Memory effects of sleep, emotional valence, arousal and novelty in children

MARIJE C. M. VERMEULEN[1,2], KRISTIAAN B. VAN DER HEIJDEN[2,3], JEROEN S. BENJAMINS[1,4], HANNA SWAAB[2,3] and EUS J. W. VAN SOMEREN[1,5]

[1]Department of Sleep and Cognition, Netherlands Institute for Neuroscience, Royal Netherlands Society for Arts and Sciences, Amsterdam, The Netherlands; [2]Department of Clinical Child and Adolescent Studies, Institute of Education and Child Studies, Leiden University, Leiden, The Netherlands; [3]Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands; [4]Departments of Experimental Psychology and Social, Health and Organisational Psychology, Utrecht University, Utrecht, The Netherlands; [5]Departments of Integrative Neurophysiology and Psychiatry, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University & Medical Center, Amsterdam, The Netherlands

Keywords
emotional stimuli, explicit memory, learning, memory retention, preadolescents, word-pair recognition

Correspondence
Eus J. W. van Someren, PhD, Department of Sleep and Cognition, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, The Netherlands.
Tel.: +31 20 5665497; fax: +31 20 5666121; e-mail: e.van.someren@nin.knaw.nl

Accepted in revised form 23 December 2016; received 24 June 2016

DOI: 10.1111/jsr.12506

SUMMARY

Effectiveness of memory consolidation is determined by multiple factors, including sleep after learning, emotional valence, arousal and novelty. Few studies investigated how the effect of sleep compares with (and interacts with) these other factors, of which virtually none are in children. The present study did so by repeated assessment of declarative memory in 386 children (45% boys) aged 9–11 years through an online word-pair task. Children were randomly assigned to either a morning or evening learning session of 30 unrelated word-pairs with positive, neutral or negative valenced cues and neutral targets. After immediately assessing baseline recognition, delayed recognition was recorded either 12 or 24 h later, resulting in four different assessment schedules. One week later, the procedure was repeated with exactly the same word-pairs to evaluate whether effects differed for relearning versus original novel learning. Mixed-effect logistic regression models were used to evaluate how the probability of correct recognition was affected by sleep, valence, arousal, novelty and their interactions. Both immediate and delayed recognition were worse for pairs with negatively valenced or less arousing cue words. Relearning improved immediate and delayed word-pair recognition. In contrast to these effects, sleep did not affect recognition, nor did sleep moderate the effects of arousal, valence and novelty. The findings suggest a robust inclination of children to specifically forget the pairing of words to negatively valenced cue words. In agreement with a recent meta-analysis, children seem to depend less on sleep for the consolidation of information than has been reported for adults, irrespective of the emotional valence, arousal and novelty of word-pairs.

INTRODUCTION

Several studies suggest that sleep facilitates consolidation of declarative memory. The recall of previously acquired information following sleep was found to be superior to retention following an equivalent period of wakefulness in studies on adults (Diekelmann *et al.*, 2009; Gais *et al.*, 2006; Marshall and Born, 2007; Payne *et al.*, 2012; Rasch and Born, 2013), and children and adolescents (Ashworth *et al.*, 2014; Backhaus *et al.*, 2008; Henderson *et al.*, 2012; Potkin and Bunney, 2012; Prehn-Kristensen *et al.*, 2009; Wilhelm *et al.*, 2008). It has hardly been investigated, however, how effects of sleep compare with (and might interact with) other factors known to affect consolidation, notably the emotional valence, arousal and novelty properties of the acquired information. The overarching aim of the present study was to do so specifically for declarative learning and memory in primary school-aged children.

Adults remember emotional and/or arousing stimuli better than neutral stimuli, but this advantage has only scarcely been studied in children (Kensinger, 2004; Kensinger and Corkin, 2003; LaBar and Cabeza, 2006; Prehn-Kristensen *et al*., 2009). The first aim of the present study is therefore to evaluate the effects of emotional valence and arousal on declarative memory consolidation in children. Sleep may further add to the long-term consolidation advantage that emotional stimuli have over neutral stimuli, although the exact interaction effect remains somewhat equivocal. Some studies in adults showed that (rapid eye movement) sleep primarily enhances consolidation of emotional text (Wagner *et al*., 2001), pictures (Hu *et al*., 2006) and objects (Payne *et al*., 2008). In contrast, Atienza and Cantero (2008) found no interaction effect of sleep and emotional properties of the stimuli. They suggested an independent, additive modulation of memory consolidation by sleep, valence and arousal. Likewise, Ackermann *et al.* (2015) found that sleep benefits memory consolidation for pictures independent of their valence. Studies on the interaction effect of sleep and emotion on memory consolidation in children and adolescents are scarce. In healthy boys aged 9–13 years, a beneficial effect of sleep compared with wake on recognition memory occurred only for emotionally negative pictures and not for neutral pictures (Prehn-Kristensen *et al*., 2009, 2013). Given this scarcity, the second aim of the present study is to compare, in children, the interactive effect of sleep versus wake on the consolidation of neutral versus emotionally valenced information.

