http://dx.doi.org/10.1111/jsr.12457
Sleep-dependent memory consolidation is related to perceived value of learned material

ELAINE VAN RIJN, CARLO LUCIGNOLI, CRISTINA IZURA and MARK T. BLAGROVE
Department of Psychology, Swansea University, Swansea, UK

SUMMARY
Although many types of newly encoded information can be consolidated during sleep, an enhanced effect has been found for memories tagged as relevant to the future, such as through knowledge of future testing or payment for successful recall. In the current study, participants (n = 80) learned Welsh and Breton translations of English words, and intrinsic relevance of learned material was operationalized as perceived value of the Welsh and Breton languages. Participants were non-Welsh native English speakers who had recently arrived in Wales. Memory for the words was tested immediately and 12 h later, after either a period of wake or a period of sleep. An increase in recall for both languages was found after sleep, but not after wake. Importantly, for the sleep condition, overnight improvement in Welsh word recall was associated with participants’ level of valuing the Welsh language. This association was not found for the wake period condition. These findings support previous indications of an active role of sleep in the consolidation of memories relevant for the future, and demonstrate that this effect may be modulated by individual differences in perceived value of the learned material. It remains to be established whether this association is mediated by an emotional attachment to the language or a cognitive facility with it, or both.

INTRODUCTION
Sleep is believed to play an active role in the consolidation of newly acquired memories (Diekelmann and Born, 2010; Ellenbogen et al., 2006; Smith, 2001). During sleep-dependent memory consolidation, recent, initially unstable information is transformed into more stable representations that integrate into long-term memory networks (Diekelmann and Born, 2010). However, not all information encoded during the day is retained in memory. Instead, a process of sleep-dependent memory triage has been proposed, in which memories are processed differentially based on salience tags attached during or after encoding of new information. It has been suggested that memories that are tagged as useful undergo preferential consolidation during sleep (Stickgold and Walker, 2013). Sleep-dependent memory consolidation is thus modulated by the perceived importance of the learned material, which acts as a mechanism by which memories with future relevance are stabilized and strengthened preferentially (van Dongen et al., 2012). This discriminatory selection is necessary for rapid and effective adaptation to changes in the environment (Stickgold and Walker, 2013).

One salience factor that determines selective consolidation during sleep is knowledge of future testing (van Dongen et al., 2012; Wilhelm et al., 2011). Wilhelm et al. (2011) found that sleep improved memory performance compared to wake only when testing was expected. Memory formation also benefits from implicit emotional salience cues that may be relevant for survival (Payne et al., 2008). For example, memory for emotional stimuli, such as emotional texts (Wagner et al., 2001) and pictures containing emotional objects (Payne et al., 2008), has been shown to benefit from sleep more than neutral stimuli. However, other studies have found improved memory for neutral as well as negative stimuli, following sleep relative to wake (e.g. Baran et al., 2012). Sleep also consolidates information selectively when a potential monetary reward is expected for remembering (Feld et al., 2014; Fischer and Born, 2009); this effect has been proposed to be related to feedback from the reward system to the hippocampus (Feld et al., 2014). However,
although Tucker et al. (2011) found improvement in declarative memory following sleep, this effect was not enhanced by a monetary reward. Furthermore, sleep consolidates information when there are instructions for items to be remembered, compared to instructions for items to be forgotten (Rauchs et al., 2011; Saletin et al., 2011). Conversely, sleep has also been found to reduce or prevent forgetting (e.g. Abel and Bäuml, 2013).

These characteristics of sleep-dependent memory consolidation suggest a hierarchy of salience in which the most salient memories are processed and stored (Wilhelm et al., 2011). Material that is expected to be recalled later benefits specifically from slow wave sleep (SWS), which is linked to hippocampal-dependent memory consolidation (Wilhelm et al., 2011). A proposed mechanism is that important information is replayed during sleep to stimulate the redistribution of these representations from the hippocampus to the neocortex for long-term storage (Diekelmann and Born, 2010). This replay can also be induced experimentally. Memories associated with a low reward during encoding, for which recall accuracy declined following a nap, have been found to be rescued from forgetting when cued during sleep using targeted memory reactivation (Oudiette et al., 2013). Hippocampal activity is higher during encoding for items targeted to be remembered compared to forgotten, suggesting that this activity during learning triggers preferential memory consolidation during sleep (Rauchs et al., 2011).

