Structure and limits of
unconscious episodic memory

Inauguraldissertation der Philosophisch-humanwissenschaftlichen Fakultät der Universität Bern zur Erlangung der Doktorwürde vorgelegt von

Marc Alain Züst

Diepoldsau, SG

Public Library of Science, San Francisco CA, 2015

Oxford University Press, Oxford, 2016

Nature Publishing Group, London, 2016

Von der Philosophisch-humanwissenschaftlichen Fakultät der Universität Bern auf Antrag von Prof. Dr. Katharina Henke (Hauptgutachterin) und Prof. Dr. Björn Rasch (Zweitgutachter) angenommen.

Bern, den 07. März 2017

Der Dekan: Prof. Dr. Fred Mast

Preface

When committing a new entry to episodic memory – the memory for unique events, a fast association of stimuli is necessary. Our research shows that episodic memory formation and retrieval is possible outside of consciousness, and that unconscious episodic memory functions rely on the same brain structures.

The goal of this dissertation is the scientific investigation of structure and limits of unconscious episodic memory, which was followed in three projects. In the first project, we utilized functional magnetic resonance imaging (fMRI) to show that unconscious and conscious episodic memories are processed in the same brain structure, the hippocampus, where conscious and unconscious memories interacted. A conscious retrieval of a memory was more expedient if it followed an unconscious retrieval of a similar memory. In the second project, we found that unconsciously encoded memories endured over at least 25 minutes of resting. Unconscious memories could still be retrieved and influence behavior after this extended retrieval interval. In the third, most complex and innovative of the three projects we found that novel vocabulary can be encoded and stored during sleep. This sleep-learned vocabulary can be unconsciously retrieved after waking. Brain activity during retrieval as recorded with fMRI showed that hippocampal activity accompanied unconscious retrieval of sleep-learned vocabulary. This is evidence that vocabulary acquisition during sleep is putatively mediated by the same brain areas that mediate conscious vocabulary acquisition while awake, most prominently hippocampus.

Learning is a laborious process. The prospect of doing so unconsciously, i.e. without effort, could lead to applications in educational and clinical settings. With regards to children with learning difficulties (e.g. ADHD) being impaired in keeping conscious focus on learning material, an unconscious approach could be especially helpful.

Publications

1. **Züst, M. A.**, Colella, P., Reber, T. P., Vuilleumier, P., Hauf, M., Ruch, S., & Henke, K. (2015). Hippocampus Is Place of Interaction between Unconscious and Conscious Memories. *PLoS ONE*, *10*(3), e0122459. https://doi.org/10.1371/journal.pone.0122459   p. 1

2. Ruch, S., **Züst, M. A.**, & Henke, K. (2016). Subliminal messages exert long-term effects on decision-making. *Neuroscience of Consciousness*, *2016*(1), niw013. https://doi.org/10.1093/nc/niw013   p. 20

3. **Züst, M. A.**, Ruch, S., Wiest, R., & Henke, K. (submitted). Vocabulary acquisition during sleep. *Nature Neuroscience*.   p. 30

PLOS | ONE

Hippocampus Is Place of Interaction between Unconscious and Conscious Memories

Marc Alain Züst[1,2]*, Patrizio Colella[1,2], Thomas Peter Reber[1,2], Patrik Vuilleumier[3], Martinus Hauf[4], Simon Ruch[1,2], Katharina Henke[1,2]

1 Department of Psychology, Division of Experimental Psychology and Neuropsychology, University of Bern, Bern, Switzerland, **2** Center for Cognition, Learning and Memory, University of Bern, Bern, Switzerland, **3** Department of Neurosciences and Clinical Neurology, University of Geneva, Geneva, Switzerland, **4** Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Bern, Switzerland

* marc.zuest@psy.unibe.ch

Abstract

Recent evidence suggests that humans can form and later retrieve new semantic relations unconsciously by way of hippocampus—the key structure also recruited for conscious relational (episodic) memory. If the hippocampus subserves both conscious and unconscious relational encoding/retrieval, one would expect the hippocampus to be place of unconscious-conscious interactions during memory retrieval. We tested this hypothesis in an fMRI experiment probing the interaction between the unconscious and conscious retrieval of face-associated information. For the establishment of unconscious relational memories, we presented subliminal (masked) combinations of unfamiliar faces and written occupations ("actor" or "politician"). At test, we presented the former subliminal faces, but now supraliminally, as cues for the reactivation of the unconsciously associated occupations. We hypothesized that unconscious reactivation of the associated occupation—actor or politician—would facilitate or inhibit the subsequent conscious retrieval of a celebrity's occupation, which was also actor or politician. Depending on whether the reactivated unconscious occupation was congruent or incongruent to the celebrity's occupation, we expected either quicker or delayed conscious retrieval process. Conscious retrieval was quicker in the congruent relative to a neutral baseline condition but not delayed in the incongruent condition. fMRI data collected during subliminal face-occupation encoding confirmed previous evidence that the hippocampus was interacting with neocortical storage sites of semantic knowledge to support relational encoding. fMRI data collected at test revealed that the facilitated conscious retrieval was paralleled by deactivations in the hippocampus and neocortical storage sites of semantic knowledge. We assume that the unconscious reactivation has pre-activated overlapping relational representations in the hippocampus reducing the neural effort for conscious retrieval. This finding supports the notion of synergistic interactions between conscious and unconscious relational memories in a common, cohesive hippocampal-neocortical memory space.

PLOS | ONE

Introduction

Episodic memory is a class of declarative memory thought to depend on consciousness of encoding and retrieval [1–3]. The hippocampus is the neuroanatomical hub governing the encoding and retrieval of episodic memories. Damage to the hippocampal-anterior thalamic axis produces severe impairments of episodic memory, but leaves unconscious forms of memory such as skill-learning or priming intact because these forms of memory depend on extrahippocampal structures [1–3].

Recent evidence suggests, however, that episodic memory formation and retrieval is possible even without conscious awareness of encoding and retrieval, and that both encoding and retrieval depend on the hippocampal-anterior thalamic axis [4]. These findings question classic notions of separate memory systems [1–3] and support the processing-based memory model [5] that distinguishes memory systems based on processing modes rather than consciousness. The processing-based memory model distinguishes between memory systems with respect to 3 variables: speed of encoding (rapid versus slow), nature of representation (flexible versus rigid), and memory content (single items versus associations). This model hypothesizes the existence of both a conscious and unconscious form of episodic memory with both forms depending on the hippocampal anterior-thalamic axis. Consciousness, therefore, is not prerequisite for relational encoding and retrieval but rather an independent factor that serves the strengthening of hippocampal memory representations [4].

