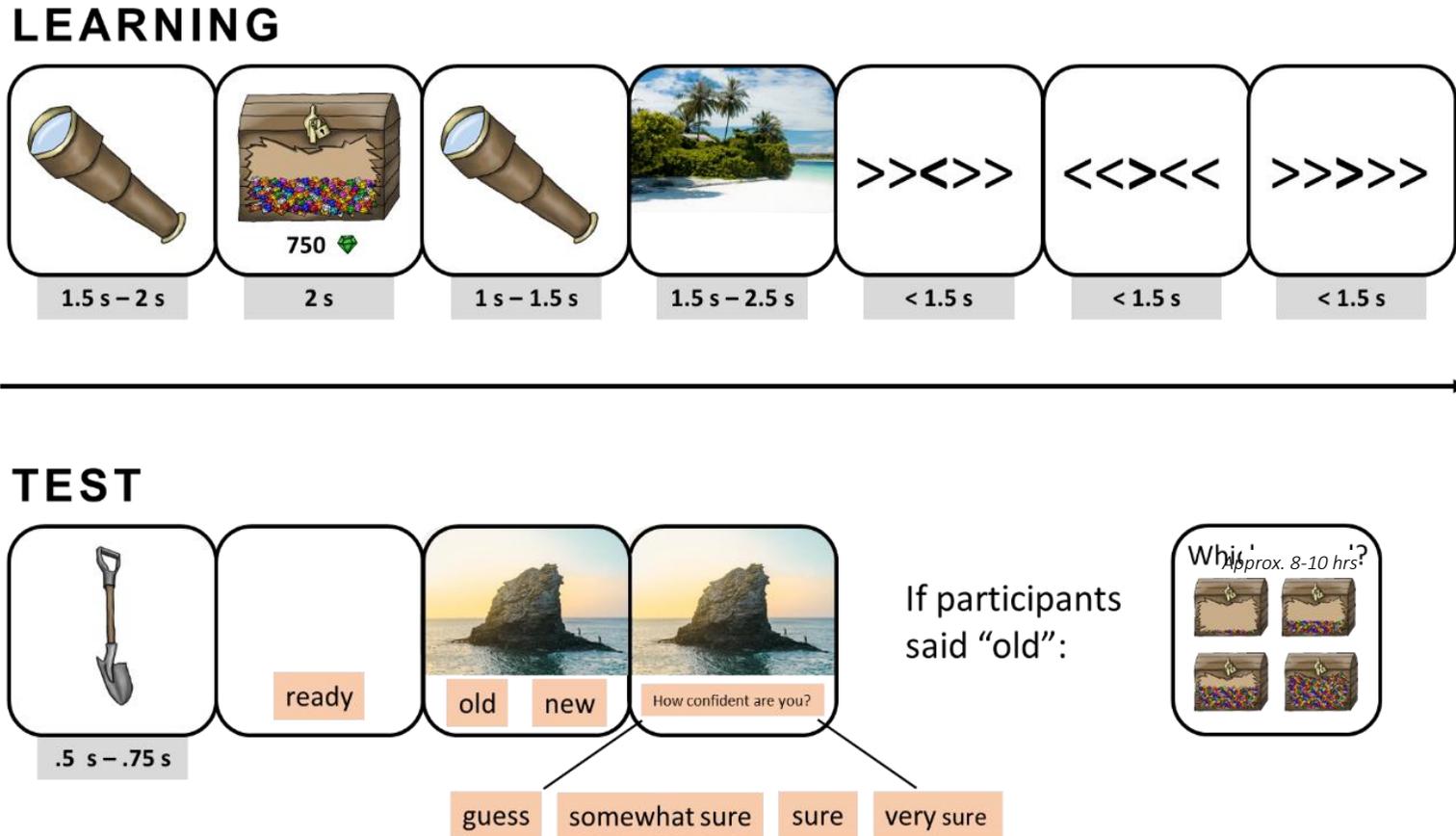
<sup>1</sup> The precise definitions of the education for highest professional qualification and highest school level qualification categories can be found at: <https://shorturl.at/lpz58>

265

266 **Design**

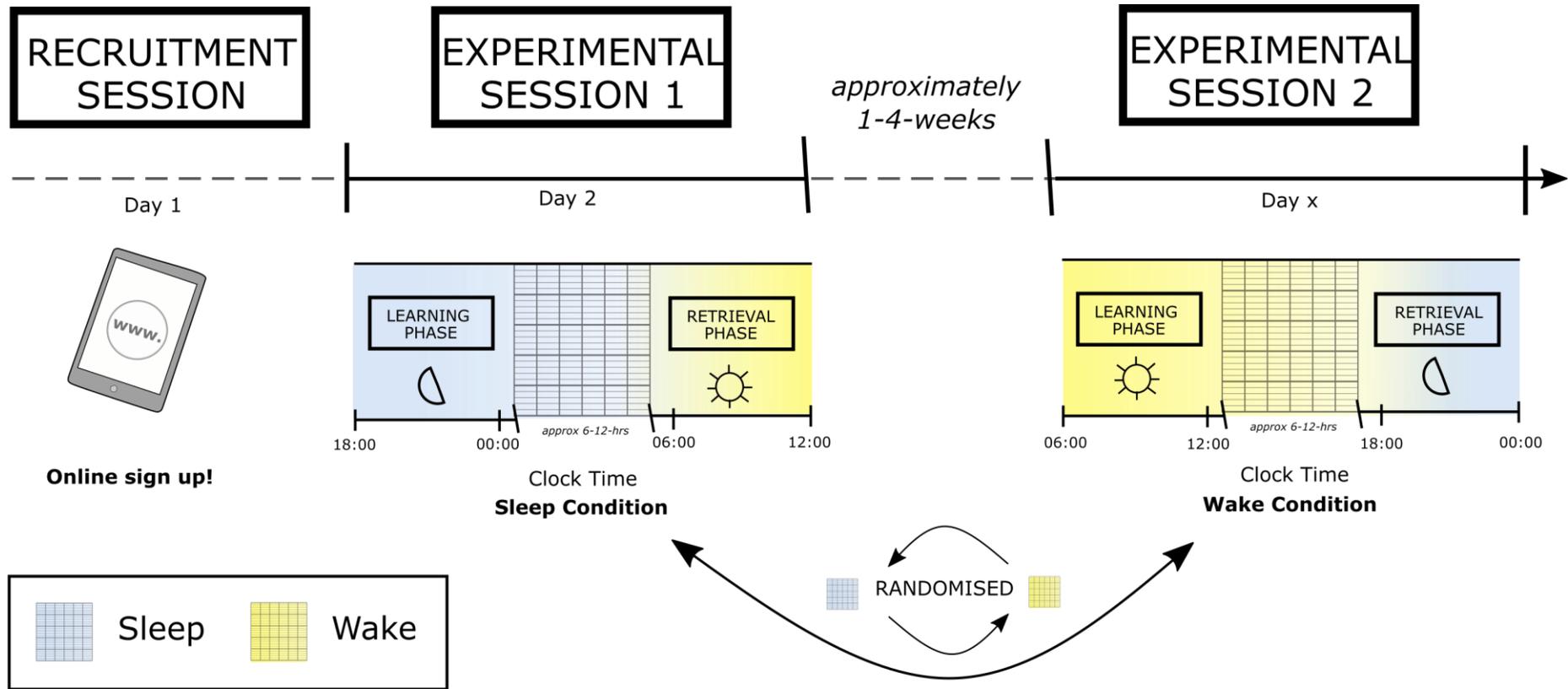
267 Participants will undergo the Motivated Learning task (see Figure 3) over two experimental  
268 sessions in a balanced AM-PM PM-AM cross-over design (see Figure 4), once with a retention  
269 interval of wake and a second time with a retention interval of sleep. From the introduction it  
270 is clear that a choice must be made to either assess memory using a procedural or a declarative  
271 task, which both have been shown to benefit from sleep in the retention interval<sup>62,64,65</sup>. We have  
272 chosen the former as in the literature there is no clear indication that a procedural task is better  
273 suited.

274 The Motivated Learning task is a recognition memory procedure and in our experiment  
275 memory is tested twice in both sleep and wake conditions, once after learning (to examine  
276 baseline performance) and again after sleep or wake. In the AM-PM PM-AM design when  
277 participants undergo the sleep condition they study images, complete an immediate recognition  
278 memory test in the evening (PM) and are tested once again the following morning (AM).  
279 Participants in the wake condition study images and complete an immediate recognition test in  
280 the morning and are subsequently tested again in the evening (PM) on the same day. Therefore,  
281 the experimental design has two within-subjects factors Retention (sleep; wake) and Time point  
282 (immediate; delayed) with two levels in each. The images that participants study are associated  
283 with rewards of four different magnitudes adding an additional within-subject factor reward  
284 with four levels to the design (50, 750, 1450, 2150). Our main analysis strategy for this design  
285 is based on linear mixed models (see Analysis Plan for details).



286 *Figure 3.* Motivated Learning Task. Example trials for the learning and recognition tasks. During learning, participants are required to memorise landscape  
 287 images. Each image is associated with a different reward shown as gems in a treasure chest before each image. During test participants' memory for those  
 288 images is tested. For each landscape image, participants decide whether an image is old (i.e., the image was shown during learning) or new (i.e., the image  
 289 was not shown during learning) and rate their confidence in their decision using a 4-point Likert scale (guess, somewhat sure, sure, very sure). If a participant  
 290 decides that an image is old, they will be asked to indicate the reward amount that image was associated with. If a participant makes a correct old/new  
 291 decision they are rewarded the amount that was presented alongside the image during learning and if the participant makes an incorrect decision, they lose the  
 292 mean value of all possible rewards (i.e. 1100 gems).

REWARD AND SLEEP-BASED MEMORY CONSOLIDATION



293

294 *Figure 4.* Experimental procedure for the proposed experiment. Before starting the experimental sessions, participants complete a recruitment  
 295 session where their demographic information is collected and a number of questionnaires are completed. If participants are eligible to participate  
 296 they undergo two experimental sessions, once with a retention interval of sleep and again with a retention interval of wake (in a counter-  
 297 balanced order). In both sessions the procedure is otherwise identical. Both sessions are separated by at least 1 week and a maximum of 4 weeks.