Whereas valence and arousal are commonly recognized as two dissociable properties of emotional stimuli, recent studies identified a third property of importance: novelty. The *relative* novelty, even of rather common stimuli, impacts the activation of circuits involved in learning, memory and emotion, including hippocampus and amygdala at least as much as valence and arousal properties do (Balderston *et al*., 2011; Blackford *et al*., 2010; Weierich *et al*., 2010). Altered brain activation with novelty detection has mostly been studied for familiar and novel stimuli *within* the same task (for review, see Murray *et al*., 2014). Task-related brain activation, however, changes as well when a task is practiced. For example, relative to when a working memory task was first presented, its practicing lowered task-related activation in extended fronto-parietal circuits, including the anterior cingulate cortex, left inferior frontal gyrus, left precentral/dorsolateral prefrontal cortex and left superior parietal cortex (Jager *et al*., 2010). This suggests that practicing decreases the dependence on attentional and working memory resources (Tomasi *et al*., 2004). Within-task item novelty versus whole-task practice-related decreases in novelty has discriminable effects on brain activation (Basso *et al*., 2013). It has been recognized that task novelty-related brain activation is relevant to classroom learning (Almarode and Miller, 2013). A possible differential dependency on the interval between learning and sleep of new material versus previously practiced material would be highly relevant for classroom schedules. However, the role of this novelty property has remained unexplored in studies on the modulatory effect of emotion on sleep-dependent memory consolidation. To compare novel learning and relearning seems highly relevant to the mechanistic understanding of the role of sleep in memory consolidation, not in the least because relearning *during* sleep is the cornerstone of the important 'reactivation' model on the role of sleep in memory consolidation. This model posits that during sleep, newly formed memory traces are reactivated, reinforced and reorganized (Born *et al*., 2006; Diekelmann and Born, 2010; Ellenbogen *et al*., 2006; Stickgold and Walker, 2007). In spite of these widely adopted concepts, few studies addressed possible similarities or differences of relearning during sleep with the relearning that we commonly do during wakefulness (De Jonge and Tabbers, 2013; Roediger and Karpicke, 2006); the very basis of our teaching methods. Also, it is not known whether sleep still supports consolidation if information has been reconsolidated already during wakefulness. A third aim of the present study is to compare, in children, the effects of sleep on the consolidation of novel information versus information that has been relearned during wakefulness.

MATERIALS AND METHODS

Participants

Participants were recruited from 53 primary schools (participation rate 36%) across The Netherlands in the periods January–June 2012 and February–July 2013. These two data waves were aggregated. Children previously diagnosed with dyslexia as reported by parents ($n = 10$) and children not meeting the age criteria ($n = 12$) were excluded, resulting in the final sample size of 386 children (45% boys, 95% Dutch nationality) aged 9–11 years (M = 10.5, SD = 0.8). These children completed the initial word-pair learning and immediate recognition (IR) task, of whom 291 also completed the delayed recognition (DR) assessment. Of these, 207 continued participation for a second assessment and 161 finished the second DR assessment. Children's estimated average sleep duration based on a 7-day sleep diary was 9 h and 59 min (SD = 34 min). The study was approved by the ethics committee of Leiden University, and informed consent was obtained for all participants.

Procedure

Parents and children filled out questionnaires and a 7-day sleep diary using the Netherlands Sleep Registry (NSR) internet platform (www.sleepregistry.nl). Within the second part of the research week (days 5 and 6 or days 6 and 7), children performed the first two sessions of a word-pair memory task at home. For the first session, children were randomly assigned to learn word-pairs between 07:00 and

08:00 hours, or between 19:00 and 20:00 hours. Recognition performance was assessed both immediately after learning (IR) and after 12 or 24 h (DR; balanced across children). The randomization yielded four conditions: '12 h-Wake' (learning and IR in the morning, DR the same evening, $n = 94$); '12 h-Sleep' (learning and IR in the evening, DR the next morning, $n = 86$); '24 h-Wake first' (learning and IR in the morning, DR the next morning, $n = 102$); '24 h-Sleep first' (learning and IR in the evening, DR the next evening, $n = 104$; Fig. 1). These four conditions were used to evaluate potential time-of-day influences. After a week, the procedures were repeated. Accordingly, the word-pair memory task consisted of four sessions: IR-1 and DR-1 (first week) and IR-2 and DR-2 (second week).

Instruments

Declarative memory was assessed using semantically unrelated word-pairs. Words were selected from Dutch nouns and adjectives using the list of Hermans and De Houwer (1994) containing 740 words with their emotional valence and familiarity ratings. Low relatedness of all selected pairs was verified using the Dutch word association project (De Deyne and Storms, 2008). Two sets of 30 word-pairs were created. In each set, 10 cue words had a positive valence (e.g. gift, funny), 10 words a negative valence (e.g. war, accident) and 10 words a neutral valence (e.g. paper, chair). Each of these was paired with a target word of neutral valence. The two sets of word-pairs were counterbalanced across participants and did not differ significantly with respect to emotional valence ratings for positive ($P = 0.285$), negative ($P = 0.984$) and neutral ($P = 0.948$) words, word length ($P = 0.164$) or familiarity ($P = 0.232$).