If episodes can be encoded with and without consciousness by way of the hippocampal anterior-thalamic axis and related cortices [6–8], the organization of consciously and unconsciously acquired information in a single, cohesive hippocampal memory space is economically and evolutionarily sensible. Linked episodic knowledge—conscious and unconscious—informs and guides us better through life than episodic knowledge that is stored separated according to levels of representation from conscious to unconscious. Episodic memories are dynamic and subject to transformation from conscious to unconscious and vice versa. Consider an unconscious memory trace that suddenly "pops" into consciousness, or implicit knowledge of a hidden sequence in a serial reaction time task [9], or a rule in the number reduction task [10,11] that become consciously accessible following sleep. Conversely, memory traces can also get purged from conscious access dropping to a pre-conscious representation [12]. Consciously encoded memories can also become inaccessible when one is instructed to forget them [13]. In all of these cases, a cohesive memory space provides for a stable organizational structure of memory that allows for shifts in the level of representation from unconscious to conscious and vice versa. Such representational shifts appear more difficult if one assumes a strict division between memory systems based on conscious access.

If conscious and unconscious episodic memories are both accommodated by the hippocampal memory system, they can be expected to interact both synergistically and competitively. For example, the activation of *unconscious* memories may facilitate the subsequent formation and retrieval of content-congruent *conscious* memories through activation of nearby or overlapping neural assemblies. We tested this hypothesis using functional magnetic resonance imaging (fMRI). In particular, the hippocampus was hypothesized to be place of interactions between unconscious and conscious retrieval processes.

Participants were first presented with subliminal combinations of unfamiliar faces and occupations (face plus the label "actor" or face plus the label "politician") for unconscious relational encoding. Due to the relational nature of unconscious memory formation, we expected the hippocampus to be activated during unconscious encoding. Following the subliminal presentation of face-occupation combinations, an unconscious-conscious retrieval interaction test was given. We studied whether the unconscious reactivation of the earlier formed face-

occupation association would facilitate or inhibit the conscious retrieval of a stored association between a celebrity's face and his occupation, namely actor or politician. We used portraits of famous actors and politicians as cues for the conscious retrieval of occupations. This relational retrieval draws on both the episodic and the semantic (facts) memory system depending on the experience that the young participants in our study had with movies and political shows/news [14]. Each test trial included the brief but visible presentation of a former subliminal face, stripped off its occupation label, followed by the presentation of the portrait of a celebrity. Participants were instructed to react to the famous face by deciding whether the depicted person was an actor or a politician (Fig 1).

We expected that the former subliminal face's occupation (actor or politician) would be re-activated unconsciously and would facilitate or inhibit the conscious retrieval of the celebrity's occupation (actor or politician) depending on whether the two faces share occupations or not. A facilitating interaction may render the conscious retrieval more efficient reducing neural activation and reaction times. Conversely, an inhibitory interaction may increase neural activation and reaction times. These two conditions were contrasted to a baseline condition that provided for unconscious face encoding and retrieval without relational demands and hence was neutral regarding occupational categorization.

The hypothesized neural basis of unconscious-conscious interactions comprises the neocortical storage sites of occupations, namely the lateral and polar temporal cortex [16,17], as well as the hippocampal-anterior thalamic axis. Evidence in favour of a common memory space for both unconscious and conscious relational memories would speak to a common memory system for conscious and unconscious relational (i.e., episodic) memories. Such evidence would challenge the traditional segmentation of memory systems according to consciousness [1–3] and would support the processing-based memory model [5].

Methods

Participants

Forty-two healthy male volunteers (age 19–32 years; $M \pm SD = 23.86 \pm 3.02$) participated in the experiment. They denied previous or current neurological or psychiatric disorders and drug abuse. Each participant fulfilled inclusion criteria and no exclusion criteria for MRI. All participants were right handed [18] and had normal or corrected-to-normal vision. Participants gave semi-informed consent. They were not informed of subliminal presentations until debriefing following the fMRI experiment. This study was approved by the local ethics committee for human studies (Kantonale Ethikkommission Bern).

Two participants were excluded from data analysis due to their insufficient acquaintance with the famous faces used in the experiment. Seven further participants were excluded because of their above-chance performance on the awareness tests (cf. section 0). Accordingly, thirty-three participants entered data analyses.