298

299 **Procedure**

300 Participants are pseudo-randomly assigned to complete the sleep or wake condition  
301 first, such that the order of the sleep and wake conditions is balanced across the sample. This  
302 is also accounted for in our stratification where half of the participants within each stratum  
303 will complete either the sleep or wake condition first. Participants will complete the sessions  
304 separated by a minimum of 1 week and a maximum of 4 weeks. In the sleep condition,  
305 participants complete the learning phase (i.e. learning task and immediate recognition task) in  
306 the evening (between 18:00 – 00:00) and the retrieval phase (i.e. delayed recognition task) in  
307 the morning (between 06:00 – 12:00). In the wake condition, participants complete the  
308 learning phase in the morning (between 06:00 – 12:00) and the retrieval phase in the evening  
309 (between 18:00 – 00:00). In both cases participants must select a two-hour window separated  
310 by 12 hours in which the learning and test phases will be completed (i.e., 06:00 – 08:00,  
311 08:00 – 10:00 or 10:00 – 12:00 and 18:00 – 20:00, 20:00 – 22:00 or 00:00). For example, if  
312 the participant completes the learning phase between 08:00 – 10:00 and the test phase  
313 between 20:00 – 22:00 in the wake condition they must also participate in both phases  
314 between 20:00 – 22:00 and 08:00 – 10:00 in the sleep condition. This will help to constrain  
315 differences in the retention interval between the sleep and wake conditions.

316 *Recruitment session.* The recruitment session can take place at any time prior to the  
317 experiment and participants will be asked to use the same device they used to sign up for all  
318 sessions. A captcha will be used on all session to avoid including bots. All data will be collected  
319 using the participant's chosen device, limited to either a computer, laptop or tablet. Therefore,  
320 participants completing the experiment on a computer or laptop will respond using their mouse  
321 and keyboard whereas those using a tablet will be able to respond using touch screen buttons.  
322 Next, they will receive information about the study and digitally sign the consent form. After

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

323 that, they will answer the demographic questions and depending on strata vacancies they will  
324 be allowed to participate. Then participants will be introduced to the cover story of the  
325 experiment.

326

327         During the recruitment session participants will also complete the screening questions,  
328 St Mary's Hospital Sleep (SMHS) Questionnaire<sup>66</sup>, Epworth Sleepiness Scale (ESS)<sup>67</sup>, reduced  
329 Morningness-Eveningness Questionnaire (rMEQ)<sup>68,69</sup>, Pittsburgh Sleep Quality Index  
330 (PSQI)<sup>70</sup>, Alcohol Use Disorders Identification Test (AUDIT)<sup>71</sup>, the Caffeine Consumption  
331 Questionnaire<sup>72</sup>, the Behavioral Inhibition System/Behavioral Activation System scale  
332 (BIS/BAS)<sup>73</sup>, the Becks Depression Inventory – Short Form (BDI - SF)<sup>74</sup>, and finally the  
333 shortened version of the Hagen Matrices Test (HMT-S)<sup>75</sup>. For the Learning phase in  
334 Experimental Session 1 (see Figure 3), participants will receive an email with a participation  
335 link and times when the experiment can be started depending on which retention condition  
336 (sleep or wake) they are assigned to first. Participants will receive automated emails shortly  
337 before each part of the experimental procedure to remind them to participate. Participants will  
338 be asked to refrain from drinking alcohol (i.e. participants should not drink alcohol 24-hours  
339 prior to and during the first or second experimental sessions) and consume no more or less than  
340 their usual caffeine intake whilst they are actively participating in this experiment.

341         *Experimental sessions.* The first experimental session will take place the earliest within  
342 24-hours of participants completing the recruitment session. At the beginning of the  
343 experimental session participants first indicate when they last consumed caffeine or alcohol  
344 and how much they consumed. Then they indicate their subjective sleepiness (SSS)<sup>76</sup> and  
345 complete a vigilance task (PVT)<sup>77</sup>. Next, they are presented with instructions describing the  
346 Motivated Learning task and how they should perform the first and second parts of the learning  
347 phase, the learning task (duration approximately 19 minutes) and the immediate recognition

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

348 task (duration approximately 14 minutes). In those instructions participants are explicitly  
349 informed about the reward contingencies described on p. 24. Participants are then asked the  
350 validation questions to ensure that they understand the task and then undergo the learning phase  
351 (consisting of the learning task and the immediate recognition task). At this point, participants  
352 in the sleep condition will be instructed to go to sleep at their usual bedtime and wake up at  
353 their usual waking time and participants in the wake condition will be asked not to nap, since  
354 even ultra-short naps may allow for sleep-dependent consolidation<sup>78</sup>. After at least twelve  
355 hours, participants will return to the experiment. Participants completing the sleep condition  
356 will first answer questions about their sleep quality (SMHS)<sup>66</sup> and will answer the sleep related  
357 questions. Participants completing the wake condition will also be asked the sleep related  
358 questions (pertaining to the night before participating) and will be asked “have you taken a nap  
359 today?” and if so “How long did the nap last, in minutes?”. All participants are once again  
360 asked if they have consumed any alcohol or caffeine, how much they consumed, rate their  
361 subjective sleepiness and vigilance is assessed a second time. They then receive instructions  
362 on completing the retrieval phase, answer the validation questions a second time, complete the  
363 delayed recognition task (duration approximately 14 minutes) and complete a verbal fluency  
364 task<sup>79</sup>. This concludes one session of the experiment and participants will then receive further  
365 instructions about the second session. Participants repeat the experimental session, known as  
366 Experimental Session 2 (see Figure 4), but at different times depending on which retention  
367 condition they completed first. At the end of the second session, participants will be debriefed  
368 and receive a profile of the questionnaire data they have provided relative to population. This  
369 feedback on their questionnaire data is used as a further incentive for participants to complete  
370 the study.

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

371 *Follow-up.* Participants will complete a long-term retrieval phase of the Motivated  
372 Learning Task 3-months after completing the second experimental session. Data collected on  
373 this part of the task will be used for exploratory purposes only.

**374 Materials**

375 *Motivated Learning Task.* This task was adapted from a previous study and the  
376 appearance of the task has been adapted to map onto the cover story below<sup>34</sup>.

377 *Cover Story.* To enhance motivation in the Motivated Learning Task we gamified it  
378 with a cover story, where the participants are part of a pirate ship crew. As a crewmember the  
379 aim of the participant is to scout treasure hidden in different locations (landscape images)  
380 occupied by their allies, corresponding to the first part of the learning phase of the experiment.  
381 Participants scout at those locations alone so they cannot take the treasure with them.  
382 Participants navigate between the locations in the form of the Flankers Task embedded amongst  
383 a treasure map. Thus, they must remember the locations and scavenge the treasure when they  
384 return with the crew, which corresponds to the second part of the learning phase (immediate  
385 recognition) and the retrieval phase (delayed recognition). When participants return for the  
386 immediate or delayed recognition along with their crew, they revisit “old” locations (the  
387 locations shown during the first part of the learning phase) and “new” locations (locations that  
388 were not shown during the first part of the learning phase and that are known to be inhabited  
389 by rival pirate clans). The goal of the participant is to “dig” at “old” locations as that is a hit  
390 and they will be rewarded treasure. They should avoid choosing to “not dig” at “old” locations  
391 as that is a miss and the crew captain will punish the crew with a loss of treasure. Since the  
392 “new” locations are occupied by rival pirate crews “digging” at those locations, a false alarm,  
393 is costly, as the rival pirates will take treasure from them. However, if participants choose to  
394 “not dig” at the “new” locations, a correct rejection, the crew captain will reward them with  
395 treasure as digging there could have risked the crew’s safety.

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

396 *Task description.* First, a fixation spyglass is shown during a jittered inter trial interval  
397 (ITI, 1500 – 2000 ms) indicating that it is time for the participant to begin scouting for treasure  
398 at a new location (see Figure 3). Then a treasure chest is shown (2000 ms) indicating how much  
399 treasure can be gained for correct recognition of this image during immediate or delayed  
400 recognition using one of four reward magnitudes (50, 750, 1450, 2150). This is followed by an  
401 additional fixation spyglass. Next, the image of the location is presented. **Each image is only**  
402 **shown once during the learning task.** After viewing each image, participants complete three  
403 trials of the flanker task to prevent rehearsal<sup>70</sup>. Participants are informed that their chances of  
404 winning a monetary bonus increase the more gems they collect. We will use four different  
405 image exposure durations (in ms 1500, 1833, 2167, 2500) to control for encoding strength. The  
406 durations and rewards associated with each image are counterbalanced so that all reward  
407 magnitudes are presented with each duration. Each of the sixteen reward × duration  
408 combinations are implemented eight times (using different images) therefore participants are  
409 shown 128 images during the first part of the learning phase. The images are pseudo randomly  
410 presented to ensure that the same reward or duration do not occur consecutively. The learning  
411 task is split into eight blocks with sixteen images presented per block and in each block at least  
412 six and at most 10 images will be associated with high reward (either 750 or 1100 gems) and  
413 at least six and at most 10 images will be associated with a long duration (either 1500 or 2000  
414 ms). At the beginning and the end of the learning task participants will complete 4 additional  
415 trials, each with pseudo-random rewards and durations occurring only once, to buffer primacy  
416 and recency effects.

417 In the flanker task arrows will be presented to the participant and the direction that the  
418 middle arrow faces will correspond to the directional button which the participant must press<sup>69</sup>  
419 whereas the arrows adjacent to the middle arrow must be ignored. There are congruent  
420 (i.e. flanking arrows face the same direction, >>>>>) and incongruent (i.e. the flanking arrows

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

421 face the opposite direction, >><>>) trials that will be split across all trials of the learning phase.  
422 If participants respond too slowly (i.e. >1.5s) they will be asked to speed up, [participants who](#)  
423 [respond too slowly after three consecutive trials of the learning task \(i.e., on nine consecutive](#)  
424 [flankers\) will be excluded from the data analysis.](#) The flankers trials are pseudo randomized  
425 such that a maximum of six trials can be of the same congruency and orientation in a row.

426 Participants' memory for half of the learned images (i.e., 64 of 128) is tested in the immediate  
427 recognition task directly after the learning task and the other half of the images is tested in the  
428 delayed recognition task in the retrieval phase. One trial of the test phase is shown in Figure 3.  
429 A recognition trial begins with a shovel during a jittered ITI (500 - 1000ms), indicating to  
430 participants that they will begin collecting treasure. Next participants must click a 'continue'  
431 button to ensure that the mouse pointer or finger (when using a tablet) is in approximately the  
432 same position for all trials. Then participants are shown the image of the location and make  
433 three decisions.