The present study aimed to examine the selective sleep-dependent consolidation of memories perceived as valuable but, in contrast to previous work, used an individual difference rather than an experimentally induced salience variable. Native English-speaking university students, recently arrived in Wales and not having lived previously in Wales or learned Welsh, were taught Welsh and Breton translations of English words. Memory for the translations was tested after a period of sleep or wake. Memory was hypothesized to improve across a period of sleep for translations of both languages, as participants expected retrieval testing, but to not change for either language across a period of wake. Crucially, within the sleep group, participants’ perceived value of the Welsh language was hypothesized to correlate with Welsh memory improvement across sleep, whereas this correlation was not predicted to be significant for the wake condition.

MATERIALS

The Stanford Sleepiness Scale

Participants completed the Stanford Sleepiness Scale (SSS) (Hoddes et al., 1973) at the start of session 1 (learning and immediate recall) and session 2 (delayed recall). This scale assesses how alert one is feeling at that moment on a scale from 1 (feeling active, vital, alert or wide awake) to 7 (no longer fighting sleep, sleep onset soon; having dream-like thoughts).

Stimuli

Twenty-eight written English words were selected from the Medical Research Council (MRC) Psycholinguistic Database (Wilson, 1988). All words used were selected based on their familiarity rating, concreteness rating and number of letters. To reduce the likelihood that participants had encountered any of the Welsh translations of the English words in their daily lives, highly familiar words were excluded. All words had a familiarity score between 321 and 499 (on a rating scale of 0–700). To control for concreteness of the words, words with a concreteness score lower than 300 (on a rating scale of 0–700) were excluded. Half the words had a concreteness score of between 338 and 472, the other half had a concreteness score of between 515 and 631. Seventeen words were five letters long and 11 words had six letters. All words were paired with their respective Welsh and Breton translations. Breton was chosen because, like Welsh, it is a Brythonic language. Brythonic is one of two groups of the modern Celtic languages. The words were presented on a

© 2016 The Authors. Journal of Sleep Research published by John Wiley & Sons Ltd on behalf of European Sleep Research Society.
5 × 6 grid, using a K010 (TF103CX) Asus tablet with a 10.1-inch screen, using the Android operating system. The experiment was programmed as an Android application. The code was written in Java, which was converted into Android using the Eclipse Mars (version 4.5.0) software and the Android Studio (version AI-141.2017176) software. The background colour of the screen was grey, or red–green–blue (RGB) colour model value (199, 199, 199).

Procedure
Participants participated in two sessions spaced 12 h apart for each participant, both taking place at home. Participants in the sleep group completed session 1 between 20:30 and 22:30 hours in the evening, whereas participants in the wake group completed session 1 between 08:30 and 10:30 hours in the morning. Participants were instructed to complete both parts by themselves in a quiet environment, with their mobile phones on silent. Before each session, participants reported how many hours they had slept in the previous 12 h. This was to ensure that participants in the wake group had not napped between the two sessions, and that the participants in the sleep group had slept for at least 6 h between the two sessions. During session 1, all words were learned and immediate cued recall was performed. During session 2, 12 h later, participants performed delayed cued recall.

Learning
In session 1, participants learned the Welsh and Breton translations of 28 English words. The translated language of each word was selected randomly for each participant, resulting in 14 of the English words being paired with the Welsh translation and 14 words being paired with the Breton translation. To distinguish between the two languages, participants were informed that English words with the Welsh translation would be written on screen in red, whereas English words with the Breton translation would be written in blue. Participants were instructed to remember the translation of all the words. All words were presented one-by-one on the squares of a grid. For each participant separately, location of the words on the grid was determined randomly, with each word appearing on a different square. Two words never appeared on the same square. Because there was a total of 30 squares but 28 words, two squares were empty, determined randomly for each participant. When the first written English word appeared participants pressed its square on the grid, after which the written translation appeared underneath the English word and the spoken audio of the corresponding translation was played. Each word had to be pressed four times, and the audio was thus heard four times before the word disappeared and the next word appeared. The interval between words was 2 s. This step was repeated until all 28 words and their translations had been presented. This learning phase was performed three times. No feedback was given on performance. Button press time was monitored to ensure that all participants spent the same amount of time learning the words.