Material

Stimuli.   For subliminal encoding, 216 frontal portraits of unknown male faces were retrieved from the internet using Google (http://images.google.com) and the FERET database [15]. The portraits were converted to grayscale, realigned and contrast-reduced. They were then validated concerning their prototypicality for the occupations "actor" and "politician" by 32 (separate) students, who agreed to evaluate the faces in an online experiment using a forced-choice task. The 148 least prototypical portraits were used as a stimulus pool for the experiment. Forty-eight portraits were used in the main experiment and the remaining 100 portraits were used in the awareness tests. The assignment of faces to these two lists was randomized for
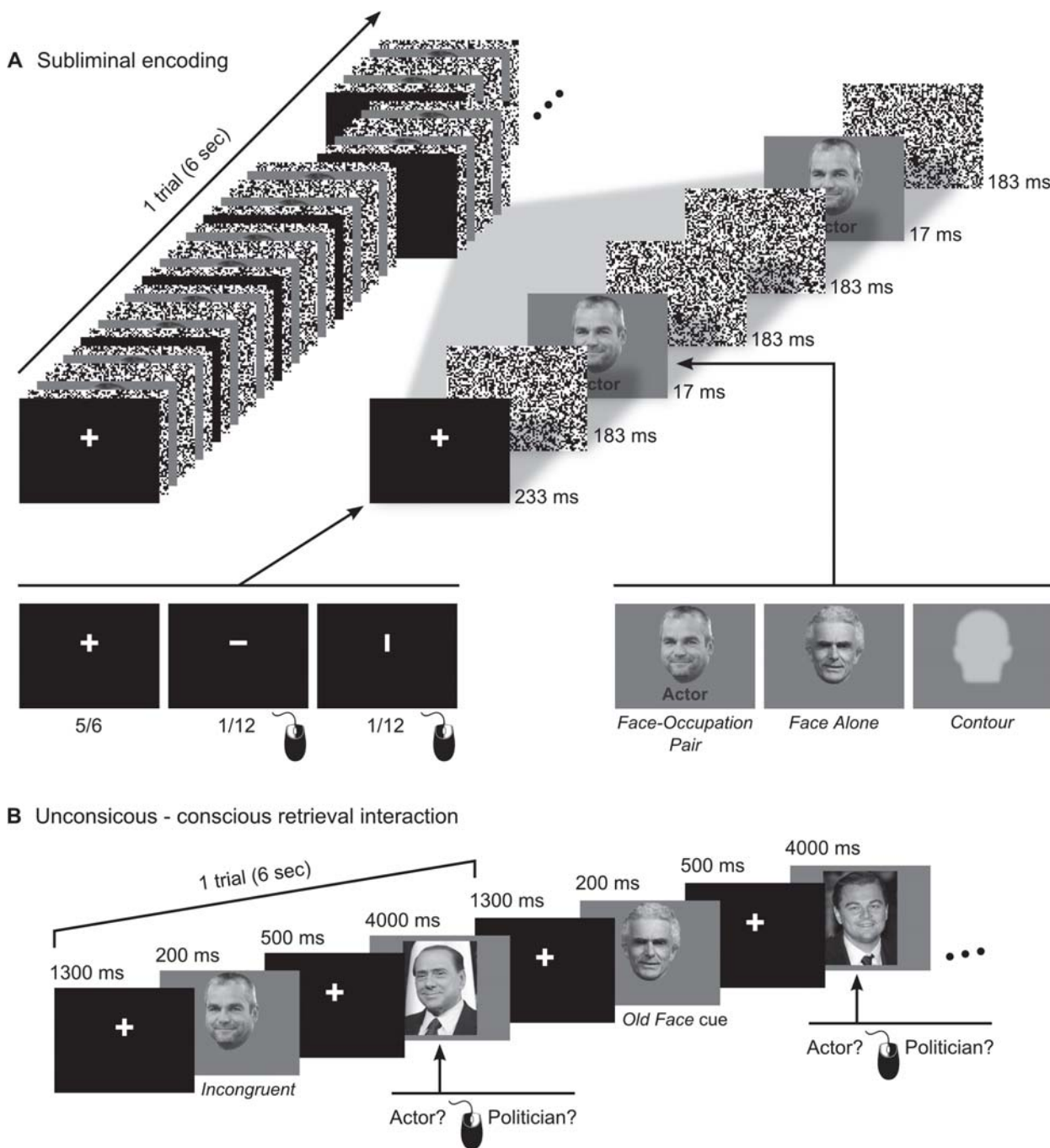
Fig 1. Experimental design. A: Attention task during subliminal encoding. Participants saw a flickering stream of black-and-white pixel masks. Subliminal stimuli were presented between masks. The top left depicts one encoding trial containing twelve repetitions of one subliminal stimulus. Four encoding trials constitute a condition block in this fMRI design. On the top right, a section of an encoding trial is highlighted with indicated presentation durations. To the lower left, the used fixation screens are displayed with their respective frequencies of appearance. Each encoding trial contained one response slide (either a vertical or horizontal line segment). To the lower right, we display the three stimulus categories that belong to the three experimental encoding conditions (from left to right): *Face-Occupation Pairs* for associative encoding, *Faces Alone* for single item encoding (non-associative baseline) and *Contour* for a non-encoding baseline (not discussed in this paper). Portraits belong to the FERET database *[15]*. **B:** Unconscious-conscious retrieval interaction with indicated presentation durations. A former subliminal face is briefly presented to cue the unconscious reactivation of previously formed face-occupation association. Next, a portrait of a celebrity comes up for the conscious retrieval of the celebrity's occupation (actor or politician) Participants were required to recognise the famous person and to indicate his occupation by button press. Each condition block contained four trials. Fig 1B illustrates a trial of the associative retrieval condition *Incongruent* and a trial of the *Old Faces* baseline condition, where no unconscious-conscious interaction was possible. Celebrities' portraits were taken from Wikimedia Commons (http://commons.wikimedia.org/wiki/Main_Page). Berlusconi: public domain; DiCaprio: Siebbi (http://www.ipernity.com/home/siebbi).

each participant. A contour of a human head was reduced in contrast and blurred to be used in a baseline condition in the subliminal encoding part of the fMRI experiment

For the test of an unconscious-conscious interaction during retrieval we collected frontal portraits of 32 famous male actors and 32 famous male politicians from the internet. These portraits were also grayscaled and realigned but not contrast-reduced because they were not used for subliminal presentation.

Setup. Stimuli were presented with a Benq© WXGA SP830 DLP video projector using a resolution of 1024 × 768 pixels and a screen refresh rate of 60 Hz. Stimuli were projected onto a backlit screen with a viewing angle of 16° width and 9° height. Stimulus presentation was programmed with the software Presentation Version 11.3 (Neurobehavioral Systems, http://www. neurobs.com). Participants responded by key press on a Lumina Response Pad LU400-Pair by Cedrus (www.cedrus.com/lumina) while lying in the MR Scanner.

Experimental procedure

The experiment was carried out in a dimmed MRI chamber. The study encompassed the following phases in this order: 1) a conscious memory task was given to establish a task-set that prepares participants for unconscious associative encoding, 2) the fMRI experiment encompassing subliminal encoding and a test of unconscious-conscious interaction during retrieval, 3) a test of stimulus awareness, and 4) the explicit identification of famous faces. Phases 1 through 3 were carried out while the participants where situated inside the MR scanner.

The experiment was designed to suit an fMRI block design with alternating condition blocks. There were two fMRI time-series, one for subliminal encoding and the other for the interaction test. All condition blocks took 24 seconds and contained four trials spanning 6 seconds each. The assignment of stimuli to conditions and of occupations to faces was pseudo-randomized. Condition blocks alternated regularly in a fixed order. The starter block varied between participants to distribute over experimental conditions certain psychological dispositions such as stress or fatigue and the pervasive scanner drift.

Subliminal encoding. We used our established presentation protocol with subliminal stimuli embedded in an attention task [7] (Fig 1A). Initially, a fixation cross (F) was presented for 233 ms. Four noise masks (M) were then presented for 183 ms each. Between the noise masks, stimuli (S) were presented subliminally for 17 ms. Stimuli were either *Face-Occupation Pairs*, *Faces Alone* (= non-associative baseline) or *Contour* (not discussed in this study). The noise masks served as forward- and backward masks [7,19]. One trial took six seconds, consisted of 6 sub-trials and ran down in the following order: 6 × (F-M-S-M-M-S-M). This resulted in 12 consecutive subliminal presentations of a stimulus. In each trial, one of the six fixation crosses was replaced by either a horizontal or a vertical line segment. These replacements had to be acknowledged by participants with key press responses. This attention task ensured that participants' attention remained focused on the centre of the screen throughout the task.

Subliminal encoding was implemented as a block design with three alternating conditions, namely *Face-Occupation Pairs*, *Faces Alone* baseline and *Contour*. Each condition embraced four blocks with four trials each. Hence, we presented 16 *Face-Occupation Pairs*, 16 *Faces Alone* and 16 times the *Contour*. According to this scheme, 32 of the 48 experimental unfamiliar portraits were presented during the encoding task. The remaining 16 portraits were later used for the unconscious-conscious retrieval interaction test in the *Novel Faces* condition (see next section).