In the first session, children were instructed to learn one set of 30 word-pairs to the best of their ability because their memory would be tested immediately after presentation of the word-pairs. They were also told that the task consisted of more parts, but the exact content of the other sessions was not explained yet. In the learning phase, the word-pairs were presented visually for 2 s each on the computer screen. After each word-pair, the children indicated the emotional feeling they experienced for the cue word by selecting one out of seven pictograms (valence: 1 very unpleasant to 7 very pleasant), adapted from the Self-Assessment Manikin (SAM; Bradley and Lang, 1994; Fig. 2a). They likewise rated the emotional intensity they experienced (arousal: 1 least

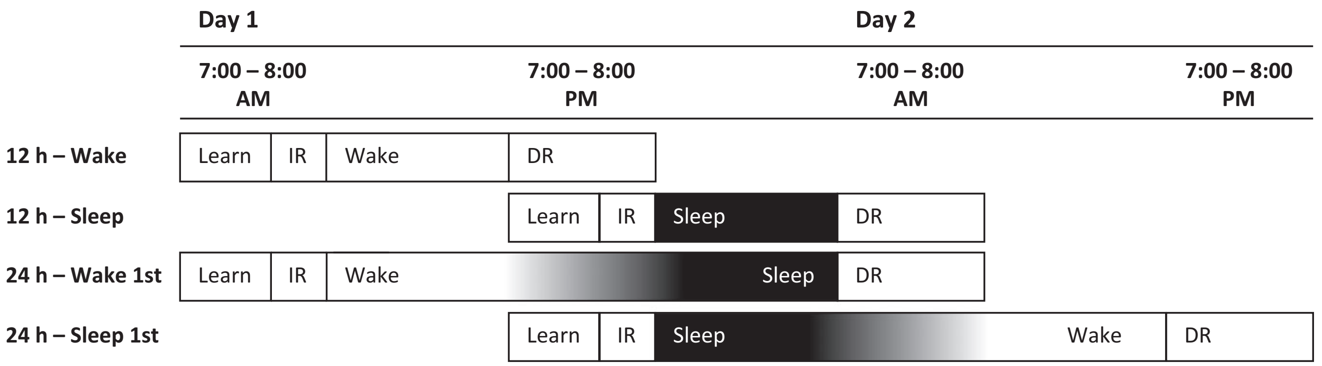

Figure 1. Study design showing the four different conditions children were randomized to. Either in the morning, or in the evening, children learned 30 unrelated word-pairs, directly followed by an immediate recognition (IR) assessment. After either 12 or 24 h, a delayed recognition (DR) assessment took place. The design is repeated a week after.

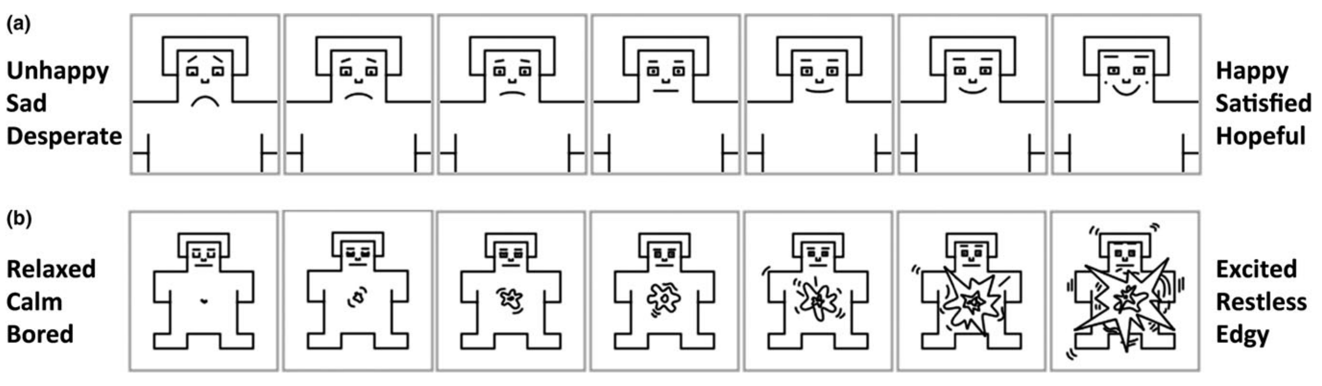

Figure 2. Adapted version of the Self-Assessment Manikin (SAM; Bradley and Lang, 1994). After each word-pair, children indicated the emotional feeling they experienced for the cue word by selecting one out of seven pictograms (a: valence: very unpleasant to very pleasant). They likewise rated the emotional intensity they experienced (b: arousal: least arousing to very arousing).

arousing to 7 very arousing; Fig. 2b). After completing these self-paced judgements, the screen was blank for 3 s before the next word-pair appeared. The word-pairs were presented once in quasi-randomized order; limiting the number of consecutive cue words with similar emotional valence to two. After this learning phase, IR was assessed by presenting the 30 cue words in quasi-randomized order. For each cue word, children had to identify the correct target presented among two foils of neutral valence. These foils were randomly selected from the remaining incorrect target words. DR was assessed 12 or 24 h later by first presenting the cue words in quasi-randomized order for 2 s, then asking children to rate their valence and arousal using the SAM, and finally asking them to identify the correct target presented among two foils. The recognition task of Session 2 (DR without learning) is similar to Session 1 (IR), using the same word-pairs but different foils. Responses on IR and DR were self-paced. After finishing the sessions, feedback was given on the number of correctly remembered word-pairs.