434 First, participants must indicate if the image is "old" or "new" [to measure memory](#)  
435 [performance.](#) If the image is "old" and the participant decides the image is "old", then that is a  
436 hit and participants are rewarded the number of gems that the image is associated with. If the  
437 image is "new" and the participant decides that the image is "new" then that is a correct  
438 rejection and they are rewarded the average reward (1100 gems). If the image is "old" and the  
439 participant decides that the image is "new" then that is a miss and the participant loses the  
440 average reward. If the image is "new" and the participant decides that the image is "old" then  
441 that is a false alarm and they lose the average reward amount. The second question participants  
442 are asked is "how certain are you?" using a four-point Likert scale ("guess", "somewhat sure",  
443 "sure", "very sure"). [Confidence is routinely measured in recognition memory tasks and we](#)  
444 [have decided to keep this assessment, as in some cases reward effects have been reported to be](#)  
445 [more pronounced for high confidence items](#)<sup>34</sup>. Finally, if the participant decided that the image

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

446 is “old” they are asked “which treasure do you think can be found here?” and must select one  
447 of the four reward options that they believe the current image is associated with. [This question](#)  
448 [will measure source memory for the reward categories](#). Participants are asked to decide if the  
449 image is “old” or “new”, rate their confidence and select the associated reward as fast as  
450 possible. Each decision must be made within 5000 ms. If participants do not respond within  
451 that time to either of the questions they will receive a warning message. After three warning  
452 messages participants will be reminded that they will be excluded from the experiment if they  
453 do not respond fast enough. When participants are shown a warning message, participants are  
454 still able to respond. Response times starting from the presentation of the location to the time  
455 at which an old/new decision is made and from that decision to the time at which participants  
456 rate their confidence and from that time until a reward is selected will be recorded for  
457 exploratory analyses.

458 For the immediate recognition task there will be eight blocks with 16 trials each,  
459 equaling 128 trials. Sixty-four of the 128 trials are old landscape images (i.e., half of the learned  
460 images) and the remaining 64 are new landscape images. Old and new trials are pseudo  
461 randomly presented such that no more than four target or lure trials can occur in a row and the  
462 same reward and duration can also not appear in a row. Between each block, participants will  
463 be shown an animation of the number of gems they have collected so far. However, this mock  
464 feedback is not influenced by true performance but rather corresponds to the slightly jittered  
465 mean number of gems that could be earned  $\pm 1$  SD with 50% accuracy during the task. This is  
466 done to keep motivation high for all participants irrespective of their true performance. In our  
467 pilot experiments (see supplementary material; [https://cloud.zi-](https://cloud.zi-mannheim.de/index.php/s/jDnY35CM4WMdQCg)  
468 [mannheim.de/index.php/s/jDnY35CM4WMdQCg](https://cloud.zi-mannheim.de/index.php/s/jDnY35CM4WMdQCg)), none of the participants noticed this was  
469 mock feedback.

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

470 In the delayed recognition task, participants complete the same procedure as in the  
471 immediate recognition task, except that participants are presented with a different set of images,  
472 i.e., the remaining 64 old images and 64 completely new images. In the follow-up, participants  
473 are shown all 256-target images they were shown during the first and second recognition phases  
474 in sessions 1 and 2 and will be shown 256 completely new images as lures.

475 ***Landscape images.*** The images are allocated in a way which means that each image is  
476 balanced across the combinations of reward and duration as well as the different time points  
477 and old/new assignments. The landscape images were collected from the creative commons  
478 online repository (<https://search.creativecommons.org/>). A pilot study conducted on Prolific  
479 (<https://www.prolific.co/>;  $N = 152$ , see supplementary material; [https://cloud.zi-](https://cloud.zi-mannheim.de/index.php/s/jDnY35CM4WMdQCg)  
480 [mannheim.de/index.php/s/jDnY35CM4WMdQCg](https://cloud.zi-mannheim.de/index.php/s/jDnY35CM4WMdQCg)) assessed those images in terms of  
481 aesthetics, composition, memorability, familiarity, whether or not the exact images have been  
482 seen before and memory accuracy. Participants rated images on those factors and subsequently  
483 completed a recognition memory test. This pilot allowed us to balance out differences on those  
484 factors between the images across the conditions as well as reward and duration categories in  
485 the Motivated Learning Task and eliminate images that are extremely recognizable.

486 ***Demographic information.*** All participants will be asked the following questions in a  
487 custom online questionnaire: What is your age?; What is your biological sex?; Which gender  
488 do you identify as?; What is your ethnicity?; What is your highest level of school education?;  
489 What is your highest professional qualification?; What is your aspired level of education?;  
490 Which type of school did you go to?; What is your current occupation? What is your  
491 relationship status?; Do you have children and if so how old are they?; Are you currently living  
492 in Germany? If yes, what are the first two numbers of your postcode? Do you live in an urban  
493 or rural area?

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

494            *Participant questionnaire.* Participants will be asked the following yes/no questions:  
495 Do you currently smoke cigarettes? If “yes” how long have you smoked them for?; Do you  
496 currently take any recreational drugs? If “yes” which drugs do you take and how long have you  
497 taken them for?; Do you currently suffer from a diagnosed sleep disorder?; Do you currently  
498 suffer from a diagnosed neurological disorder?; Are you currently taking any prescribed  
499 medication?; Do you currently suffer from a diagnosed addiction disorder?; Do you currently  
500 suffer from a diagnosed mental health disorder?; Participants will be asked to state, which  
501 disorder they suffer from and which medication they are taking if they answer “yes” to the  
502 questions regarding sleep, neurological, addiction or mental health disorders or those who  
503 answer “yes” to taking medication will be asked to indicate which medication they are taking;  
504 Have you traveled across time zones within the past three weeks? If “yes”, where did you travel  
505 to?; Do you currently work as a shift worker? Or have you ever worked as a shift worker?. If  
506 “yes”, how long have/did you work(ed) as a shift worker? and have you worked night shifts  
507 within the past 6 months?. The data collected from this questionnaire will be used for  
508 exploratory purposes only. Of note, we will not use these questions to exclude participants  
509 even though this is done in similar research. We do this to enable exploring moderators.

510 *Sleep questionnaires.*

511 The data collected from the following sleep related questionnaires will be used for exploratory  
512 analyses only, [examples of those analyses are provided underneath each questionnaire.](#)

513            *Caffeine Consumption Questionnaire.* Participants will indicate which caffeinated  
514 products they have consumed throughout the day before participating (including coffee,  
515 decaffeinated coffee, espresso, black, green, white, or mate tea, cocoa drink, iced tea, drinks  
516 with tea extract, cola and mixed cola beverages, energy drink, energy shot, alcopops with  
517 energy drink, cola or coffee and chocolate) and will indicate when they consumed those

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

518 products (breakfast, between breakfast and lunch, lunch, between lunch and dinner, dinner and  
519 after dinner)<sup>72</sup>. The caffeine consumption questionnaire will be translated directly from English  
520 into German. The amount of caffeine which participants have consumed will be used in  
521 exploratory analyses to determine whether or not memory performance in the sleep and wake  
522 conditions for high and low rewards is moderated by caffeine consumption.

523 *Sleep Related Questions.* All participants will indicate their bedtime, rising time and the  
524 length of time that they spent asleep. Participants will also indicate if they had any awakenings  
525 and if so, how many they had.

526 *Epworth Sleepiness Scale (ESS).* The ESS asks participants to rate their general  
527 sleepiness in eight everyday scenarios using a four-point scale (0 = would never doze, 1 = slight  
528 chance of dozing, 2 = moderate chance of dozing and 3 = high chance of dozing)<sup>67</sup>. Scores on  
529 the ESS range from 0-24, a low ESS score indicates low levels of general sleepiness and a high  
530 score indicates high levels of general sleepiness. The German version of the ESS will be used  
531 in this experiment<sup>80</sup>. The Epworth sleepiness scale will be used to determine whether higher  
532 levels of sleepiness cause detrimental effects to the relationship between sleep and memory  
533 consolidation for rewarded information.

534 *Stanford Sleepiness Scale (SSS).* The SSS asks participants to rate their current level of  
535 subjective sleepiness on a seven-point scale (1 = feeling active, vital, alert or wide awake; 7 =  
536 no longer fighting sleep, sleep onset soon, having dream-like thoughts)<sup>76</sup>. A low score on the  
537 SSS indicates a low level of state sleepiness and a high score on the SSS indicates a high state  
538 level of sleepiness. SSS scores will be used to determine whether differences in memory  
539 performance between the sleep and wake condition may be attributed to differences in  
540 subjective sleepiness. The SSS will be directly translated from English to German<sup>as in 41</sup>.

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

541 *Reduced Morningness-Eveningness Questionnaire (MEQr)*. The MEQr is a reduced  
542 version of the full Morningness-Eveningness Questionnaire (MEQ), which uses only 5-items  
543 from the MEQ (i.e. items 1, 7, 10, 18 and 19: e.g., “During the first half hour after having  
544 woken in the morning, how tired do you feel?”)<sup>69,81</sup>. The MEQr measures an individual’s  
545 chronotype (i.e. the time of day that an individual feels most alert)<sup>69</sup> and scores on the MEQr  
546 range from 4-26. Scores below 12 are indicative of a morning type whereas scores greater than  
547 17 are indicative of an evening type. Scores between 12 and 17 are indicative of neither type<sup>69</sup>.  
548 The German version of the MEQr will be used in this experiment<sup>81</sup>. *The MEQr will be used to*  
549 *determine whether chronotype synchrony (i.e., whether you are participating at a time that*  
550 *matches your chronotype) impacts the relationship between sleep and memory consolidation*  
551 *for reward.*

552 *St Mary’s Hospital Sleep (SMHS) Questionnaire*. The SMHS is a subjective measure of  
553 sleep quality over the last 24 hours<sup>66</sup>. Participants will answer items 6 (“How many times did  
554 you wake up?; using a 7-point Likert scale ranging from 1 = “Not at all” to 7 = “More than six  
555 times”) and 9 (“How well did you sleep last night?”; using a 6-point Likert scale ranging from  
556 1 = “very badly” to 6 = “Very well”). Lower scores on item six indicate high sleep quality and  
557 higher scores indicate poor sleep quality. Higher scores on item nine indicate high sleep quality  
558 and lower scores indicate poor sleep quality. The selected SMHS items will be directly  
559 translated from English to German. *Ratings for both items will be used to see if memory*  
560 *performance for high to low reward items is correlated with the level of sleep quality*  
561 *experienced between the learning and testing phases of the sleep condition.*

562 *The Pittsburgh Sleep Quality Index (PSQI)*. The PSQI is another subjective measure of  
563 sleep quality, except participants are asked about their sleep habits and over the past month  
564 (e.g., “During the past month, how often have you had trouble sleeping because you cannot get

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

565 to sleep within 30 minutes?")<sup>70</sup>. The PSQI consists of 18 items, which are clustered into seven  
566 component scores, which each range from 0-3 and are summed. Thus, PSQI scores can range  
567 from 0-21, where lower scores indicate poor sleep quality and higher scores indicate good sleep  
568 quality. The German version of the PSQI will be used in this experiment<sup>82</sup>. Like the SMHS  
569 scores on this scale will be used to see if memory performance for high to low reward items is  
570 correlated with participants general level of sleep quality experienced over the past month.