Test of an unconscious-conscious interaction during retrieval. In the test of interaction during retrieval, participants had to categorize celebrities with respect to their occupation—actor or politician (Fig 1B). This test encompassed four conditions: *Congruent*, *Incongruent*,

Old Face and *Novel Face*. Each condition embraced four blocks of four trials. Sixty-four portraits of famous actors and politicians were presented as targets. The presentation of a portrait was preceded by the brief but clearly visible presentation of one of 48 non-famous faces. Of the 48 non-famous faces, 32 had previously been shown in the subliminal encoding task. The previously presented 16 *Face-Occupation Pairs* were assigned to both the *Congruent* and the *Incongruent* condition. The previously presented *Faces Alone* were assigned to the no-interaction baseline condition of *Old Faces*. The remaining 16 faces had not been presented for encoding; they were presented in the condition of *Novel Faces* (not discussed in this paper). The apparent discrepancy between the number of non-famous faces (48: 16 associative old, 16 single old, 16 not presented for encoding) and the number of famous faces (64) is explained by the fact that each of the 16 former subliminal *Face-Occupation Pairs* was used twice, namely once in the congruent condition and once in the incongruent condition. Accordingly, 16 famous faces were preceded by a congruent associative old face, 16 by an incongruent associative old face, 16 by a non-associative old face, and 16 by a new (not previously presented) face. A trial ([Fig 1B](#)) started with the presentation of a fixation cross for 1300 ms. This was followed by a 200 ms presentation of a non-famous face. Next, a fixation cross appeared again for 500 ms (= cue-target interval). Finally, a famous face was presented for 4000 ms. Participants were asked to indicate as quickly as possible whether the famous face was an actor or a politician.

Test of stimulus awareness. Following the fMRI experiment, participants were asked whether they had noticed something during the attention task that they performed in the first part of the fMRI experiment. When they denied, they were further asked whether they might have perceived faces or words between or within the noise masks. A yes answer led to the exclusion of this participant's data set. Following this inquiry, all participants were informed of the subliminal presentation paradigm. Next, we administered two objective tests of stimulus awareness. In these tests, participants' potential awareness of subliminal stimuli was assessed based on their choice behaviour. We first applied a test of face awareness that tested for the awareness of subliminally presented individual faces. Next, we applied a test of occupation awareness that tested for the awareness of subliminally presented faces plus written occupations. Each awareness tests comprised 50 trials. A trial consisted of the 12-fold subliminal presentation of a stimulus (procedure adopted from subliminal encoding in the main experiment) followed by the forced-choice test concerning this stimulus. Hence, unlike the experiment, there was no encoding-test interval. The immediate succession of a subliminal stimulus and its test facilitates the behavioural expression of stimulus awareness. In the test of face awareness, we presented 50 subliminal unfamiliar faces, each followed by the supraliminal side-by-side presentation of the target face plus a distractor face (presentation duration: 5 s). Subjects were asked to indicate which of the two faces had just been presented subliminally. In the test of occupation awareness, half of 50 faces were presented subliminally with the written occupation "actor" and the other half with "politician". Each subliminal face-occupation pair was followed by a forced-choice test that required participants to choose between the two occupations. Participants were given 5 s to indicate which of the two occupations was just presented subliminally. In both awareness tests, participants received direct test instructions: they were instructed to base their decisions between faces or occupations on their previous conscious perception of shapes or fragments of subliminal stimuli. Direct test instructions such as this are known to be more sensitive to conscious than unconscious perception and memory [20,21], which allows measuring stimulus awareness. On the other hand, indirect retrieval tests such as the one used in the main experiment (evaluating a famous faces), are more sensitive to unconscious processing. If subjects performed above chance (binomial test; $p < .2$) in either of these two awareness tests, their experimental data were excluded from analysis.

Page 6

Test of knowledge of celebrities. Because our participants were young and unexperienced, we needed to ensure that they knew the politicians and actors used in the second part of the fMRI experiment. An interaction between unconscious and conscious retrieval could only occur if participants were able to identify our portraits of celebrities. To this end, participants were instructed at the end of the session to classify the previously used portraits of celebrities according to "politician" and "actor" and to retrieve the celebrities' names. If participants claimed to know a celebrity but failed to retrieve his name, they described the celebrity and/or where they knew them from to prove identification. All participants but two were able to identify the celebrities. The experimental data of those two participants who failed were excluded from analysis.

MRI data acquisition

Anatomical and functional images were acquired with a 3T Siemens Magnetom Trio whole-body scanner. Anatomical T1-weighted image acquisition followed a 3D-gradient echo-sequence with a spatial resolution of $1 \times 1 \times 1$ mm^3 (acquisition matrix = 256×256 voxels, 176 sagittal slices; time of repetition (TR) = 7.92 ms; echo time (TE) = 2.48 ms; flip angle (FA) = 16°; field of view (FOV) = 256×256 mm^2). Structural image acquisition was carried out during the awareness tests.

Functional T2*-weighted images were acquired using a blood-oxygen-level-dependent (BOLD) sensitive, interleaved 2D-gradient echo planar single-shot pulse (EPI) sequence with a spatial resolution of $1.8 \times 1.8 \times 4$ mm^3 (acquisition matrix = 128×128 voxels, 34 transversal slices; TR = 4000 ms; TE = 32 ms; FA = 90°; FOV = 230×230 mm^2).

Behavioral data analysis

Choice reaction times (RT) acquired during the interaction task were analysed with IBM SPSS (version 20). Trials with RT deviating more than 2 *SD* from the individual mean were excluded. Because RTs were not normally distributed (Kolmogorov-Smirnov tests, all $p < .001$; skewness > 0), nonparametric statistics were computed (Wilcoxon signed rank exact test). However, parametric testing yielded comparable results.

fMRI data analysis

Preprocessing of volumes was carried out with the software SPM8 (Wellcome Department of Cognitive Neurology, London, UK). Volumes were slice-time corrected, realigned to the first volume, coregistered to the anatomical volume, normalized to the MNI T1 template and finally smoothed with an 8 mm (FWHM) isotropic Gaussian kernel.

First, we computed independent component analyses (ICA) and correlated the extracted components of brain activity with the time-course of the alternations between condition blocks in each of the two fMRI time-series, i.e., the encoding time-series and the interaction time-series. This analysis yields a model-and hypothesis-free estimate of functionally coupled brain areas that were engaged during unconscious associative encoding and retrieval. We computed group-level ICAs using the GIFT toolbox (http://mialab.mrn.org/software/gift/index.html). The optimal number of independent components was estimated according to the minimum description length criteria [22] in advance of the actual analysis, which was run with the Info-max algorithm [23]. This procedure resulted in the extraction of 17 independent components for the encoding time-series and 19 independent components for the interaction time-series. We were interested in components reflecting unconscious relational memory processes that co-vary with the occurrence of associative condition blocks. Independent components were thus sorted with respect to their regression fit with the modelled time course of associative condition

blocks. Associative condition blocks contained *Face-Occupation Pairs* at encoding and *Congruent* and *Incongruent Faces* at test. False discovery rate (FDR)-corrected [24] one-sample t-tests were computed on the β-weights of sorted components to determine whether a component was significantly associated with a time-series. Significant components were subsequently checked for a-priori regions of interest, namely hippocampus and lateral- and polar temporal neocortices. Significant components containing a-priori regions of interest where then tested for a regression fit with their baseline condition (i.e., *Faces Alone* at encoding; *Old Faces* at test) to ensure that functional coupling of these components was specific for unconscious associative memory processes. Hence, a non-significant regression fit with baseline conditions was expected. Cluster statistics were calculated in SPM8 with a height threshold of $p = .05$ (family-wise-error corrected). Component images were thresholded at $Z > 2$ for visualisation. Labelling and visual inspection of the activation patterns was carried out with xjView8 (http://www.alivelearn.net/xjview8/).