Testing
Cued recall was performed immediately after learning, and exactly 12 h later. For both cued recall tests, participants were presented with each translation, one-by-one, written in black on the middle of the screen. They were instructed to type the corresponding English word. All words were presented in a random order and there was no time limit for typing each response.

Statistical analyses
To analyse the difference in cued recall between the wake and sleep groups, the number of correctly remembered Welsh and Breton translations before and after the 12-h delay was calculated for each participant. Responses were scored by a judge blind to both group and session. Responses were scored as correct if they were identical to the target or the target word was referred to clearly. All other responses, including synonyms or misspellings where a different word was produced, were scored as incorrect. The data were then analysed using a 2 × 2 × 2 mixed analysis of variance (ANOVA), with session (before versus after the 12-h delay) and language (Welsh versus Breton) as within-subject variables, and group (sleep versus wake) as between-subjects variable. Paired-samples and independent-samples t-tests were used to analyse any significant interactions. For effect size measures, partial eta squared ($\eta_p^2$) was calculated for ANOVA results and Cohen’s $d$ for t-test results.

Change scores were calculated for the number of correctly remembered Welsh and Breton words across the 12-h period of sleep or wake. Pearson’s $r$ correlational analyses were then performed for the sleep and wake groups separately, to examine the relationship between change in participants’ scores and how much they value the Welsh and Breton languages. A significant correlation was predicted only for change in recall of Welsh words across sleep and level of valuing the Welsh language, and not for the other correlations, and so no correction for multiple correlations was performed. A stepwise multiple regression analysis was performed to examine further if overnight change in Welsh words cued recall was predicted by valuing Welsh, valuing Breton, sleepiness at immediate testing and sleepiness at delayed testing.

RESULTS
Testing-related variables
One participant in the sleep group indicated consuming alcohol during the experiment; data from this participant were excluded from the analyses. Means and standard deviations of duration of the two sessions, time between the two
sessions and the SSS scores, for the wake and sleep groups, are presented in Table 1.

The groups did not differ in time between the two sessions ($t_{(77)} = -0.23$, $P = 0.82$), time to complete session 1 ($t_{(77)} = 0.22$, $P = 0.83$) or time to complete session 2 ($t_{(77)} = 0.04$, $P = 0.97$). For the SSS, there was no significant main effect of session ($F_{(1,77)} = 0.01$, $P = 0.91$) and no significant main effect of group ($F_{(1,77)} = 0.16$, $P = 0.69$). The interaction between session and group approached significance ($F_{(1,77)} = 3.55$, $P = 0.06$). The means indicate that sleepiness decreased at delayed testing for the wake group, whereas for the sleep group sleepiness increased at delayed testing. The two groups did not differ in terms of how many months participants had been living in Wales ($t_{(77)} = -0.14$, $P = 0.90$), nor in how many Welsh words ($t_{(77)} = 1.20$, $P = 0.23$) or Breton words ($t_{(77)} = 0.99$, $P = 0.33$) they knew at screening (see Table 1). Furthermore, no gender differences were found for months living in Wales ($t_{(77)} = -1.50$, $P = 0.14$) or in how many Welsh words ($t_{(77)} = -0.19$, $P = 0.85$) or Breton words ($t_{(77)} = -0.96$, $P = 0.34$) were known at screening.