571 ***Psychomotor Vigilance Task (PVT)***. The PVT is a sustained attention task used to  
572 measure participants' objective vigilance<sup>77</sup>. We will use a 3-minute version of the Psychomotor  
573 Vigilance Task adapted from a 5-minute version of the task<sup>83</sup>. In this reaction time task,  
574 participants have to press the space bar as soon as a millisecond clock appears on the screen.  
575 The following measures will be analysed: median reaction speed (1/reaction time in ms) and  
576 percentage of lapses (number of lapses divided by the number of valid stimuli, excluding false  
577 starts; lapse = reaction time  $\geq$  500 ms). Reaction times shorter than 100 ms will be regarded as  
578 anticipated responses and treated as errors of commission. Participants will be instructed to  
579 respond as soon as a stimulus is shown on the screen. That is, they should shorten their reaction  
580 times as best they can but should not press the response button too early - this is a false start.  
581 Alongside the SSS, data collected from this task will be used to determine whether differences  
582 identified between sleep and wake conditions are due to that manipulation alone and not  
583 differences in subjective and objective vigilance, respectively.

584 ***Regensburger Wortflüssigkeits-Test***. The Regensburger Wortflüssigkeits-Test<sup>79</sup>, is a  
585 measure of verbal fluency. Participants are asked to type as many words as possible within a  
586 two-minute time window. They do this twice: once for words beginning with the letter p or m  
587 and once for words belonging to the category professions or hobbies. The order of the letter  
588 and category version will be balanced across participants. The different versions are used for

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

589 the two experimental sessions respectively. The data will be used to determine differences in  
590 general retrieval performance between sleep and wake groups. The order of the cued letters  
591 and categories will be randomised for each participant between the retrieval sessions.

592 ***BIS/BAS scale.*** The BIS/BAS scale<sup>73</sup> is a measure of both the behavioral inhibition  
593 system (BIS) and behavioral activation system (BAS), both of which are related to motivation  
594 towards moving away from aversive outcomes and moving towards goal-oriented outcomes  
595 respectively. The scale consists of a total of 24 self-report items, 7 of the items are associated  
596 with the BIS component (e.g., “criticism or scolding hurts me quite a bit”) and 13 of the items  
597 are associated with the BAS component (e.g., “I go out of my way to get things I want”). The  
598 other 4 items are fillers. For all of the item’s participants respond using a 4-point Likert scale  
599 (1 = “very true for me” – 4 = “very false for me”). Higher scores on the BIS component of this  
600 scale indicate that an individual is more likely to experience negative feelings when pursuing  
601 a goal. The BAS component can be broken down further into three categories, the BAS  
602 responsiveness score, drive score and fun-seeking score. Generally higher scores on the BAS  
603 components indicated that an individual is more likely to seek out a goal because it is  
604 rewarding. This scale will be used to perform exploratory analyses to determine if there are  
605 any relationships between BIS and BAS scores on memory performance. The German version  
606 of this scale will be used in this experiment<sup>84</sup>.

607 ***Becks Depression Inventory – Short Form (BDI - SF).*** The BDI – SF is a shortened  
608 version of the original BDI<sup>85</sup> containing only 13 items instead of 21. The BDI – SF is a measure  
609 of depressive symptoms, which are indicative of depression. For each item on this scale  
610 participants respond using a 4-point Likert scale (e.g., 1 = “I do not feel sad” - 4 = “I am so sad  
611 or unhappy I can’t stand it”). Scores on the BDI-SF range from 0 – 39, where lower scores are  
612 indicative of fewer depressive symptoms. This scale will be directly translated from English

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

613 into German. This scale will be used to determine whether there is a reduced effect of high  
614 rewards on memory after sleep for participants who report higher levels of depressive  
615 symptoms.

616 ***Hagen Matrices Test Short Version (HMT-S)***. The HMT-S<sup>75</sup> is an adapted shorter  
617 measure of the HMT which measures intelligence, specifically induction and fluid reasoning.  
618 In this task participants are required to identify patterns and rules in a series of puzzles. They  
619 are shown six 3 X 3 matrices of patterns that are incomplete and are missing one part.  
620 Participants must select the correct solution to the matrix from 8 potential options. In order to  
621 successfully complete the task participants must be able to identify the rules that govern the  
622 matrices they are shown. Each item is given a score of 1 if it is answered correctly and a score  
623 of 0 if the response is incorrect or missing. Thus, participants scores can range from 0 to a total  
624 of 6. Higher scores in this task indicate that an individual has greater induction and fluid  
625 reasoning abilities. Given that intelligence has been broadly related to the benefits of sleep on  
626 memory consolidation and the occurrence of neurophysiological activity that occurs during  
627 sleep<sup>86</sup> it is plausible that performance on this task might influence memory performance on  
628 the MLT. Therefore, we will conduct exploratory analyses to determine whether or not the  
629 benefits of sleep on reward related memory performance is associated with performance on the  
630 HMT-S.

631 ***Validation questions***. To ensure that participants understand the Motivated Learning  
632 Task they are provided with information about how they will be rewarded gems or how they  
633 might lose gems as described previously (see Table 3). To test their understanding, they are  
634 asked the following questions: You see a landscape image that you already saw when you were  
635 scouting for treasures. The treasure chest belonging to it contains 750 gems. You correctly  
636 identify the picture as “old” - what happens?; You see a landscape image that you already saw

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

637 when you were scouting for treasures. The treasure chest belonging to it contains 2150 gems.  
 638 You make a mistake and identify the picture as “new” - what happens?; You see a landscape  
 639 image that you did not see when you were scouting for treasures. The treasure chest belonging  
 640 to it contains 2150 gems. You correctly identify the picture as "new" - what happens?; and  
 641 finally you see a landscape image that you did not see when you were scouting for treasures.  
 642 The treasure chest belonging to it contains 1450 gems. You make a mistake and identify the  
 643 picture as “old” - what happens? The chances of getting all four validation questions that have  
 644 four options to answer each correct by chance is  $0.25^4 = 0.004$ . On the occasions which  
 645 participants will answer the validation questions they will be given two opportunities to  
 646 correctly answer all of them. If they incorrectly answer **at least one of** the validation questions  
 647 **on their first attempt** they will be given a second opportunity to answer them. If they incorrectly  
 648 answer **at least one of the questions** on the second try they will be excluded from the  
 649 experiment. **If the participants answer all of the validation questions correctly on their first or**  
 650 **second attempt they will be able to continue the experiment. On the second attempt participants**  
 651 **are also shown the instructions for the motivated learning task a second time.** Given that  
 652 participants will complete the validation questions four times, the gems referred to in the  
 653 questions will be adjusted each time.

654 Table 3. Reward contingencies for the Motivated Learning Task.

	<b>Trial Type</b>		
	<i>Reward Contingencies</i>	Target	Lure
<b>Response</b>	“Yes”	Hit (win n gems)*	False Alarm (lose 1100 gems)
	“No”	Miss (lose 1100 gems)	Correct Rejection (win 1100 gems)

655 \*n refers to the number of gems which are associated with a given target image.

656           *Wake experience.* Participants will be asked to document their wake experience in the  
657 wakefulness condition. For this, they will be asked, “please provide a short description of your  
658 activity during each hour of the retention period”. *To document their wake experience*  
659 *participants will be asked to recall and approximate their activity of each hour during the day*  
660 *by typing it into the relevant fields.* Based on the received answers we will categorize the data  
661 and perform exploratory analyses.

662           *Seriousness check.* To ensure that participants performed the experiment seriously and  
663 did not engage in nefarious activities, such as repeat participations or masking of their true  
664 location via VPN we will perform a seriousness check at the end of the study. Participants will  
665 be asked, “It would be very helpful if you could tell us at this point whether you have taken  
666 part seriously, so that we can use your answers for our scientific analysis, or whether you were  
667 just clicking through to take a look at the survey? *Please note that any answer that you provide*  
668 *to this question will not impact your chances of winning in the prize draw or prevent you from*  
669 *being added to the prize draw*” and can respond with “I have taken part seriously” or “I have  
670 just clicked through, please throw my data away”. This approach has been shown to improve  
671 data quality in online studies<sup>87</sup>. This information will be used in exploratory analyses to reveal  
672 whether the seriousness of participant’s responses impacts the confirmatory analyses described  
673 below in the analysis plan.

## 674 **Analysis Plan**

675

676           The data simulations presented below and in Figure 1 were carried out using R (version  
677 4.2.0) running in RStudio<sup>88</sup>. All analyses will be performed in Rstudio after the data have been  
678 collected.

679

680 **Data Pre-processing.**

681 **Hit and False Alarm Rates.** To compute the hit rate, the number of hits will be divided by the  
682 corresponding number of target trials. Comparatively, to compute the false alarm rate the  
683 corresponding number of false alarms will be divided by the number of lure trials. Hit and false  
684 alarm rates will be computed for all combinations of retention (sleep vs. wake), timepoint  
685 (immediate vs. delayed) and reward (50 vs. 750 vs. 1450 vs. 2150). This means that hit and  
686 false alarm rates are computed for each participant are collapsed across all durations for all  
687 levels of interest and duration conditions will be used to perform exploratory analyses. **The**  
688 **main focus of analyses of the duration conditions will be to confirm that low memory strength**  
689 **items (those that were shown for the shortest time) benefit most from sleep-dependent**  
690 **consolidation, as has been reported before<sup>7,89</sup>. The duration conditions will also allow us to**  
691 **perform exploratory analyses that take into account differences in memory performance due to**  
692 **age or other demographic variance.** Following the original paper which developed the MLT<sup>30</sup>,  
693 the hit rate will be used as our main outcome variable. In this task, the hit rate is the most  
694 conceptually relevant outcome measure. This is because only targets are associated with a  
695 reward and the lures are not, since they are only shown in the test phase. Therefore, only the  
696 hit rate should be modulated by the reward, not the false alarm rate.

697 **Discriminability.**  $d'$  will be computed from the hit and false alarm rates (ignoring the  
698 reward categories) for each participant as follows:

$$699 \quad d' = z(HR) - z(FAR)$$

700 The measure  $d'$  is an operationalization of discriminability (i.e., participants ability to  
701 discriminate between old and new images)<sup>90</sup>.  $d'$  will be used to calculate participants general  
702 memory performance and exclude outliers as described above.