While ICA is able to uncover global-scale networks, it only allows to plot the strength of association of single voxels with these networks, but is limited in providing insight into how local neural groups relate to behaviour directly. Therefore, we regressed retrieval performance (reaction time differences) onto fMRI contrasts to reveal signal changes that relate linearly to the behavioural evidence of unconscious-conscious retrieval interactions. SPM8 was used for first and second level analyses of contrasts between conditions. In the first level analysis, the time-series of each participant were modelled with a box car function convolved with a canonical hemodynamic response function. In the second level analysis, group level statistics were computed on first level contrasts using within-subject one-way ANOVAs. We entered the RT-differences recorded at test as a covariate of interest into the second-level GLM.

For subliminal encoding, the contrast (*Face-Occupation Pairs > Faces Alone* baseline) was correlated with the difference in reaction times in the incongruent versus congruent condition. This RT difference was chosen as a regressor over the RT difference of *Old Faces* baseline—*Congruent* because there is no difference between a prospectively congruent and incongruent face at the time of encoding, and because each *Face-Occupation Pair* was used in both the *Incongruent* and the *Congruent* condition.

For the interaction test, the contrast (*Congruent > Old Faces* baseline) was correlated with the difference in reaction times recorded in the *Congruent* versus the *Old Faces* condition. Both this RT measure and the fMRI contrast reflect facilitating interactions between unconscious and conscious associative retrieval. The fMRI data were not analysed regarding interfering interactions (*Incongruent* condition) because the behavioural data (see below) showed no evidence of interference between unconscious and conscious retrieval. No corrections for multiple comparisons were applied due to the small signals associated with unconscious processing [7,8,19]. The height threshold was $p = .001$ for the whole brain and $p = .005$ for the hippocampus, which was the a priori key region of interest. The extent threshold was four voxels. Labelling and visual inspection of the activation patterns was carried out with xjView8.

Results

Awareness tests

Participants were oblivious of both the fact of subliminal stimulation and the subliminal stimuli. For the analysis of data obtained in the two objective awareness tests, we took a conservative approach analysing the data of each individual using binomial testing. Participants with a performance above the upper 20%- cut-off of the chance distribution of correct responses were considered potentially aware of subliminal stimuli. Their data acquired in the fMRI experiment were therefore excluded from analysis. The 20%-cut-off corresponded to a hit rate of

56% (50% = chance level). Seven participants performed better than expected by chance either on the face or on the occupation awareness test. The remaining participants performed at chance level as individuals and as a group on both the face awareness test (49 ± 7.3% ($M ± SD$); one-sample t-test against 0.5: $t(32) = -.583$, $p = .564$) and the occupation awareness test (49 ± 7.2%; $t(32) = -.527$, $p = .603$). Hence, these remaining participants were unable to consciously detect subliminal faces or words or fragments thereof.

Main Experiment: Behavioural performance

Subliminal processing and attention task. Participants were simultaneously processing two different streams of information at the unconscious and the conscious level. At the conscious level, participants engaged in the attention task. At the unconscious level, they processed subliminal faces and written occupations. We collected behavioural data on the attention task and calculated accuracy scores. Participants performed the attention task with high accuracy (hit rate = 94 ± 23.7%, $M ± SD$) indicating that they focused gaze at the middle of the screen and paid attention to the masked presentations during the whole stimulation sequence.

Interaction of unconscious with conscious retrieval of occupations. In the critical retrieval interaction test, participants responded to the presentation of portraits of famous actors and politicians by manually indicating their occupational category "actor" versus "politician". Two participants performed poorly (46.9% and 65.6% correct) because they were not familiar with the celebrities; these two participants were excluded from data analysis. The remaining participants identified celebrities with 91 ± 10% ($M ± SD)$ (*Congruent*), 93 ± 9% (*Incongruent*) and 92 ± 8% (*Old Faces*) correct responses. Because accuracy of choice did not differ between conditions ($F(2,64) = 1.315$, $p = .276$), reaction times were the dependent variable that could be modulated by the preceding unconscious retrieval processes.

A Wilcoxon exact test revealed a significant difference in reaction times between the congruent and the incongruent condition ($Z = -2.850$, $p = .002$, one-tailed, effect size $r = .50$) with faster responses to congruent versus incongruent famous faces. We further investigated whether this effect was due to congruence gains or incongruence costs by comparing the two conditions to the non-associative baseline condition *Old Faces*. This analysis showed that response latencies were significantly shorter to *Congruent* than *Old Faces* ($Z = -1.689$, $p = .047$, one-tailed, $r = .29$). There was no statistical difference between *Incongruent* and *Old Faces* ($Z = -.777$, $p = .437$) ([Fig 2](#)). Hence, the above behavioural effect was due to congruence gains rather than incongruence costs. The $M ± SD$s of the RTs in the interaction task were: 1279 ± 313 ms (*Congruent*), 1351 ± 352 ms (*Incongruent*) and 1337 ± 373 ms (*Old Faces* baseline). In conclusion, unconscious-conscious interactions were only apparent in the *Congruent* condition, where the unconsciously encoded and retrieved occupations were identical with the consciously retrieved occupations.

As all former subliminal faces in the congruent condition were also used in the incongruent condition and vice versa, a bias could arise due to their repeated presentation. This was, however, not the case: Wilcoxon exact tests showed that RTs to famous faces did not differ between the first versus second presentation of face cues in the congruent condition ($Z = -1.099$, $p = .280$, two-tailed) nor the incongruent condition ($Z = -0.116$, $p = .916$, two-tailed).