Cued recall

Means and standard deviations of the cued recall measures for the wake and sleep groups are presented in Table 2. The three-way interaction (session × language × group) was not significant ($F_{(1,77)} = 0.26$, $P = 0.61$). There was a significant main effect of session ($F_{(1,77)} = 9.43$, $P = 0.003$, $\eta_p^2 = 0.11$), demonstrating that significantly more words were recalled after the delay than before [mean cued recall: session 1 = 13.49 (SD = 6.39), session 2 = 14.34 (SD = 6.80)]. There were no significant main effects for language ($F_{(1,77)} = 1.70$, $P = 0.20$) or group ($F_{(1,77)} = 0.05$, $P = 0.83$). The interaction between session and group was significant ($F_{(1,77)} = 12.37$, $P = 0.001$, $\eta_p^2 = 0.14$). All other two-way interactions were not significant (all $P_s > 0.53$). Paired-samples t-tests demonstrated that the sleep group remembered significantly more Welsh words ($t_{(38)} = 3.18$, $P = 0.003$, $d = 0.21$) and Breton words ($t_{(38)} = 3.93$, $P < 0.001$, $d = 0.30$) after the 12-h delay compared to before, whereas there was no change in Welsh recall ($t_{(38)} = -0.34$, $P = 0.75$) or Breton recall ($t_{(38)} = -0.18$, $P = 0.86$) for the wake group. Independent-samples t-tests comparing overnight change score in Welsh and Breton recall between the sleep and wake groups demonstrated a significant increase for the sleep group in Welsh recall ($t_{(28)} = 2.61$, $P = 0.011$, $d = 0.59$) and Breton recall ($t_{(28)} = 2.87$, $P = 0.005$, $d = 0.65$) compared to the wake group. Mean change scores are presented in Table 2. There were no gender differences in cued recall or change in cued recall measures (all $t_s < 1.26$ and all $P_s > 0.21$).

Value of language

Welsh was rated as significantly more valuable than Breton [$F_{(1,77)} = 63.82$, $P < 0.001$, $\eta_p^2 = 0.45$, mean valuing of Welsh = 3.27 (SD = 0.87), mean valuing of Breton = 2.54 (SD = 0.98)]. There was no group difference ($F_{(1,77)} = 0.12$, $P = 0.73$) or a significant interaction between language and group ($F_{(1,77)} = 0.27$, $P = 0.60$) for valuing Welsh and Breton, and no gender difference in valuing Welsh or Breton ($F_{(1,77)} = 0.20$, $P = 0.66$) or a significant interaction between language and gender ($F_{(1,77)} = 0.27$, $P = 0.60$).

Correlational analyses

Pearson’s $r$ correlations were performed to assess the relationship between participants’ level of valuing the Welsh or Breton language and cued recall scores (immediate, delayed and change scores over the 12-h period). Correlation

| Table 1 | Means and standard deviations of testing-related measures for the wake and sleep groups |
|------------------|------------------|------------------|
| **Measure** | **Wake** | **Sleep** |
| Mean time (decimal hours) between sessions | 12.10 (0.20) | 12.11 (0.18) |
| Mean time (decimal hours) to complete session 1 | 0.46 (0.12) | 0.45 (0.14) |
| Mean time (decimal hours) to complete session 2 | 0.17 (0.07) | 0.17 (0.12) |
| Mean sleepiness score at immediate testing$^*$ | 2.55 (0.75) | 2.36 (0.81) |
| Mean sleepiness score at delayed testing$^*$ | 2.30 (1.09) | 2.64 (1.40) |
| Months living in Wales | 10.63 (11.34) | 10.95 (11.72) |
| Number of Welsh words known | 3.55 (4.16) | 2.64 (2.29) |
| Number of Breton words known | 0.03 (0.16) | 0.00 (0.00) |

$^*$Stanford Sleepiness Scale score.

| Table 2 | Means and standard deviations for the immediate and delayed cued recall measures for the wake and sleep groups |
|------------------|------------------|------------------|
| **Language** | **Immediate cued recall** | **Delayed cued recall** | **Change score** |
| **Wake** | **Sleep** | **Wake** | **Sleep** |
| Welsh | 7.18 (3.27) | 6.77$^*$ (3.80) | 0.46 (1.40) | 0.79$^*$ (1.56) |
| Breton | 6.65 (3.51) | 6.38$^*$ (3.53) | -0.05$^*$ (1.74) | 1.05$^*$ (1.67) |