Statistical analyses

Planned analyses utilized three-level random intercept mixed-effect logistic regression models to evaluate whether the probability of correctly remembering a word-pair was predicted by sleep; by the valence, arousal and novelty of the cue word; and by the interaction of sleep with these cue word properties. Mixed-effect analyses take into account the interdependence of the data points inherent to the hierarchical structure of our data: words (level 1 or i) are nested within sessions in week 1 (newly learned) or week 2 (relearned; level 2 or j), which are nested within children (level 3 or k; Twisk, 2006). Because the response variable was binary (word-pair is remembered correct = 1 or not correct = 0), logistic binomial analyses were applied (logit model: $\log[\pi_{ij}/(1-\pi_{ij})] = \beta_0 + \beta_1 x_{ij} + u_{0j}$). We used the second-order Penalized Quasi-Likelihood (PQL-2) approximation procedure, after initial estimates were obtained by running the model using first-order Marginal Quasi-Likelihood (MQL-1) estimation. Age and sex were included as covariates. Furthermore, for analyses of DR performance, we only included word-pairs that were correctly recognized at IR. All non-binary predictor variables were standardized. To evaluate whether IR and DR performance depended on familiarity versus novelty of word-pairs, dummy coding was applied as follows for each individual word-pair for each individual child. Word-pairs presented during the first week were all coded as relearning = 0. Word-pairs presented during the second week were coded with relearning = 1 if they had been correctly recognized during IR-1, and with relearning = 0 if they had not been learned correctly during week one. The regression coefficients were tested for significance with the Wald test (Rasbash *et al.*, 2012) on a probability level of $\alpha = 0.05$ (two-sided), and transformed to odds ratios (ORs) and their 95% confidence intervals for meaningful reporting. All analyses were performed using MLwiN software version 2.27 (Centre for Multilevel Modelling, University of Bristol, Bristol, UK). A detailed explanation of the models used and results based on mixed ANOVAS conducted with SPSS software version 23 can be found in the Supporting Information.

RESULTS

Table 1 provides an overview of the number of correctly recognized word-pairs for IR and DR for each of the four conditions. Response patterns for IR and DR for negative, neutral and positive word-pairs are shown in Table 2. Table 3 shows the valence and arousal ratings for negative, neutral and positive words averaged over all sessions and subjects. In order to test the convergent construct validity of the children's valence and arousal ratings, we estimated the percentage of variance in their ratings that could be explained by the normative values of word valence (1 very negative/unpleasant to 7 very positive/pleasant) and arousal (1 very passive/calm to 7 very active/aroused) obtained from the database of Moors *et al.* (2013). This database contains the same words as the list of Hermans and De Houwer (1994), but is more recent and reliable. Linear mixed-effect regression analyses showed that the normative values explained only 42% and 3% of the children's valence and arousal ratings, respectively. This suggests a possibly limited validity of children's ratings. Therefore, we decided to exclude the subjective valence and arousal ratings from further analyses, and to use the normative valence and arousal values.

Main effect of sleep versus wake

Mixed-effect logistic regression modelling (see Supporting Information, model a) was used to estimate the effect of sleep versus only wakefulness during the retention interval on DR. DR performance was not affected by the presence versus absence of sleep between learning and DR [OR = 1.06 (0.82–1.38), $P = 0.660$].

Main effects of arousal, valence and novelty

Mixed-effect logistic regression models (see Supporting Information, models b and c) showed that a higher normative arousal value of the cue word significantly increased the probability that a word-pair was correctly recognized at IR [OR = 1.09 (1.05–1.13), $P < 0.001$] and DR [OR = 1.14 (1.08–1.21), $P < 0.001$]. Likewise, a higher normative valence value of the cue word significantly increased the probability that a word-pair was correctly recognized at IR [OR = 1.09 (1.05–1.13), $P < 0.001$] and DR [OR = 1.18 (1.11–1.25), $P < 0.001$]. To obtain a better insight in the relative contribution of negative versus positive valence to these effects, ancillary analyses were performed separately on the subset of negative and neutral words, and on the subset of positive and neutral words. These analyses

Table 1 Number of correctly recognized word-pairs (out of 30) for IR and DR for each of the four conditions

Novel learned *(n = 265)*								***Relearned*** *(n = 149)*							
		*Correct IR-1**		*Correct DR-1†*						*Correct IR-2**		*Correct DR-2†*			
	N	*M*	*SD*		*M*	*SD*	*%‡*		*N*	*M*	*SD*		*M*	*SD*	*%‡*
12h-Wake	62	18.55	4.58	Correct IR-1	12.69	5.32	68%	12h-W	27	21.78	5.14	Correct IR-2	17.67	5.55	81%
				Incorrect IR-1	5.76	2.78						Incorrect IR-2	4.56	3.29	
12h-Sleep	64	19.17	4.68	Correct IR-1	14.42	5.60	75%	12h-S	38	22.50	5.54	Correct IR-2	18.79	6.19	84%
				Incorrect IR-1	5.27	2.41						Incorrect IR-2	4.45	3.36	
24h-Wake 1st	76	18.92	4.57	Correct IR-1	12.86	5.46	68%	24h-W 1st	49	22.67	5.21	Correct IR-2	18.06	6.45	80%
				Incorrect IR-1	5.49	2.83						Incorrect IR-2	4.08	3.12	
24h-Sleep 1st	63	17.78	4.09	Correct IR-1	12.46	5.10	70%	24h-S 1st	35	20.83	5.30	Correct IR-2	15.66	6.36	75%
				Incorrect IR-1	6.27	2.65						Incorrect IR-2	4.63	2.85	
Total	265	18.62§	4.50	Total	18.78§	4.81		Total	149	22.03¶	5.30	Total	22.00¶	4.54	
				Correct IR-1	13.10	5.40						Correct IR-2	17.61	6.25	
				Incorrect IR-1	5.68	2.69						Incorrect IR-2	4.39	3.13	