Main Experiment: fMRI data

Independent component analyses. We performed an independent component analysis (ICA) on the fMRI data acquired during the encoding fMRI time-series to explore the functional connectivity of brain regions during the subliminal processing of *Face-Occupation Pairs*. The subliminal presentation of *Face-Occupation Pairs* was associated with decreased activity in
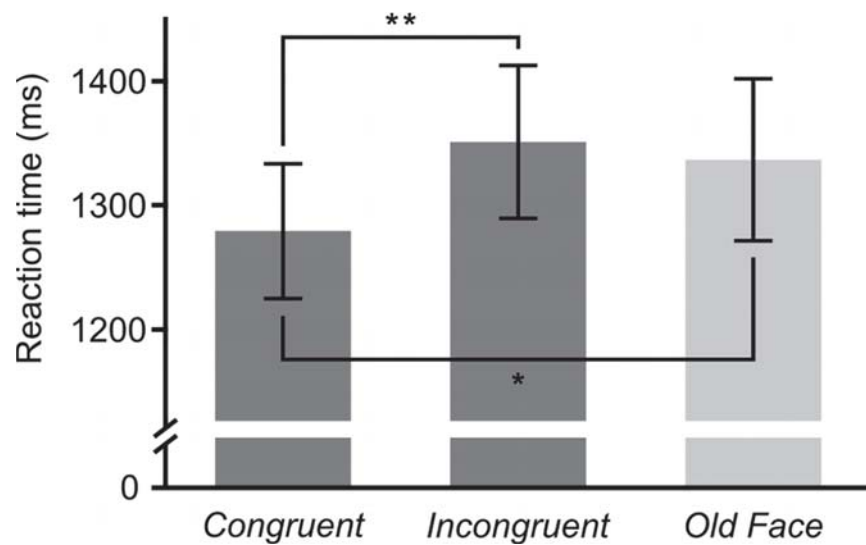
PLOS | ONE

Fig 2. Reaction times at categorizing famous faces. Group means and SEM are displayed. *Old Face* trials (no association) are used as baseline. * Mean difference *(ΔM)* = 57 ms, *p* = .047, effect size *r* = .29; ** *ΔM* = 71 ms, *p* = .002, *r* = .50; one-tailed Wilcoxon signed rank exact tests.

doi:10.1371/journal.pone.0122459.g002

a number of functionally connected brain areas that constituted one of the obtained components (*r* = -.16, *t*(32) = -3.1, *p* = .004 < FDR critical *p* = .006) (Table 1 and Fig 3). This component included bilateral areas in the superior temporal sulcus (extending into superior and middle temporal gyrus) and temporal pole, which harbour storage sites of lexical-semantic information such as occupations [16]; bilateral hippocampus and ventromedial thalamus, required for encoding of new information [25]; and bilateral amygdala, which is considered to play an important role in face perception and evaluation. As faces convey highly significant

Table 1. Idependent component analysis of subliminal encoding: Functional network coupled with presentation of subliminal *Face-Occupation Pairs* (*r* = -.16, *p* = .004).

Region of activation	L/R	Brodmann area	X	Y	Z	N of voxels	T
Insula, Superior temporal sulcus (extending into Superior and Middle temporal gyrus), Temporal pole	L	21, 22, 34, 35, 38, 47	-40	10	-16	4260	15.98
Brain Stem (multiple local maxima)	bilat		6	-28	-20		15.79
Hippocampus, Amygdala	L		-20	-12	-16		9.12
Hippocampus, Amygdala	L		-26	-10	-20		8.58
Entorhinal cortex	L	28	-22	-20	-24		8.54
Insula	L	13	-42	-2	-4		7.17
Superior temporal sulcus (extending into superior temporal gyrus), Temporal pole	R	21, 22, 28, 34, 38, 47	40	8	-14	850	12.45
Insula	R	13	42	0	-6		8.73
Hippocampus, Amygdala	R		24	-12	-16		8.01
Hippocampus, Amygdala	R		30	-6	-22		7.40
Amygdala	R		30	-6	-16		7.35
Cerebellum, vermis	L		-4	-68	-18	29	8.80
Ventromedial thalamus	bilat		2	-12	-8	31	7.55

p <. 05 (FWE). L, left; R, right; bilat, bilateral.