$^*$ $P = 0.003$; $^1P < 0.001$; $^2P = 0.011$; $^§P = 0.005$.© 2016 The Authors. Journal of Sleep Research published by John Wiley & Sons Ltd on behalf of European Sleep Research Society.
The present study aimed to examine the effect of sleep on the consolidation of memories and the association between this consolidation and level of valuing the learned material. An increase in memory for both Welsh and Breton words was found after sleep, but not after wake. Importantly, overnight increase in Welsh cued recall was correlated with how much participants valued the Welsh language, whereas no such correlation was found for the wake group. These findings accord with previous research demonstrating that sleep is involved in consolidating memory (e.g. Diekelmann and Born, 2010; Ellenbogen et al., 2006; Smith, 2001), but furthers that work by suggesting that the motivated selectivity of which memories are processed (Stickgold and Walker, 2013) can extend to individual differences in the value that participants attribute to the material being learned. This association was, however, found between valuing Breton and change in recall of Breton words. The mean valuing of Breton was low, and significantly lower than for Welsh. The absence of a significant correlation between valuing Breton and Breton overnight improvement is interpreted by us as due to a floor effect of valuing Breton.

Regression analysis

A stepwise multiple regression analysis was conducted to examine further whether overnight Welsh change score was predicted by valuing Welsh, valuing Breton, sleepiness at immediate testing and sleepiness at delayed testing (SSS scores). The model was significant ($F_{(1,38)} = 4.83$, $P = 0.01$). Overnight Welsh change score was predicted by valuing Welsh ($P = 0.01$) and by sleepiness at immediate testing ($P = 0.04$). The results from the regression are presented in Table 4.

**DISCUSSION**

coefficients are presented in Table 3. For the sleep group only there was a significant positive correlation between valuing the Welsh language and change score of Welsh cued recall ($r = 0.34$, $n = 39$, $P = 0.04$), indicating that a higher valuing of Welsh was associated with an increase in number of Welsh words recalled correctly after the 12-h delay. This correlation was not significant for the wake group ($r = -0.02$, $n = 40$, $P = 0.89$). There was no significant correlation between valuing Breton and change in Breton words for the wake or sleep group. No significant correlations were found between valuing either language and scores on immediate or delayed cued recall for either group.

Regression analysis

A stepwise multiple regression analysis was conducted to examine further whether overnight Welsh change score was predicted by valuing Welsh, valuing Breton, sleepiness at immediate testing and sleepiness at delayed testing (SSS scores). The model was significant ($F_{(1,38)} = 4.83$, $P = 0.01$). Overnight Welsh change score was predicted by valuing Welsh ($P = 0.01$) and by sleepiness at immediate testing ($P = 0.04$). The results from the regression are presented in Table 4.

**Table 3 Correlations between valuing Welsh or Breton and cued recall scores for the wake ($n = 40$) and sleep ($n = 39$) groups**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wake</th>
<th>Sleep</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1. Value Welsh</td>
<td>0.16</td>
<td>-0.01</td>
</tr>
<tr>
<td>2. Welsh immediate cued recall</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td>3. Welsh delayed cued recall</td>
<td>0.91**</td>
<td>0.92**</td>
</tr>
<tr>
<td>4. Change score Welsh</td>
<td>-0.02</td>
<td>-0.12</td>
</tr>
<tr>
<td>5. Value Breton</td>
<td>0.56**</td>
<td>0.56**</td>
</tr>
<tr>
<td>6. Breton immediate cued recall</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>7. Breton delayed cued recall</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>8. Change score Breton</td>
<td>0.32*</td>
<td>0.32*</td>
</tr>
</tbody>
</table>

* $P < 0.05$; ** $P < 0.01$.