### 703 **Model Specification and Hypothesis Testing**

704 Our hypotheses (shown in Table 1 and Figure 1) will be tested and formalised using the  
 705 R-package lmerTest<sup>91</sup> in the following linear mixed effects model with a maximal random  
 706 effects structure as is recommended in the literature<sup>92</sup>:

```
707 hit rate ~ timepoint * retention * reward + ((timepoint + retention + rewa  

  708 rd) ^ 2 | subject)
```

709 This maximal linear mixed effects model includes all interactions **and** main effects **as**  
 710 **well as** random **intercepts and** slopes for each participant for all parameters, with the exception  
 711 of the three-way interaction where only one data point per participant exists, as the slope for  
 712 that interaction and the random residual error would be indistinguishable. Deviation coding  
 713 will be used for all categorical predictors in this model (See Table 4). Reward will be **scaled**  
 714 **such that a change in reward values reflects an increase of 1000 gems** collapsed across duration  
 715 categories. p-values produced by the lmerTest package using Satterthwaites degrees of freedom  
 716 will be used to evaluate relevant parameters in this model (see below).

717 Table 4: Coding scheme.

predictor	-0.5	0.5
timepoint	immediate	delayed
retention	wake	sleep

718 *Note.* Deviation coding of predictors for the main analysis.

719

720 We will use the maximal model to give us an indication of whether our prediction that  
 721 the magnitude of decline in memory for high vs. low rewarded images will be greater after a  
 722 period of wake compared to a period of sleep at delayed recognition. This is represented in the  
 723 timepoint × retention × reward parameter. **If the timepoint × retention × reward is non-**

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

724 significant and an equivalence test suggests equivalence, we will conclude that there is no effect  
725 of reward on sleep-dependent memory consolidation. If the timepoint x retention parameter is  
726 non-significant and an equivalence test suggests equivalence, we will conclude that reward  
727 does not affect consolidation and reward effects are due to processes during encoding alone. If  
728 the timepoint x retention x reward is significant, we will followed it up with additional tests  
729 since the interaction could be taking place in any combination of those variables (for example  
730 at both immediate and delayed recognition). Therefore, we will reduce the model by the  
731 timepoint parameter and examine two linear mixed effects model for both immediate and  
732 delayed recognition with the following maximal effects structure:

```
733 hit rate ~ retention * reward + ((retention + reward) ^ 2 | subject)
```

734 First, we will examine the reduced model for immediate recognition to determine  
735 whether or not the retention × reward parameter is significant, which it would be if an  
736 interaction was present at immediate recognition. We do not expect that this will be a  
737 significant interaction since sleep is not expected to exert any impact on reward memory here  
738 as it has not yet occurred. However, we do expect that the reward parameter will be significant,  
739 where memory for high rewards will be greater than memory for low rewards at immediate  
740 recognition. If counter to this expectation the reward parameter at immediate recognition is not  
741 significant, this in combination with the timepoint × retention × reward of the full model would  
742 mean that rewards at encoding do not suffice to explain the memory enhancing effects of  
743 rewards.

744 Moreover, we will also examine whether the retention condition parameter is  
745 significant. If so, this would indicate that there is a time of day effect at immediate recognition  
746 between the sleep and wake conditions, which may occur due to learning either in the morning  
747 (in the wake condition) or the evening (in the sleep condition). For instance, one might expect

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

748 that participants in the wake condition that learn in the morning perform better as they are well  
749 rested, whereas those in the sleep condition learn in the evening and are tired and as a result  
750 they perform worse. If that is the case then baseline hit rates will be added as a covariate to the  
751 reduced model of delayed recognition to determine whether they have an impact on our  
752 interpretation of the findings as they pertain to the impact of retention on reward.

753 **To test our main hypotheses** we examine the delayed recognition model and predict  
754 that the retention and reward parameters will be significant, where the sleep group collapsed  
755 across reward categories will have greater memory compared to the wake group and high  
756 reward images will be better remembered than low reward images collapsed across retention  
757 conditions, respectively (**H1 and H2; see Table 1**). Consequentially we will examine the  
758 retention  $\times$  reward parameter to assess our final hypothesis, the interaction between retention  
759 and reward at delayed recognition (**H3: see Table 1**). We will follow up this interaction by  
760 performing linear mixed models on each unit of reward between sleep and wake groups as  
761 follows:

```
762 hit rate ~ retention + (retention | subject)
```

763 We expect in this analysis that significant differences between sleep and wake groups  
764 at delayed recognition will emerge in the highest reward categories (1450 and 2150) and that  
765 performance at low reward categories (50 and 750) will be statistically equivalent. Simulated  
766 data demonstrating this pattern of results are shown in Figure 1. If memory performance at  
767 immediate recognition is included as a covariate and is significant and the interaction is not  
768 significant then it will be concluded that the covariate explains more variation in our data than  
769 the interaction between sleep and reward, and the data will be interpreted as such and  
770 explanations will be explored. However if the covariate is included and is significant or non-  
771 significant and the interaction term remains significant then we will conclude that after

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

772 accounting for the variation explained by baseline scores, the retention  $\times$  reward interaction  
773 persists. This pattern of results would replicate previous research indicating that sleep benefits  
774 memory performance over a period of wakefulness and that sleep exerts its influence on  
775 memory for high rewards only<sup>44,49,50</sup>.

776 Evidence indicating that the parameters (including covariates) described above are  
777 unlikely under the null hypothesis will be determined via p-values  $< .020$  and all follow-up  
778 tests will use the same threshold.

779 **Resolving Model Convergence Issues.** It is possible that our maximal models will not  
780 converge due to “overparamitization” within models containing all possible parameters, such  
781 as random intercepts and slopes<sup>93</sup>. Yet, it makes sense to start with the maximal model, since  
782 not including those parameters can yield an increased risk of Type I error<sup>94</sup>. We will perform  
783 30,000 iterations of the maximal model. If the maximal model still fails to converge after  
784 30,000 iterations or the estimated correlation parameters lie at 0 or +1, the data will be fitted  
785 using a zero correlation model. If after 30,000 iterations a model is not identified, random  
786 slopes per participant starting with the highest order components will be excluded until a model  
787 is identified. It is also possible that a model is identifiable, but overparamatization is indicated  
788 in a random-effects Principle Component Analysis (implemented using the rePCA() function  
789 in the lme4 R-package)<sup>95</sup>.

790

Table 5: Simulation model parameters

parameter name	parameter value
fixed effects	
intercept	0.60
timepoint	-0.10
reward	0.015
timepoint:retention:reward	0.015
random effects	
participant intercept	0.12
participant timepoint sd	0.25
participant reward sd	0.02
error	0.12

791

*Note.* Model parameters used in the data-generating model. The parameter value for reward reflects the change in reward per 1000 gems.

792

793 If this occurs, components of the same order which have the smallest variance will be removed

794 from the model. In this scenario, the fixed effects parameters will be evaluated using p-values

795 calculated using lmerTest. To prevent p-hacking, p-values will only be calculated once a model

796 with good convergence is identified. *If either of the following scenarios occur it will be*

797 *concluded that our model derived from the lmer package does not have good convergence: 1)*

798 *the package is unable to converge on a final model and no output is produced; and 2) a model*

799 *is produced but a singular fit is identified indicating that the model has been overfitted to the*

800 *data.*

801 Main effects estimated to be 0 or close to 0 will not be removed, thus ensuring that p-

802 values derived from the identified model can be meaningfully interpreted and confidence

803 intervals can be used in equivalence tests. Lastly, models that fail to converge will be

804 documented and presented in the Supplement upon Stage 2 submission.

805 **Data Simulation.** Model convergence aside, linear mixed effects models also present a

806 unique challenge for estimating effect sizes and analyses, since variance is shared and

807 partitioned amongst all parameters. Therefore, there is no consensus as to how one should

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

808 compute effect size estimates for main effects and interactions<sup>96</sup>. Our solution was to create a  
809 data generating model including parameters that reflect our hypotheses (see Table 5) and  
810 simulate data to calculate an appropriate sample size<sup>97,98</sup>. In this data generating model, the  
811 data produced are aggregated over duration so the hit rate is computed across each timepoint ×  
812 reward × retention combination. The reward parameter was scaled such that a change in reward  
813 values reflects an increase of 1000 gems collapsed across duration categories. Thereby, we  
814 could set the hypothetical reward of zero, immediate recognition and the wake condition to  
815 baseline, which could then be compared to an increase of 1000 gems (reward effect), delayed  
816 recognition and the sleep condition, respectively. In other words, the main effect of reward  
817 reflects an increase in hit rate for every 1000 gems (per reward category). The Main effect of  
818 timepoint reflects the change in hit rate between immediate and delayed testing. Finally, the  
819 main effect of retention reflects the change in hit rate between the sleep and wake condition.  
820 Additionally, unlike the data analysis model shown above, this data generating model uses  
821 dummy coding and not deviation coding for each of the predictors (See Table 5).

822 To make our data generating model as realistic as possible we included by-subject  
823 random intercepts and slopes for main effects and interactions that are expected to be non-zero  
824 in addition to the residual error that is normally distributed. The residual error in this model  
825 was based upon pilot data which were collected in an online environment and can therefore  
826 approximate the error that we may encounter in this experiment using the same task. Moreover,  
827 for memory measures it is known that the measurements between immediate and delayed  
828 recognition will be correlated, so a high correlation between the time points was included  
829 (simulated correlation from the first run of the data simulation:  $r = 0.87$ ). With that in mind  
830 we also assumed that memory would decay for each participant between immediate and  
831 delayed recognition so a reduction in memory performance is also included in the data  
832 simulation.

833

Table 5: Coding scheme.

predictor	0	1
timepoint	immediate	delayed
retention	wake	sleep

834

*Note.* Dummy coding of predictors for the data-generating model.

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

The parameters shown in Table 5 were also derived from our predictions described previously and the existing literature and data. For example, the parameter reflecting the impact of reward on memory in Table 5 was derived from the pilot study which was conducted to validate our task (see supplemental material; <https://cloud.zi-mannheim.de/index.php/s/jDnY35CM4WMdQCg>). The estimated size of that effect was decreased slightly from that data to account for the fact that we will be using a more heterogenous sample. However, deriving parameters for the precise impact of sleep on memory for high rewards at delayed recognition proved challenging. A recent meta-analysis across different tasks indicated that the impact of sleep on memory at delayed recognition ranges between  $d = -.252 - 1.14$ <sup>8</sup> in young and older adults. A meta-analysis on emotional memory found that the effect lies at  $d = .470$ <sup>99</sup> indicating that there is much variability in the size of the sleep effect on memory. This variability is likely increased by the heterogeneous sample we are collecting. The challenge is further complicated by effect size inflation in meta-analyses due to possible publication bias. However, in the face of uncertain effect sizes the goal of our simulation was not to determine a precise a priori effect size. Instead, our approach was to find a compromise between resource constraints and achieving 95% power to detect a broad range of effect sizes.