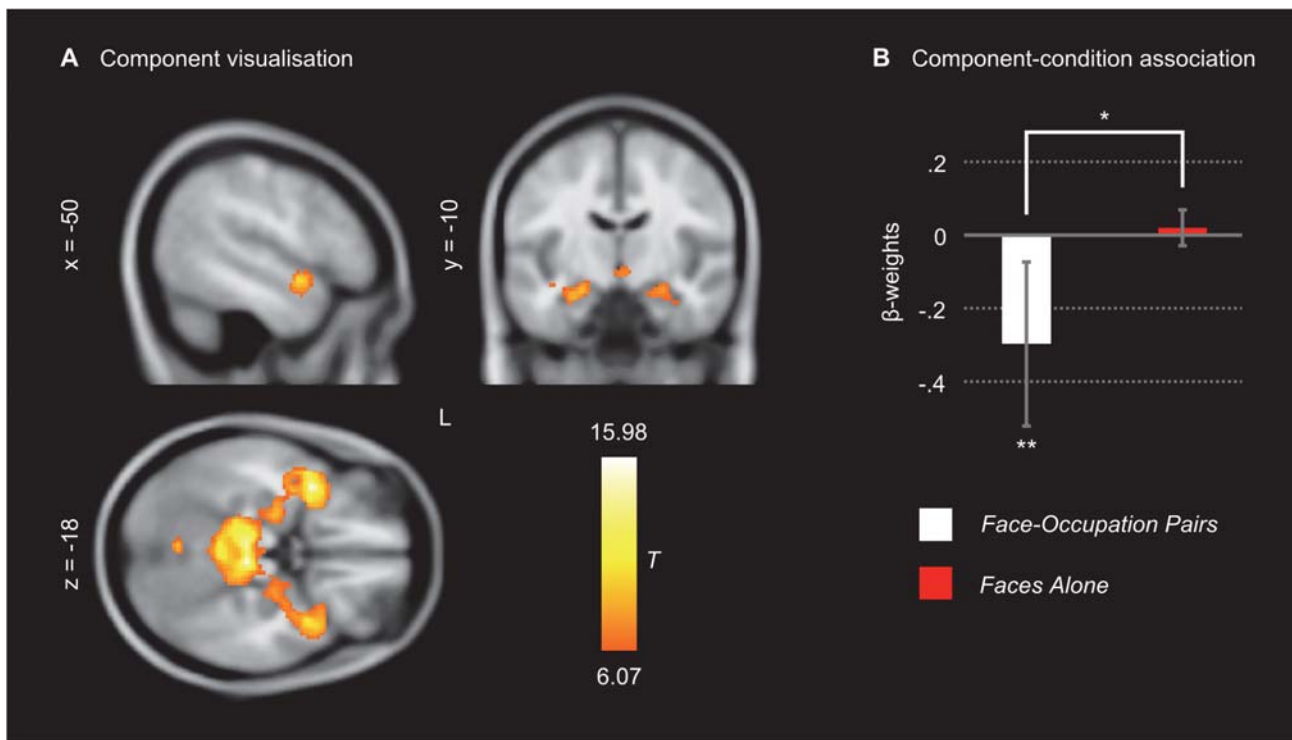doi:10.1371/journal.pone.0122459.t001

Fig 3. Independent component analysis (ICA) on encoding time-series. The depicted component is significantly associated with the occurrence of subliminal *Face-Occupation Pairs*. **A:** Clusters within the component encompassing bilateral hippocampus, amygdala, superior temporal sulcus, and temporal pole. These brain regions are important for episodic and semantic memory. Coordinates are according to MNI space; left is left on the coronal slice and upwards is left on the transversal slice. **B:** The temporal coupling of this network is specific for unconscious associative encoding. The component is significantly associated with the occurrence of *Faces-Occupation Pairs* (Pearson's r = -.16, ** p = .004) but not with *Faces Alone* (Pearson's r = .13, p = .69). The regression fit of the component is significantly better with *Face-Occupation Pairs* than with *Faces Alone* (* p = .044, effect size r = .35). The bar plot shows β-weights of the time course modelled specifically to the associative and the baseline condition. Error bars indicate the SEM.

doi:10.1371/journal.pone.0122459.g003

social and emotional information, the amygdala is automatically engaged when faces are perceived [26]. Importantly, this component did not covary with the *Faces Alone* baseline (r = .13, $t(32)$ = 0.4, p = .69), and the regression-fit of this component with *Face-Occupation Pairs* was significantly better than with *Faces Alone* ($t(32)$ = -2.1, p = .044, effect size r = .35) (Fig 3B). In conclusion, we can assume that this component was specifically related to the semantic associative binding of subliminal faces with written occupations.

The negatively (rather than positively) deflected fMRI signal during subliminal relational versus single face encoding calls for an explanation. Such negative deflections are in fact a replicable phenomenon observed during subliminal associative encoding relative to a non-associative baseline [6,7,27]. The hippocampus is active whenever an event is experienced [28] and also during rest because it retrieves and stores memories in the stream of spontaneous conscious mentation [29]. During subliminal encoding, the hippocampus may split its processing capacity between conscious spontaneous mentation and unconscious encoding of subliminal stimuli. Because backward masks interrupt the firing response of activated neurons [30,31], the processing of subliminal face-occupation pairs gets interrupted by backward masks. The more neurons are recruited to encode subliminal stimuli instead of spontaneous conscious thoughts, the more spiking activity is interrupted in the hippocampus, which reduces the fMRI signal. Relative to a relational condition, where many hippocampal neurons are recruited for unconscious encoding, a non-relational baseline condition frees hippocampal neurons from

Table 2. Subliminal encoding-related fMRI signal correlates with behavioural facilitation during the interaction test.

Region of activation	L/R	Brodmann area	X	Y	Z	N of voxels	T	$r_{cluster}$
Negative correlation: *Faces Alone > Face-Occupation Pairs × ΔRT(Incongruent—Congruent)*								
Lentiform nucleus	L		-22	2	2	27	4.14	-.589
Hippocampus *	L		-32	-10	-14	147	3.92	-.564
Amygdala	L		-26	-2	-20			
Lentiform nucleus	R		24	2	-2	20	3.77	-.582
Insula, frontal operculum	R	13 / 44	46	-4	12	8	3.60	-.545
Positive correlation: *Face-Occupation Pairs > Faces Alone × ΔRT(Incongruent—Congruent)*								
No Suprathreshold clusters								

p < .001 (unc.);
*p < .005 (unc.).
L, left; R, right; ΔRT, reaction time difference

doi:10.1371/journal.pone.0122459.t002

subliminal processing and makes them available for encoding the stream of spontaneous conscious thoughts. Hence, the firing of more hippocampal neurons is interrupted in the experimental than the baseline condition reducing the fMRI signal.

We also performed an ICA on the fMRI data acquired during the interaction time-series. This ICA yielded no significant component. Because all experimental conditions included the conscious inspection of famous faces and the conscious retrieval of their occupations—i.e., tasks associated with strong signal changes in the brain—, we suspect that the superimposed signal changes associated with unconscious processes were too weak and sparse for a clear modulation of the global signal.

Correlation of fMRI data with behavioural performance. The reaction time difference *Incongruent—Congruent* was regressed onto the subliminal encoding contrast (*Face-Occupation Pairs > Faces Alone*) to reveal brain activation underlying successful unconscious associative encoding (Table 2). Activity reductions during subliminal associative encoding were related to faster responses at test in the congruent versus incongruent condition. Significant correlations were located in a large area that included the left hippocampus and amygdala (r = -.564; Fig 4A) corroborating the results of the subliminal encoding ICA, and extending them by linking brain activity at encoding with behavioural facilitation at test. Further inverse correlations were located in the right frontal operculum/insula (BA 44/13) and bilateral lentiform nuclei, i.e., putamen and globus pallidus. Activity in or near these two regions has been shown to be associated with word reading [32], more consistently within the left, rather than the right hemisphere [33]. It should be noted that unconscious information processing is often strongly supported by the right hemisphere, especially if the encoded information is emotionally relevant [34]. Hence, unconsciousness of processing might explain why we found a right hemisphere focus of activation in areas usually displaying left hemisphere dominance.