**Table 4 Predictors of overnight Welsh change score for the sleep group**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$B$</th>
<th>SE $B$</th>
<th>$\beta$</th>
<th>$P$</th>
<th>$r_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-3.00</td>
<td>1.24</td>
<td></td>
<td>0.01</td>
<td>0.40</td>
</tr>
<tr>
<td>Value Welsh</td>
<td>0.72</td>
<td>0.28</td>
<td>0.39</td>
<td>0.01</td>
<td>0.40</td>
</tr>
<tr>
<td>Sleepiness at immediate testing</td>
<td>0.61</td>
<td>0.29</td>
<td>0.32</td>
<td>0.04</td>
<td>0.33</td>
</tr>
</tbody>
</table>

The dependent variable was overnight Welsh change score. $R^2 = 0.21$; adjusted $R^2 = 0.17$; $r_p$ is the partial correlation.
We acknowledge that there was no significant difference between the languages in performance change across the night, even though Welsh was valued more than Breton. This may be because the effect size for the association between Welsh language valuing and change in performance, although significant, is smaller than the across-night effect. This accords with the findings of Bennion et al. (2016), that sleep prioritizes top-down, task goal cues (for example, expected testing) over bottom-up, stimulus-driven cues (in their study, emotion associated with cues). In the present study, the top-down cue of expectation of retrieval testing may serve as a more important determinant of sleep-dependent consolidation than does the bottom-up stimulus-driven cue of valuing Welsh.

For limitations of the study, it is possible that the cued recall results may have been confounded by circadian effects, in that learning and testing did not occur at the same times for both groups. However, this confound has not been found in previous studies that have tested for it (e.g. Payne et al., 2008; Walker et al., 2002), and sleepiness scores did not differ significantly between the groups at learning or testing in the present study. A second possible confound is that the wake group was exposed to more interference than was the sleep group as a result of waking life activities. However, if sleep only protects memories from interference, a decrease in memory following wake together with a stabilization of memory following sleep would be expected. The present finding of an increase in memory performance following sleep favours an active role of sleep in the consolidation of memory, rather than a passive stabilizing function. Furthermore, as the main finding here is the correlation between valuing Welsh and overnight improvement in Welsh learning, it is not clear how this could be accounted for by any circadian or interference confound.

In the present study, the colours in which the words from the two languages were presented were the same for all participants (Welsh presented in red and Breton presented in blue). Although this arrangement of colours in the present study would not explain the correlation between valuing Welsh and overnight increase in memory performance, for future studies it would be recommended to counterbalance the presentation colours across the participants. Furthermore, the present study used a between-subjects design, using one list of English words. Future work could consider using a within-subjects design in which each participant completes the wake and sleep condition, which would require having two lists of English words. This would remove interindividual differences in learning. However, consideration would be needed for the feasibility of producing two matched lists, given the multiple constraints required on English words used and their translations.

To summarize, there is considerable evidence that recently acquired memory traces are not all consolidated equally during sleep (Rasch and Born, 2013). The findings here indicate a sophisticated selectivity and processing of memories during sleep as a function of value ascribed to the learned material. The findings support previous indications of an active role of sleep in the consolidation of memories perceived as valuable and relevant (Stickgold and Walker, 2013), and extend that work to an association with individual differences in value attributed to a minority language by majority language speakers. Two non-exclusive mechanisms could be proposed for this association. There may be an emotional attachment to the language and its history and country, or there may be a cognitive facility with the language, including a liking for its orthography and sound (Csizér and Dőrményi, 2005; Masgoret and Gardner, 2003; Mori and Calder, 2015). It may be that these mechanisms overlap and have their own emotional and cognitive components. Future work should address in what ways the language is valued, or not valued, by participants, and the emotional and cognitive components of this valuing, in order to elucidate the mechanism by which this valuing is associated with sleep-dependent memory consolidation.

ACKNOWLEDGEMENTS

We thank Dr Cynfael Lake and Dr Rhisiart Hinks for their assistance with the Welsh and Breton translations of the English words. This work was supported by a PsyPAG research grant bursary.

AUTHOR CONTRIBUTIONS

EVR and MTB developed the study concept; EVR, MTB and CI contributed to the study design; CL programmed the experiment; EVR collected and analysed the data; EVR drafted the manuscript and MTB edited it.

CONFLICT OF INTEREST

No conflicts of interest declared.

REFERENCES