853

854

855

Therefore, under the reasonable assumption that the impact of sleep on memory and more specifically its impact on high reward at delayed recognition could also be smaller than is reported in the literature we chose parameters for the effect of sleep on reward at delayed

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

856 recognition which yielded an overall small effect. This effect emerges at high rewards and is 0  
857 at low rewards, because we modeled the data in such a way that there is no difference between  
858 wake and sleep for the lowest reward categories (as is shown in Figure 1). The simulation thus  
859 replicates previous research in line with our predictions,<sup>44,49,50</sup> but at a much smaller effect size  
860 than has previously been reported for both the impact of sleep on memory at delayed  
861 recognition and the timepoint  $\times$  reward  $\times$  retention interaction. Note that we do not include a  
862 parameter for simulating the effect of retention since we assume that the impact of sleep exerts  
863 itself on delayed recognition **modulated by** rewards (see Table 1 for further information). **This**  
864 **is supported by the finding that cues bias reactivation for cued memories at the expense of non-**  
865 **cued memories**<sup>101</sup>. Overall, the decisions used to create the model led to the chosen parameters  
866 shown in Table 5 and the data generating model below:

```
867 hit rate ~ timepoint + reward + timepoint:retention:reward + (timepoint + r  
868 eward + timepoint:retention:reward | subject)
```

869 **Power Analysis and Sample Size.** To calibrate our power analysis to achieve at least  
870  $\beta = .95$  to detect our interactions of interest, we first simulated 1000 data sets using the  
871 parameters presented in Table 5 and the data generating model described above starting with a  
872 sample size of 1500 participants. However, since the maximal model does not always converge,  
873 for ease of simulation a simpler data analyzing model was used:

```
874 hit rate ~ timepoint:reward:retention + (1|sub_id)
```

875 Once the first 1000 simulations were complete the proportion of data sets, which yielded  
876 our significant interactions of interest at  $\alpha > .020$  was calculated. If the number of data sets  
877 which yielded our interaction of interest at  $\alpha > .020$  was lower than 95% the simulation was  
878 repeated and the number of participants included was increased by 50. We repeated that process  
879 until at least 95% power was achieved, which was the case at 1750 participants.

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

880 To establish the plausibility of our power analysis, in terms of identifying the timepoint  
881  $\times$  reward  $\times$  retention interaction parameter shown in table 5, we calculated the mean 90%  
882 confidence intervals across all 1000 simulated data sets for our last simulated scenario of 1750  
883 participants. This was carried out to rule out the possibility that the confidence intervals of that  
884 effect include zero as this would indicate that it is possible that no effect would be identified  
885 and would mean our predictions are implausible. This was not the case, the lower and upper  
886 90% confidence intervals were .009 and .021 respectively. We also investigated the  
887 vulnerability of our analysis to Type I errors. This was achieved by simulating 1000 data sets  
888 using the identical parameters above shown in Table 5 with the exception of the timepoint  $\times$   
889 reward  $\times$  retention interaction, which was set to a value of 0. In those simulated data sets 2.10%  
890 incorrectly identified an effect at  $p < .020$ .

891 **Positive Controls.** To ensure that the data we have collected are of sufficient quality  
892 for testing our hypotheses presented in Table 1, we will perform the following positive controls:  
893 1) we will use a repeated measures t-test to confirm that memory between the lowest and  
894 highest reward categories is significantly different such that hit rate is greater for high rewards  
895 compared to low rewards in the delayed recognition test; 2) we will confirm that a retention  
896 interval of 12 hours yields a significant decline in memory performance between immediate  
897 and delayed testing by comparing  $d'$  (collapsed over the other conditions) between immediate  
898 and delayed testing using a repeated measures t-test; and, finally, 3) we will confirm that  
899 participants memory performance as measured using  $d'$  is significantly different from zero at  
900 delayed testing collapsed across all other conditions, which it should be if participants are  
901 capable of discriminating between targets and lures. In the event that one of these tests yields  
902 a statistically non-significant result (as determined using an alpha of  $p > .020$ ) then equivalence  
903 tests<sup>100</sup> will be used and carried out against an equivalence bound of Cohens  $d = -0.10 - 0.10$ .

904 If any of the above analyses are found to be equivalent then it will be concluded that our data  
905 cannot be used to test our hypotheses.

906 **Control Analyses.** Equivalence tests<sup>101</sup> will be carried out to determine whether control  
907 variables across sleep and wake conditions are statistically equivalent, and therefore can be  
908 ruled out as variables that may otherwise explain differences in memory performance between  
909 those conditions. These tests will be carried out against an equivalence bound of Cohens  $d = -$   
910  $0.10 - 0.10$ , we consider effects within this range to be unlikely to influence the main analyses  
911 proposed in our analysis plan. Variables that are not equivalent will be considered in any  
912 interpretation of differences in memory performance between sleep and wake conditions and  
913 will be added as covariates to the model specified above to determine whether our initial  
914 interpretation of the model changes. **Therefore, after evaluating our model without any**  
915 **covariates**, the covariates will be added sequentially to determine the relative impact of each  
916 of them individually on our interpretation of the data. For example, if a given covariate explains  
917 a significant amount of variability in our data such that the remaining variance explained by  
918 our predictions is no longer significant, then it will be concluded that in our design the predicted  
919 effect is not detectable.

920 Equivalence tests will be used to compare SSS scores as well as median reaction speed  
921 and the percentage of lapses in the PVT during the learning phase and the retrieval phase and  
922 the number of words generated in the Regensburger Wortfluessigkeitstest during the retrieval  
923 phase between sleep and wake conditions. It is possible that the performance on the motivated  
924 learning task improves in the second relative to the first session. We will perform an  
925 equivalence test to determine whether there are differences between Experimental Sessions 1  
926 and 2 in immediate recognition and if so session number will be included as a predictor in our  
927 model to determine whether our conclusions change. **The remaining equivalence test that will**

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

928 be performed will be used to examine whether or not the hit rate for low reward items at delayed  
929 testing is statistically equivalent.

930 Since our sample might include individuals who have a sleep disorder (including  
931 insomnia, somnambulism, sleep apnea, REM sleep behavior disorder, narcolepsy or restless  
932 legs syndrome) or neurological disorder known to impact memory (specifically dementia,  
933 Alzheimer's disease and amnesia), we will include the presence or absence of these afflictions  
934 as predictors in an additional model to determine whether our conclusions change. If so, we  
935 will explore reasons why this may be.

936 Moreover, it is also unclear whether sleep benefits memories generally across low and  
937 high rewards, or more highly rewarded items benefit more at the cost of no sleep benefit for  
938 the lowest rewarded items. Therefore, we will conduct a repeated-measures t-test to compare  
939 the hit rate for the lowest reward category between the sleep and wake conditions at delayed  
940 testing. If that t-test is significant at  $p < .020$ , it will be concluded that sleep actively  
941 consolidates information which individuals are not highly motivated to learn even in  
942 competition to more highly rewarded information. If it is not significant, then equivalence tests  
943 will be conducted against an equivalence bound of Cohens  $d = -0.10 - 0.10$  to conclude that the  
944 sleep effect for lowest reward category is not meaningfully higher than 0.

945 In addition, if the main analysis for H3 is not significant (i.e., we do not find a timepoint  
946 x retention x reward effect), we will conduct an analysis on a restricted sample. In this sample  
947 we will exclude all participants with mental disorders (including sleep disorders) and limit the  
948 age range to 20-39 years old. This will allow us to control whether the effect of rewards on  
949 sleep-dependent consolidation is only evident for young healthy adults.

950 Additionally, given that the original experiment which used the Motivated Learning  
951 Task found that the benefits of reward on memory were present at high levels of confidence,

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

952 we will also perform a similar analysis on our data. This will enable us to determine whether  
953 the impact of sleep-based memory consolidation on reward information is modulated by  
954 confidence, a known proxy of memory strength<sup>102</sup>. Therefore, we will add the participants  
955 confidence ratings for correct responses on ‘old’ trials (i.e. hits) as a parameter to our original  
956 analysis and evaluate whether the reward by confidence interaction at delayed recognition is  
957 significant. Data at delayed recognition will be used since the original experiment identified  
958 the effect after a 24-hour retention interval. It is expected that, if the original study’s findings  
959 replicate, the magnitude of the benefit of reward on memory will be larger for high confidence  
960 responses compared to lower confidence responses. If it is the case that the interaction is  
961 significant then further exploratory analyses will be conducted to consider other variables of  
962 interest such as retention and timepoint, although we have no strong predictions at this point in  
963 time.

964 Finally, we will also perform an analysis using the sensitivity measure  $d'$  to enable  
965 comparisons between our analyses and previous research on recognition memory (that did not  
966 manipulate reward values) and sleep dependent memory consolidation. To do this we will  
967 perform t-tests on the delayed recognition data as well as the difference score.