DR, delayed recognition; IR, immediate recognition; *N*, number of children included; *M*, mean number of recognized word-pairs; *SD*, standard deviation; W, Wake; S, Sleep.
*Time of day at which children learned the words-pairs had no effect on IR performance: IR-1, $F_{3,261} = 1.18$, $P = 0.318$; IR-2, $F_{3,145} = 0.96$, $P = 0.414$.
†Correct word-pairs at DR were presented separately for those that were already correctly recognized at IR and those that were not recognized at IR.
‡Percentage of word-pairs correctly remembered at DR, which were already correctly recognized at IR.
§No significant difference was found between IR-1 and DR-1, $t_{264} = -0.61$, $P = 0.542$.
¶No significant difference was found between IR-2 and DR-2, $t_{148} = 0.10$, $P = 0.923$.

Table 2 Response patterns for IR and DR for negative, neutral and positive word-pairs

	Response pattern		*Novel learned** *(n = 265)*		*Novel learned†* *(n = 149)*		*Relearned†* *(n = 149)*	
Word type	IR	DR	*M*	*SD*	*M*	*SD*	*M*	*SD*
Negative (*n* = 10)	Incorrect	Incorrect	2.0	1.5	2.0	1.5	1.3	1.3
	Incorrect	Correct	1.9	1.3	1.9	1.3	1.5	1.3
	Correct	Incorrect	2.0	1.3	2.0	1.3	1.8	1.4
	Correct	Correct	4.2	2.0	4.1	2.0	5.5	2.3
Neutral (*n* = 10)	Incorrect	Incorrect	1.9	1.4	2.0	1.3	1.1	1.1
	Incorrect	Correct	1.9	1.2	1.9	1.3	1.5	1.3
	Correct	Incorrect	1.8	1.3	1.5	1.2	1.4	1.2
	Correct	Correct	4.5	2.1	4.6	2.1	6.1	2.3
Positive (*n* = 10)	Incorrect	Incorrect	1.9	1.5	1.8	1.4	1.2	1.4
	Incorrect	Correct	1.9	1.3	1.9	1.3	1.4	1.3
	Correct	Incorrect	1.8	1.5	1.8	1.5	1.3	1.2
	Correct	Correct	4.5	2.2	4.5	2.1	6.0	2.3

DR, delayed recognition; IR, immediate recognition; *M*, mean number of words; *SD*, standard deviation, for each combination of correctness on IR and DR for novel learning (*based on 265 participants who completed IR-1 and DR-1), and again for novel learning and relearning (†based on 149 participants who completed all four sessions IR-1, DR-1, IR-2 and DR-2).

showed that the probability a word-pair was correctly remembered at IR and DR was affected across the negative to neutral valence range of cue words (i.e. higher negative valence, worse recognition) [IR: OR = 1.09 (1.01–1.17), $P = 0.027$; DR: OR = 1.18 (1.04–1.33), $P = 0.008$], but not across the neutral to positive valence range of cue words [IR: OR = 0.93 (0.86–1.01), $P = 0.073$; DR: OR = 0.88 (0.77–1.01), $P = 0.073$]. Finally, the probability of correctly recognizing a word-pair was significantly higher for relearned words, both at IR [OR = 2.23 (1.99–2.48), $P < 0.001$] and at DR [OR = 1.90 (1.65–2.19), $P < 0.001$] (see Supporting Information, models d and e).

Table 3 Normative and assessed valence and arousal ratings for cue words with negative, neutral and positive valence

	Cue word valence									
	Negative (n = 20)		*Neutral (n = 20)*		*Positive (n = 20)*					
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*df*	*P*	η_p^2
Normative values*										
Valence	1.94	0.35	4.22	0.35	5.79	0.42	539.38	2,57	<0.001	0.950
Arousal	5.25	0.66	3.73	0.78	4.06	0.98	18.93	2,57	<0.001	0.399
Child ratings†										
Valence	2.85	0.28	4.34	0.34	5.38	0.29	343.34	2,57	<0.001	0.923
Arousal	3.81	0.30	3.02	0.16	3.47	0.22	57.07	2,57	<0.001	0.667

M, mean number; *SD*, standard deviation.
*Valence (1 very negative/unpleasant to 7 very positive/pleasant) and arousal (1 very passive/calm to 7 very active/aroused) for each of the cue word types according to the normative database of Moors *et al.* (2013). Bonferroni *post hoc* analyses showed that all three word types significantly differed in valence. Regarding arousal: negative words were significantly more arousing than neutral and positive words. No significant difference in arousal was found between neutral and positive words.
†Valence ratings (1 very unpleasant to 7 very pleasant) and arousal (1 least arousing to 7 very arousing) as indicated on the Self-Assessment Manikin (SAM) averaged over negative, neutral and positive words across participants and sessions. Bonferroni *post hoc* analyses showed that valence and arousal ratings differed between all three cue word types.

Interactions of sleep with arousal, valence and novelty

Interaction terms were added to the mixed-effect logistic regression models in order to investigate whether a period of sleep versus wake moderated effects of arousal, valence and novelty on the probability a word-pair was remembered correctly at DR (see Supporting Information, models f and g). Results showed no significant interaction of sleep with arousal [OR = 1.01 (0.88–1.17), $P = 0.882$], with valence [OR = 1.10 (0.95–1.27), $P = 0.196$] or with relearning [OR = 0.85 (0.59–1.23), $P = 0.384$].