968 **References**

969

- 970 1. Diekelmann, S. & Born, J. The memory function of sleep. *Nature Reviews*  
971 *Neuroscience* **11**, 114–126 (2010).
- 972 2. Feld, G. B. & Born, J. Neurochemical mechanisms for memory processing during  
973 sleep: Basic findings in humans and neuropsychiatric implications.  
974 *Neuropsychopharmacology* **45**, 31–44 (2020).
- 975 3. Stickgold, R. Sleep-dependent memory consolidation. *Nature* **437**, 1272–1278 (2005).
- 976 4. Barret, T. & Ekstrand, B. Effect of sleep on memory. III. Controlling for time of day  
977 effect. *Journal of Experimental Psychology* **96**, 321–327 (1972).
- 978 5. Brawn, T. P., Fenn, K. M., Nusbaum, H. C. & Margoliash, D. Consolidating the  
979 Effects of Waking and Sleep on Motor-Sequence Learning. *Journal of Neuroscience*  
980 **30**, 13977–13982 (2010).
- 981 6. Drosopoulos, S., Wagner, U. & Born, J. Sleep enhances explicit recollection in  
982 recognition memory. *Learning & Memory* **12**, 44–51 (2005).
- 983 7. Drosopoulos, S., Schulze, C., Fischer, S. & Born, J. Sleep's function in the  
984 spontaneous recovery and consolidation of memories. *Journal of Experimental*  
985 *Psychology: General* **136**, 169 (2007).
- 986 8. Gui, W.-J. *et al.* Age-related differences in sleep-based memory consolidation: A  
987 meta-analysis. *Neuropsychologia* **97**, 46–55 (2017).
- 988 9. Jenkins, J. G. & Dallenbach, K. M. Obliviscence during Sleep and Waking. *The*  
989 *American Journal of Psychology* **35**, 605–612 (1924).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 990 10. Tucker, M. A., Tang, S. X., Uzoh, A., Morgan, A. & Stickgold, R. To Sleep, to Strive,  
991 or Both: How Best to Optimize Memory. *PLOS ONE* **6**, e21737 (2011).
- 992 11. Van Ormer, E. B. Retention after intervals of sleep and waking. *Archives of*  
993 *Psychology* **137**, 49 (1932).
- 994 12. Wagner, U., Kashyap, N., Diekelmann, S. & Born, J. The impact of post-learning  
995 sleep vs. Wakefulness on recognition memory for faces with different facial  
996 expressions. *Neurobiology of Learning and Memory* **87**, 679–687 (2007).
- 997 13. Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A. & Stickgold, R. Practice  
998 with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron* **35**, 205–211  
999 (2002).
- 1000 14. Walker, M. P., Stickgold, R., Alsop, D., Gaab, N. & Schlaug, G. Sleep-dependent  
1001 motor memory plasticity in the human brain. *Neuroscience* **133**, 911–917 (2005).
- 1002 15. Wilson, J. K., Baran, B., Pace-Schott, E. F., Ivry, R. B. & Spencer, R. M. Sleep  
1003 modulates word-pair learning but not motor sequence learning in healthy older adults.  
1004 *Neurobiology of aging* **33**, 991–1000 (2012).
- 1005 16. Bailes, C., Caldwell, M., Wamsley, E. J. & Tucker, M. A. Does sleep protect  
1006 memories against interference? A failure to replicate. *Plos one* **15**, e0220419 (2020).
- 1007 17. Morgan, D. P., Tamminen, J., Seale-Carlisle, T. M. & Mickes, L. The impact of sleep  
1008 on eyewitness identifications. *Royal Society Open Science* **6**, 170501 (2019).
- 1009 18. Pöhlchen, D., Pawlizki, A., Gais, S. & Schönauer, M. Evidence against a large effect  
1010 of sleep in protecting verbal memories from interference. *Journal of Sleep Research*  
1011 e13042 (2020).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1012 19. Rickard, T. C., Cai, D. J., Rieth, C. A., Jones, J. & Ard, M. C. Sleep does not enhance  
1013 motor sequence learning. *Journal of Experimental Psychology: Learning, Memory,*  
1014 *and Cognition* **34**, 834–842 (2008).
- 1015 20. Jenkins, J. G., Dallenbach, K. M. Oblivescence during sleep and waking. *American*  
1016 *Journal of Psychology* **32**, 605–612 (1924)
- 1017 21. Mednick, S. C., Cai, D. J., Shuman, T., Anagnostaras, S., Wixted, J. T. An  
1018 opportunistic theory of cellular and systems consolidation. *Trends in Neurosciences*  
1019 **34**, 504-14 (2011)
- 1020 22. Dudai, Y. Karni, A., Born, J. The consolidation and transformation of memory.  
1021 *Neuron* **88**, 20-32 (2015)
- 1022 23. O’Neill, J., Pleydell-Bouverie, B., Dupret, D. & Csicsvari, J. Play it again:  
1023 Reactivation of waking experience and memory. *Trends in Neurosciences* **33**, 220–  
1024 229 (2010).
- 1025 24. Carr, M. F., Jadhav, S. P., Frank, L. M. Hippocampal replay in the awake state: a  
1026 potential substrate for memory consolidation and retrieval. *Nature Neuroscience* **2**,  
1027 147-53 (2011)
- 1028 25. Wilson, M. A. & McNaughton, B. L. Reactivation of hippocampal ensemble  
1029 memories during sleep. *Science* **265**, 676–679 (1994).
- 1030 26. Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G. & Zugaro, M. B. Selective  
1031 suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience* **12**,  
1032 1222–1223 (2009).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1033 27. Marshall, L., Helgadóttir, H., Mölle, M. & Born, J. Boosting slow oscillations during  
1034 sleep potentiates memory. *Nature* **444**, 610–613 (2006).
- 1035 28. Ngo, H.-V. V., Martinetz, T., Born, J. & Mölle, M. Auditory closed-loop stimulation  
1036 of the sleep slow oscillation enhances memory. *Neuron* **78**, 545–553 (2013).
- 1037 29. Ven, G. M. van de, Trouche, S., McNamara, C. G., Allen, K. & Dupret, D.  
1038 Hippocampal Offline Reactivation Consolidates Recently Formed Cell Assembly  
1039 Patterns during Sharp Wave-Ripples. *Neuron* **92**, 968–974 (2016).
- 1040 30. Ackermann, S., Hartmann, F., Papassotiropoulos, A., Quervain, D. J. F. de & Rasch,  
1041 B. No Associations between Interindividual Differences in Sleep Parameters and  
1042 Episodic Memory Consolidation. *Sleep* **38**, 951–959 (2015).
- 1043 31. Mantua, J. Sleep Physiology Correlations and Human Memory Consolidation: Where  
1044 Do We Go From Here? *Sleep* **41**, (2018).
- 1045 32. Feld, G. B., Weis, P. P. & Born, J. The limited capacity of sleep-dependent memory  
1046 consolidation. *Frontiers in Psychology* **7**, 1368 (2016).
- 1047 33. Feld, G. B. & Born, J. Sculpting memory during sleep: Concurrent consolidation and  
1048 forgetting. *Current Opinion in Neurobiology* **44**, 20–27 (2017).
- 1049 34. Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B. & Gabrieli, J. D. E.  
1050 Reward-motivated learning: Mesolimbic activation precedes memory formation.  
1051 *Neuron* **50**, 507–517 (2006).
- 1052 35. Braun, E. K., Wimmer, G. E. & Shohamy, D. Retroactive and graded prioritization of  
1053 memory by reward. *Nature Communications* **9**, 4886 (2018).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1054 36. Loftus, G. R. & Wickens, T. D. Effect of incentive on storage and retrieval processes.  
1055 *Journal of Experimental Psychology* **85**, 141–147 (1970).
- 1056 37. Madan, C. R., Fujiwara, E., Gerson, B. C. & Caplan, J. B. High reward makes items  
1057 easier to remember, but harder to bind to a new temporal context. *Frontiers in*  
1058 *Integrative Neuroscience* **6**, (2012).
- 1059 38. Gruber, M. J., Ritchey, M., Wang, S.-F., Doss, M. K. & Ranganath, C. Post-learning  
1060 Hippocampal Dynamics Promote Preferential Retention of Rewarding Events. *Neuron*  
1061 **89**, 1110–1120 (2016).
- 1062 39. Weiner, B. & Walker, E. L. Motivational factors in short-term retention. *Journal of*  
1063 *Experimental Psychology* **71**, 190–193 (1966).
- 1064 40. Asfestani, M. A. *et al.* Consolidation of reward memory during sleep does not require  
1065 dopaminergic activation. *Journal of Cognitive Neuroscience* 1–16 (2020).
- 1066 41. Feld, G. B., Besedovsky, L., Kaida, K., Münte, T. F. & Born, J. Dopamine D2-like  
1067 receptor activation wipes out preferential consolidation of high over low reward  
1068 memories during human sleep. *Journal of Cognitive Neuroscience* **26**, 2310–2320  
1069 (2014).
- 1070 42. Studte, S., Bridger, E. & Mecklinger, A. Sleep spindles during a nap correlate with  
1071 post sleep memory performance for highly rewarded word-pairs. *Brain and Language*  
1072 **167**, 28–35 (2017).
- 1073 43. Miendlarzewska, E. A., Bavelier, D. & Schwartz, S. Influence of reward motivation  
1074 on human declarative memory. *Neuroscience & Biobehavioral Reviews* **61**, 156–176  
1075 (2016).

1076

1077 44. Igloi, K., Gaggioni, G., Sterpenich, V. & Schwartz, S. A nap to recap or how reward  
1078 regulates hippocampal-prefrontal memory networks during daytime sleep in humans.  
1079 *eLife* **4**, e07903 (2015).

1080 45. Lisman, J. E. & Grace, A. A. The Hippocampal-VTA Loop: Controlling the Entry of  
1081 Information into Long-Term Memory. *Neuron* **46**, 703–713 (2005).

1082 46. Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L. & Pennartz, C.  
1083 M. Hippocampus leads ventral striatum in replay of place-reward information. *PLoS*  
1084 *biology* **7**, (2009).

1085 47. Oudiette, D., Antony, J. W., Creery, J. D. & Paller, K. A. The Role of Memory  
1086 Reactivation during Wakefulness and Sleep in Determining Which Memories Endure.  
1087 *Journal of Neuroscience* **33**, 6672–6678 (2013).

1088 48. Baran, B., Daniels, D. & Spencer, R. M. C. Sleep-Dependent Consolidation of Value-  
1089 Based Learning. *PLoS ONE* **8**, (2013).

1090 49. Fischer, S. & Born, J. Anticipated reward enhances offline learning during sleep.  
1091 *Journal of Experimental Psychology: Learning, Memory, and Cognition* **35**, 1586–  
1092 1593 (2009).