DISCUSSION

The current study investigated how the memory consolidating effect of sleep compares with, and might interact with, other factors known to affect consolidation, notably the arousal, valence and novelty properties of the acquired information. Children aged 9–11 years participated in the assessment of IR and DR of learned word-pairs, across intervals of 12 or 24 h that included periods of wakefulness, sleep or both. Cue words covered a range of arousal and valence values according to a normative database, and a distinction was made between previously learned and novel word-pairs.

Corresponding to previous studies (Atienza and Cantero, 2008; Kensinger and Corkin, 2003), we found that more arousing cue words resulted in better IR and DR. However, in contrast to the consolidation advantage of emotional versus neutral word-pairs previously reported for adults (Kensinger, 2004; Kensinger and Corkin, 2003; Walker and Van Der Helm, 2009), our main finding was that the more negatively valenced the cue word was, the less likely children recognized its associated target. Because the effect was consistent across IR and DR, the findings seem highly relevant to follow up on. Future studies may address whether the disadvantage of coupling target words to negatively toned cue words occurs during encoding, or during subsequent consolidation. Interestingly, a previous study on memory for neutral and negatively toned pictures in children reported *better* recognition of negative pictures (Prehn-Kristensen *et al.*, 2009). This previous study concerned mere recognition of visual information that had been seen before, but not the coupling of cues and targets. Accordingly, a negative emotional tone might enhance storage of the cue, but simultaneously impede processing of information associated with it. Alternatively, the striking differences in findings may be related to differential engagement of attention-related brain regions in processing pictures compared with words (Flaisch *et al.*, 2015).

Possibly related to the aforementioned differences in cognitive and brain functional processes, the present study did not find the previously suggested beneficial effect of a period including sleep (Prehn-Kristensen *et al.*, 2009, 2013). In fact, the present study did not show any influence of post-learning sleep on the DR of the right neutral target that was previously presented in pair with a cue word. In addition to the lack of a main effect of sleep, there were no interaction effects either, indicating that sleep does not seem to play a major role irrespective of the valence, arousal or novelty of the cue word. This lack of a beneficial effect of sleep contrasts with previous studies in adults (Diekelmann *et al.*, 2009; Gais *et al.*, 2006; Marshall and Born, 2007; Payne *et al.*, 2012; Rasch and Born, 2013) and children or adolescents (Backhaus *et al.*, 2008; Potkin and Bunney, 2012; Prehn-Kristensen *et al.*, 2009; Wilhelm *et al.*, 2008). Interestingly, performance did not deteriorate over a period of 12–

24 h between IR and DR, suggestive of an immediate maximally effective storage that could only be improved further by a second learning opportunity. Stabile memory performance across a period of wake was also present in a previous study in children (Ashworth *et al*., 2015; Backhaus *et al*., 2008), in contrast to the deterioration in declarative memory across wakefulness in adults (Gais *et al*., 2006; Payne *et al*., 2012).

Of note, we used normative valence and arousal values rather than the children's ratings because these seemed less reliable. Children found it difficult to indicate their feelings for the cue words, especially for arousal, because pictograms around neutral were frequently selected for all words, which resulted in little variance. Analyses using the children's ratings therefore lost significance for arousal and resulted in smaller effects of similar direction for valence [IR: OR = 1.02 (1.00–1.04), $P = 0.061$; DR: OR = 1.05 (1.02–1.08), $P = 0.002$], the more negatively the cue word was rated, the less likely children recognized its associated target.

A possible limitation of our study is that sleep-dependent memory consolidation is more consistently found with recall paradigms than with recognition paradigms. It has been suggested that the small effects or lack of effects of sleep in recognition paradigms could be due to the fact that the hippocampus is less involved in recognition than in recall (Diekelmann *et al*., 2009). In this view, the hippocampus is crucially involved in sleep-dependent memory consolidation processes, through specific coordinated neurophysiological events (slow waves, spindles, ripples) that facilitate the integration of new information into pre-existing cortical networks.

Another limitation of the present study is that only 149 of the 386 children (39%) completed the whole procedure of four sessions (IR-1, DR-1, IR-2 and DR-2). The protocol of the first data wave consisted of only two sessions with 82 children at IR-1 of whom 54 (66%) completed DR-1. The second data wave consisted of four sessions and included 304 children at IR-1, $n = 211$ at DR-1 (69%), $n = 197$ at IR-2 (65%) and $n = 155$ at DR-2 (51%). The missing data are mainly due to non-adherence because of the timing of the protocol and not to demanding characteristics of the memory task itself. The morning session between 07:00 and 08:00 hours was reported to be inconvenient because participants were in a rush to get ready for school. The evening session between 19:00 and 20:00 hours was reported to interfere with scheduled activities, such as sports, hobbies and clubs. Of note, no difference in IR performance was found between the children who completed the whole procedure and the ones that dropped out ($t_{384} = -0.04$, $P = 0.972$). Moreover, the number of completers is still substantially larger than in previous studies in children (Backhaus *et al*., 2008; Potkin and Bunney, 2012; Prehn-Kristensen *et al*., 2009; Wilhelm *et al*., 2008). The relatively large sample was possible because we used online assessment at home, even though this has the disadvantage that it is less controlled compared with laboratory assessments.

The importance of sleep and valence on memory consolidation is well studied, and often confirmed in adults and adolescents. Results for children, however, are more equivocal. We therefore performed a comprehensive study that allowed us to compare the robustness of the effects of sleep relative to the effects of stimulus characteristics valence, arousal and novelty. The key finding suggested by the current study is that the negative valence of a cue word strongly determines how well a neutral word paired to it can be recognized, both immediately after learning and in addition after an interval of 12–24 h. No main or interaction effects of sleep could be found for the present declarative memory paradigm. The lack of effect of sleep on memory performance is in line with a meta-analysis in children aged 5–12 years (Astill *et al*., 2012) and studies in adolescents aged 14–16 years (Kopasz *et al*., 2010; Voderholzer *et al*., 2011). One possible interpretation of the lack of sleep effects is that consolidation in children might occur more independently of sleep than the sleep-dependency suggested by several studies in adults. The same interpretation was made by Astill *et al.* (2012), based on a meta-analysis that included, among other studies, six experimental studies with designs that allowed for direct comparison between wake and sleep intervals. It may be as well that specific consolidation advantages of intervals with sleep can surface only under more controlled or specific conditions in children and are more easily detectable in adults.

ACKNOWLEDGEMENTS

The authors want to thank all schools and families for participation, and all students who helped collecting the data. This work was supported by the Netherlands Organization of Scientific Research (NWO), The Hague: National Initiative Brain & Cognition Research Program 'Innovative Learning Materials and Methods' under Grant 056-32-013; and VICI Innovation Grant 453-07-001.

AUTHOR CONTRIBUTIONS

All authors have contributed to this manuscript by one or more of the following activities: study design, data collection, data analysis, interpretation of results, preparation of the manuscript.

CONFLICT OF INTEREST

All authors declare no conflicts of interest.

REFERENCES

Ackermann, S., Hartmann, F., Papassotiropoulos, A., De Quervain, D. J. and Rasch, B. No associations between interindividual