1093 50. Bennion, K. A., Payne, J. D. & Kensinger, E. A. The impact of napping on memory  
1094 for future-relevant stimuli: Prioritization among multiple salience cues. *Behavioral*  
1095 *neuroscience* **130**, 281 (2016).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1096 51. Nieuwenhuis, S., Forstmann, B. U. & Wagenmakers, E.-J. Erroneous analyses of  
1097 interactions in neuroscience: A problem of significance. *Nature neuroscience* **14**,  
1098 1105–1107 (2011).
- 1099 52. Diekelmann, S., Wilhelm, I. & Born, J. The whats and whens of sleep-dependent  
1100 memory consolidation. *Sleep Medicine Reviews* **13**, 309–321 (2009).
- 1101 53. Cordi, M. J. & Rasch, B. How robust are sleep-mediated memory benefits? *Current*  
1102 *Opinion in Neurobiology* **67**, 1–7 (2020).
- 1103 54. Button, K. S. *et al.* Power failure: Why small sample size undermines the reliability of  
1104 neuroscience. *Nature Reviews Neuroscience* **14**, 365–376 (2013).
- 1105 55. Lakens, D. & Evers, E. R. K. Sailing From the Seas of Chaos Into the Corridor of  
1106 Stability: Practical Recommendations to Increase the Informational Value of Studies.  
1107 *Perspectives on Psychological Science* **9**, (2014).
- 1108 56. Davidson, P., Jönsson, P., Carlsson, I., & Pace-Schott, E. Does sleep selectively  
1109 strengthen certain memories over others based on emotion and perceived future  
1110 relevance?. *Nature and Science of Sleep* **13**, 1257-1306 (2021).
- 1111 57. Berridge, K. C., Robinson, T. E. Liking, wanting, and the incentive-sensitization  
1112 theory of addiction. *American Psychologist* **71**, 670-679 (2016).
- 1113 58. Oudiette, D., Paller, K. A. Upgrading the sleeping brain with targeted memory  
1114 reactivation. *Trends in Cognitive Sciences*, **17**, 142-149 (2013).
- 1115 59. Hauner, K. K., Howard, J. D., Zelano, C., Gottfried, J. A. Stimulus-specific  
1116 enhancement of fear extinction during slow-wave sleep. *Nature neuroscience*, **16**,  
1117 1553-1555 (2013).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1118 60. Ashton, J. E., Cairney, S. A. Future-relevant memories are not selectively  
1119 strengthened during sleep. *Plos One* **16**, e0258110 (2021).
- 1120 61. Kroneisen, M., Kuepper-Tetzl, C. E. Using day and night–Scheduling retrieval  
1121 practice and sleep. *Psychology Learning & Teaching* **20**, 40-57 (2021).
- 1122 62. Berres, S., Erdfelder, E. The sleep benefit in episodic memory: An integrative review  
1123 and a meta-analysis. *Psychological Bulletin*, **147**, 1309-1353 (2021).
- 1124 63. Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., ... Wittchen, H.  
1125 U. Twelve-month prevalence, comorbidity and correlates of mental disorders in  
1126 Germany: the Mental Health Module of the German Health Interview and  
1127 Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in*  
1128 *Psychiatric Research*, **23**, 304-319 (2014).
- 1129 64. Lipinska, G., Stuart, B., Thomas, K. G., Baldwin, D. S., Bolinger, E. Preferential  
1130 consolidation of emotional memory during sleep: a meta-analysis. *Frontiers in*  
1131 *Psychology* **10**, 1014 (2019).
- 1132 65. Schmid, D., Erlacher, D., Klostermann, A., Kredel, R., Hossner, E. J. Sleep-dependent  
1133 motor memory consolidation in healthy adults: A meta-analysis. *Neuroscience &*  
1134 *Biobehavioral Reviews*, **118**, 270-281 (2020).
- 1135 66. Ellis, B. W. *et al.* The St. Mary's Hospital Sleep Questionnaire: A Study of  
1136 Reliability. *Sleep* **4**, 93–97 (1981).
- 1137 67. Johns, M. W. A New Method for Measuring Daytime Sleepiness: The Epworth  
1138 Sleepiness Scale. *Sleep* **14**, 540–545 (1991).

- 1139 68. Adan, A. & Almirall, H. Horne & Östberg morningness-eveningness questionnaire: A  
1140 reduced scale. *Personality and Individual Differences* **12**, 241–253 (1991).
- 1141 69. Horne, J. A. & Ostberg, O. A self-assessment questionnaire to determine  
1142 morningness-eveningness in human circadian rhythms. *Int J Chronobiol* **4**, 97–110  
1143 (1976).
- 1144 70. Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R. & Kupfer, D. J. The  
1145 pittsburgh sleep quality index: A new instrument for psychiatric practice and research.  
1146 *Psychiatry Research* **28**, 193–213 (1989).
- 1147 71. Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R. & Grant, M.  
1148 Development of the alcohol use disorders identification test (AUDIT): WHO  
1149 collaborative project on early detection of persons with harmful alcohol consumption-  
1150 II. *Addiction* **88**, 791–804 (1993).
- 1151 72. Irons, J. G. Bassett, D. T. Prendergast, C. O. Landrum, R. E. & Heinz, A. J.  
1152 Development and initial validation of the caffeine consumption questionnaire-revised.  
1153 *Journal of Caffeine Research*, **6**, 20-25 (2016).
- 1154 73. Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective  
1155 responses to impending reward and punishment: the BIS/BAS scales. *Journal of*  
1156 *Personality and Social Psychology*, **67**, 319-333 (1994).
- 1157 74. Volk, R. J. Pace, T. M. & Parchman, M. L. Screening for depression in primary care  
1158 patients: Dimensionality of the Short Form of the Beck Depression  
1159 Inventory. *Psychological Assessment*, **5**, 173 -181 (1993).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1160 75. Heydasch, T. Haubrich, J. & Renner, K.-H. Die Kurzform des Hagener  
1161 Matrizen-Tests (HMT-S). Ein 6-Item Intelligenztest zum schlussfolgernden  
1162 Denken. *Methoden, Daten, Analysen*, **7**, 183–208 (2013).
- 1163 76. Hoddes, E., Zarcone, V., Smythe, H., Phillips, R. & Dement, W. C. Quantification of  
1164 sleepiness: A new approach. *Psychophysiology* **10**, 431–436 (1973).
- 1165 77. Dinges, D. I, & Powell, J. W. Microcomputer analysis of performance on a portable,  
1166 simple visual RT task sustained operations. *Behavior Research Methods,*  
1167 *Instrumentation, and Computers*, **17**, 652–655 (1985).
- 1168 78. Lahl, O., Wispel, C., Willigens, B., Pietrowsky, R. An ultra short episode of sleep is  
1169 sufficient to promote declarative memory performance. *Journal of Sleep Research*, **17**,  
1170 3-10 (2008).
- 1171 79. Aschenbrenner, S., Tucha, O. & Lange, K. W. *Regensburger wortflüssigkeits-test:*  
1172 *RWT*. (Hogrefe, Verlag für Psychologie, 2000).
- 1173 80. Bloch, K. E., Schoch, O. D., Zhang, J. & Russi, E. W. German version of the epworth  
1174 sleepiness scale. *Respiration* **66**, 440–447 (1999).
- 1175 81. Randler, C. German version of the reduced morningness–eveningness questionnaire  
1176 (rMEQ). *Biological Rhythm Research* **44**, 730–736 (2013).
- 1177 82. Backhaus J, Riemann D. *Schlafstörungen bewältigen [Coping with sleep disorders]*.  
1178 Weinheim: Beltz Psychologie Verlags Union; 1996.
- 1179 83. Roach, G. D., Dawson, D. & Lamond, N. Can a Shorter Psychomotor Vigilance Task  
1180 Be Used as a Reasonable Substitute for the Ten-Minute Psychomotor Vigilance Task?  
1181 *Chronobiology International* **23**, 1379–1387 (2006).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1182 84. Strobel, A. Beauducel, A. Debener, S. & Brocke, B. (2001). Eine deutschsprachige  
1183 version des BIS/BAS-Fragebogens von carver und white. *Zeitschrift für Differentielle*  
1184 *und Diagnostische Psychologie*, **22**, 216-227 (2001).
- 1185 85. Beck, A. T. Ward, C. H. Mendelson, M, Mock, J. & Erbaugh J. An inventory for  
1186 measuring depression. *Archives of General Psychiatry*. **4**, 561–71 (1961).
- 1187 86. Ujma, P. P. Sleep spindles and general cognitive ability–A meta-analysis. *Sleep*  
1188 *Spindles & Cortical Up States*, **2**, 1-17 (2021).
- 1189 87. Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. Seriousness checks are useful to  
1190 improve data validity in online research. *Behavior Research Methods* **45**, 527-535  
1191 (2013).
- 1192 88. R Core Team. *R: A language and environment for statistical computing*. (R  
1193 Foundation for Statistical Computing, 2019).
- 1194 89. Payne, J. D., Tucker, M. A., Ellenbogen, J. M., Wamsley, E. J., Walker, M. P.,  
1195 Schacter, D. L., Stickgold, R. Memory for semantically related and unrelated  
1196 declarative information: the benefit of sleep, the cost of wake. *PloS One* **7**, e33079  
1197 (2012).
- 1198 90. Macmillan, N. A. & Creelman, C. D. *Detection theory: A user's guide, 2nd ed.*  
1199 (Lawrence Erlbaum Associates Publishers, 2005).
- 1200 91. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in  
1201 linear mixed effects models. *Journal of Statistical Software* **82**, 1–26 (2017).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

- 1202 92. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for  
1203 confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*  
1204 **68**, 255–278 (2013).
- 1205 93. Bates, D., Kliegl, R., Vasishth, S. & Baayen, H. Parsimonious mixed models.  
1206 (2015). *arXiv preprint arXiv:1506.04967* (2015).
- 1207 94. Heisig, J. P. & Schaeffer, M. Why you should always include a random slope for the  
1208 lower-level variable involved in a cross-level interaction. *European Sociological*  
1209 *Review* **35**, 258–279 (2019).
- 1210 95. Bates, D. Mächler, M. Bolker, B. & Walker, S. “Fitting Linear Mixed-Effects Models  
1211 Using lme4.” *Journal of Statistical Software*, **67**, 1–48 (2015).
- 1212 96. Rights, & S., Jason D. Quantifying explained variance in multilevel models: An  
1213 integrative framework for defining r-squared measures. *Psychological Methods* **24**,  
1214 309–338 (2019).
- 1215 97. Brysbaert, M. & Stevens, M. Power analysis and effect size in mixed effects models:  
1216 A tutorial. *Journal of Cognition* **1**, 9 (2018).
- 1217 98. Green, P. & MacLeod, C. J. SIMR: An r package for power analysis of generalized  
1218 linear mixed models by simulation. *Methods in Ecology and Evolution* **7**, 493–498  
1219 (2016).
- 1220 99. Schäfer, S. K. *et al.* Sleep’s impact on emotional memory: A meta-analysis of whole-  
1221 night, nap, and REM sleep effects. *bioRxiv* (2019).

## REWARD AND SLEEP-BASED MEMORY CONSOLIDATION

1222 100. Sterpenich, V., van Schie, M. K., Catsiyannis, M., Ramyeard, A., Perrig, S., Yang, H.  
1223 D., ... Schwartz, S. Reward biases spontaneous neural reactivation during  
1224 sleep. *Nature Communications*, **12**, 1-11 (2021).

1225 101. Lakens, D., Scheel, A. M. & Isager, P. M. Equivalence testing for psychological  
1226 research: A tutorial. *Advances in Methods and Practices in Psychological Science* **1**,  
1227 259–269 (2018).

1228 102. Wixted, J. T. & Wells, G. L. (2017). The relationship between eyewitness confidence  
1229 and identification accuracy: A new synthesis. *Psychological Science in the Public*  
1230 *Interest*, **18**, 10-65 (2017).

1231

1232 **Data availability**

1233 The authors commit to sharing the raw data and materials on acceptance of our Stage 2  
1234 manuscript.

1235 **Code availability**

1236 The authors commit to sharing all code on acceptance of our Stage 2 manuscript.

1237 **Acknowledgements**

1238 This work was supported by a grant from the German Research Foundation to GBF  
1239 (Deutsche Forschungsgemeinschaft GZ: FE 1617/2-1). The funders have/had no role in study  
1240 design, data collection and analysis, decision to publish or preparation of the manuscript.

1241 **Competing interests**

1242 The authors declare no competing interests.