differences in sleep parameters and episodic memory consolidation. *Sleep*, 2015, 38: 951–959.
Almarode, J. and Miller, A. M. *Captivate, Activate, and Invigorate the Student Brain in Science and Math, Grades 6–12*. Corwin Press, California, US, 2013.
Ashworth, A., Hill, C. M., Karmiloff-Smith, A. and Dimitriou, D. Sleep enhances memory consolidation in children. *J. Sleep Res.*, 2014, 23: 302–308.
Ashworth, A., Hill, C. M., Karmiloff-Smith, A. and Dimitriou, D. The importance of sleep: attentional problems in school-aged children with Down syndrome and Williams syndrome. *Behav. Sleep Med.*, 2015, 13: 455–471.
Astill, R. G., Van Der Heijden, K. B., Van Ijzendoorn, M. H. and Van Someren, E. J. W. Sleep, cognition, and behavioral problems in school-age children: a century of research meta-analyzed. *Psychol. Bull.*, 2012, 138: 1109–1138.
Atienza, M. and Cantero, J. L. Modulatory effects of emotion and sleep on recollection and familiarity. *J. Sleep Res.*, 2008, 17: 285–294.
Backhaus, J., Hoeckesfeld, R., Born, J., Hohagen, F. and Junghanns, K. Immediate as well as delayed post learning sleep but not wakefulness enhances declarative memory consolidation in children. *Neurobiol. Learn. Mem.*, 2008, 89: 76–80.
Balderston, N. L., Schultz, D. H. and Helmstetter, F. J. The human amygdala plays a stimulus specific role in the detection of novelty. *Neuroimage*, 2011, 55: 1889–1898.
Basso, G., Magon, S., Reggiani, F. *et al.* Distinguishable neurofunctional effects of task practice and item practice in picture naming: a BOLD fMRI study in healthy subjects. *Brain Lang.*, 2013, 126: 302–313.
Blackford, J. U., Buckholtz, J. W., Avery, S. N. and Zald, D. H. A unique role for the human amygdala in novelty detection. *Neuroimage*, 2010, 50: 1188–1193.
Born, J., Rasch, B. and Gais, S. Sleep to remember. *Neuroscientist*, 2006, 12: 410–424.
Bradley, M. M. and Lang, P. J. Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *J. Behav. Ther. Exp. Psychiatry*, 1994, 25: 49–59.
De Deyne, S. and Storms, G. Word associations: norms for 1,424 Dutch words in a continuous task. *Behav. Res. Methods*, 2008, 40: 198–205.
De Jonge, M. and Tabbers, H. K. Repeated testing, item selection, and relearning: the benefits of testing outweigh the costs. *Exp. Psychol.*, 2013, 60: 206–212.
Diekelmann, S. and Born, J. The memory function of sleep. *Nat. Rev. Neurosci.*, 2010, 11: 114–126.
Diekelmann, S., Wilhelm, I. and Born, J. The whats and whens of sleep-dependent memory consolidation. *Sleep Med. Rev.*, 2009, 13: 309–321.
Ellenbogen, J. M., Payne, J. D. and Stickgold, R. The role of sleep in declarative memory consolidation: passive, permissive, active or none? *Curr. Opin. Neurobiol.*, 2006, 16: 716–722.
Flaisch, T., Imhof, M., Schmalzle, R., Wentz, K. U., Ibach, B. and Schupp, H. T. Implicit and explicit attention to pictures and words: an fMRI-study of concurrent emotional stimulus processing. *Front. Psychol.*, 2015, 6: 1861.
Gais, S., Lucas, B. and Born, J. Sleep after learning aids memory recall. *Learn. Mem.*, 2006, 13: 259–262.
Henderson, L. M., Weighall, A. R., Brown, H. and Gareth Gaskell, M. Consolidation of vocabulary is associated with sleep in children. *Dev. Sci.*, 2012, 15: 674–687.
Hermans, D. and De Houwer, J. Affective and subjective familiarity ratings of 740 Dutch words. *Psychol. Belgica*, 1994, 34: 115–139.
Hu, P., Stylos-Allan, M. and Walker, M. P. Sleep facilitates consolidation of emotional declarative memory. *Psychol. Sci.*, 2006, 17: 891–898.
Jager, G., Block, R. I., Luijten, M. and Ramsey, N. F. Cannabis use and memory brain function in adolescent boys: a cross-sectional multicenter functional magnetic resonance imaging study. *J. Am. Acad. Child Adolesc. Psychiatry*, 2010, 49: 561–572, 572.e1–3.
Kensinger, E. A. Remembering emotional experiences: the contribution of valence and arousal. *Rev. Neurosci.*, 2004, 15: 241–251.
Kensinger, E. A. and Corkin, S. Memory enhancement for emotional words: are emotional words more vividly remembered than neutral words? *Mem. Cognit.*, 2003, 31: 1169–1180.
Kopasz, M., Loessl, B., Valerius, G. *et al.* No persisting effect of partial sleep curtailment on cognitive performance and declarative memory recall in adolescents. *J. Sleep Res.*, 2010, 19: 71–79.
LaBar, K. S. and Cabeza, R. Cognitive neuroscience of emotional memory. *Nat. Rev. Neurosci.*, 2006, 7: 54–64.
Marshall, L. and Born, J. The contribution of sleep to hippocampus-dependent memory consolidation. *Trends Cogn. Sci.*, 2007, 11: 442–450.
Moors, A., De Houwer, J., Hermans, D. *et al.* Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behav. Res. Meth.*, 2013, 45: 169–177.
Murray, R. J., Brosch, T. and Sander, D. The functional profile of the human amygdala in affective processing: insights from intracranial recordings. *Cortex*, 2014, 60: 10–33.
Payne, J. D., Stickgold, R., Swanberg, K. and Kensinger, E. A. Sleep preferentially enhances memory for emotional components of scenes. *Psychol. Sci.*, 2008, 19: 781–788.
Payne, J. D., Tucker, M. A., Ellenbogen, J. M. *et al.* Memory for semantically related and unrelated declarative information: the benefit of sleep, the cost of wake. *PLoS One*, 2012, 7: e33079.
Potkin, K. T. and Bunney, W. E. Jr. Sleep improves memory: the effect of sleep on long term memory in early adolescence. *PLoS One*, 2012, 7: e42191.
Prehn-Kristensen, A., Goder, R., Chirobeja, S., Bressmann, I., Ferstl, R. and Baving, L. Sleep in children enhances preferentially emotional declarative but not procedural memories. *J. Exp. Child Psychol.*, 2009, 104: 132–139.
Prehn-Kristensen, A., Munz, M., Molzow, I., Wilhelm, I., Wiesner, C. D. and Baving, L. Sleep promotes consolidation of emotional memory in healthy children but not in children with attention-deficit hyperactivity disorder. *PLoS One*, 2013, 8: e65098.
Rasbash, J., Steele, F., Browne, W. J. and Goldstein, H. *A User's Guide to MLwiN: Version 2.26*. Centre for Multilevel Modelling, University of Bristol, Bristol, UK, 2012.
Rasch, B. and Born, J. About sleep's role in memory. *Physiol. Rev.*, 2013, 93: 681–766.
Roediger, H. L. and Karpicke, J. D. Test-enhanced learning: taking memory tests improves long-term retention. *Psychol. Sci.*, 2006, 17: 249–255.
Stickgold, R. and Walker, M. P. Sleep-dependent memory consolidation and reconsolidation. *Sleep Med.*, 2007, 8: 331–343.
Tomasi, D., Ernst, T., Caparelli, E. C. and Chang, L. Practice-induced changes of brain function during visual attention: a parametric fMRI study at 4 Tesla. *Neuroimage*, 2004, 23: 1414–1421.
Twisk, J. W. R. *Applied Multilevel Analysis: A Practical Guide*. Cambrigde University Press, Cambridge, UK, 2006.
Voderholzer, U., Piosczyk, H., Holz, J. *et al.* Sleep restriction over several days does not affect long-term recall of declarative and procedural memories in adolescents. *Sleep Med.*, 2011, 12: 170–178.
Wagner, U., Gais, S. and Born, J. Emotional memory formation is enhanced across sleep intervals with high amounts of rapid eye movement sleep. *Learn. Mem.*, 2001, 8: 112–119.

Walker, M. P. and Van Der Helm, E. Overnight therapy? The role of sleep in emotional brain processing. *Psychol. Bull.*, 2009, 135: 731–748.
Weierich, M. R., Wright, C. I., Negreira, A., Dickerson, B. C. and Barrett, L. F. Novelty as a dimension in the affective brain. *Neuroimage*, 2010, 49: 2871–2878.
Wilhelm, I., Diekelmann, S. and Born, J. Sleep in children improves memory performance on declarative but not procedural tasks. *Learn. Mem.*, 2008, 15: 373–377.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. Detailed explanation of the three-level mixed-effect logistic regression models and results using mixed ANOVA's.