We also regressed reaction time differences at test onto brain activity underlying the interaction of unconscious with conscious associative retrieval. Because the behavioural data indicated that the unconscious-conscious retrieval interaction yielded only congruence gains, and no incongruence costs, we focussed on congruence effects. The reduction in response time in the *Congruent* versus *Old Face* condition was regressed onto the fMRI contrast (*Congruent > Old Faces*). Significant negative correlations (Table 3) were located in the right posterior hippocampus (r = -.475; Fig 4B), the left temporal pole (BA 38) (r = -.572), and within the posterior end of the right superior temporal sulcus extending into the angular region (BA 22 & 39) (r = -.556). The posterior superior temporal sulcus is a face-responsive region often associated
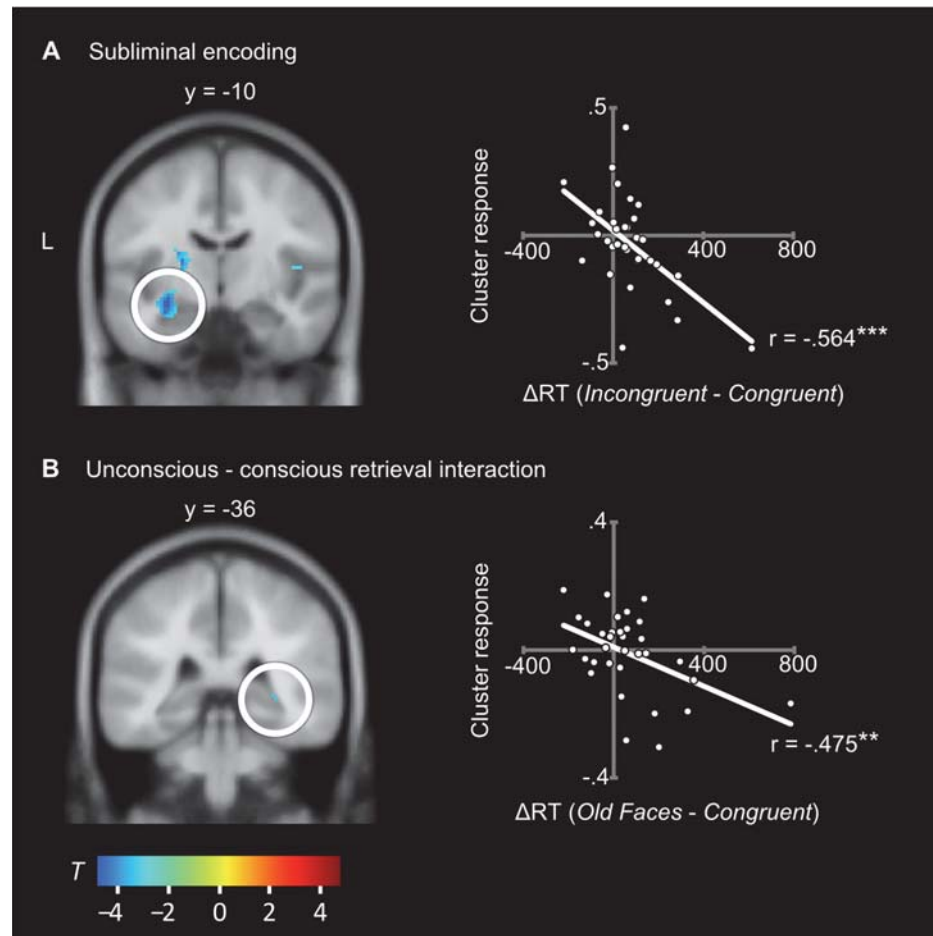
Fig 4. Hippocampal activity relates to retrieval performance. Locations of significant correlations are displayed on the left side of the figure. The circled hippocampal clusters correspond to the respective scatterplots displayed on the right side of the figure. Pearson-correlation coefficients are included. **A:** Correlation of the encoding contrast (*Face-Occupation Pairs > Faces Alone*) with the reaction time difference (ΔRT) between the *Incongruent* and *Congruent* condition. **B:** Correlation of the interaction contrast (*Congruent > Old Faces*) with the ΔRT between the *Congruent* and the *Old Faces* condition. Coordinates are in MNI space; the left side of the image corresponds to the left side of the brain. Warm colours indicate positive correlations (none present), cold colours negative correlations. ** $p < .01$, *** $p < .001$.

doi:10.1371/journal.pone.0122459.g004

with theory of mind [26,35,36], and the temporal poles have been shown to specifically account for person-related semantics such as occupations [17]. That activity in these areas predicts the magnitude of the congruence effect suggests that these areas supported the unconscious recognition of the former subliminal faces and the unconscious reactivation of associated knowledge (occupations), which could then facilitate the conscious retrieval of the famous individual's occupations. We will discuss below why this correlation was negative and what its theoretical implications are.

Discussion

Based on previous evidence that experienced episodes can be encoded with and without consciousness and recruit the hippocampal anterior-thalamic axis and related cortices in both cases [5–8], we hypothesized that consciously and unconsciously acquired relational memories

Table 3. Retrieval-related fMRI signal correlates with behavioural facilitation during the interaction test.

Region of activation	L/R	Brodmann area	X	Y	Z	N of voxels	T	$r_{cluster}$
Negative correlation: Old Faces > Congruent × ΔRT(Old Faces—Congruent)								
Temporal pole	L	38	-42	20	-26	4	3.87	-.572
Superior temporal s / angular g	R	39 / 22	44	-56	14	4	3.76	-.556
Precuneus / Calcarine sulcus	R	31 / 18	2	-72	18	22	3.71	-.583
Hippocampus *	R		32	-36	0	5	2.87	-.475
Positive correlation: Congruent > Faces Alone × ΔRT(Old Faces—Congruent)								
No suprathreshold clusters								

$p < .001$ (unc.);

$*p < .005$ (unc.);

L, left; R, right; ΔRT, reaction time difference; s, sulcus; g, gyrus.

doi:10.1371/journal.pone.0122459.t003

are harboured within a single, cohesive hippocampal-neocortical memory space, where they interact with each other. The reactivation of an unconsciously acquired relational memory facilitated the subsequent conscious retrieval of a semantically congruent relational memory. This facilitation was reflected in shortened reaction times and simultaneously recorded reductions in neural activation within hippocampus and neocortical storage sites thought to harbour lexical-semantic and person identity information. In the following, we discuss potential mechanisms that may underlie this facilitative unconscious-conscious retrieval interaction.

Our results point to facilitatory unconscious-conscious retrieval interactions in the congruent condition of our fMRI experiment. The supraliminal presentation of a non-famous person, who was previously presented subliminally with an occupation (e.g., politician), led to reduced response times to celebrities that share this occupation (e.g., also politician). The shortening of response times indicates that the celebrities' occupations were preactivated by the presentation of the former subliminal faces. Accordingly, subliminal face-occupation combinations must have been encoded and stored in the first place. This finding replicates previous demonstrations of the feasibility of subliminal semantic paired-associative encoding and long-term storage using face-occupation combinations [6,7,19,37] and word pairs [8,38–40].

Savings in response times in the congruent condition went along with modulations of neural activation within hippocampus, temporal pole, superior temporal sulcus, angular gyrus, and precuneus. These neural effects were probably due to conscious rather than unconscious retrieval processes because signals associated with conscious versus unconscious mental processes are much stronger [41–43]. We suggest two possible mechanisms that may have caused response time and neural savings in the congruent condition (Fig 5).

1) Facilitation in the congruent condition may have occurred through conceptual priming of occupational knowledge stored in the lateral temporal lobe. Following the presentation of the former subliminal face and identification through face recognition units in the fusiform face area, the face-associated occupation (e.g., politician) was retrieved through hippocampal processes (Fig 5.1.) that in turn activated occupation-relevant storage sites in the lateral temporal lobe (Fig 5.2.). This preactivation of occupational knowledge sites in the lateral temporal lobe may then have primed the conscious retrieval of the famous person's occupation (e.g., politician) reducing net activation in the lateral temporal lobe through repetition suppression due to neural sharpening or facilitation [44,45]. The reduced activation in hippocampus can be explained in terms of a sparse hippocampal recruitment for recovering the preactivated occupation of the famous face. The conceptual preactivation curtailed any unnecessary hippocampal search processes, which were necessary in the baseline condition increasing the hippocampal
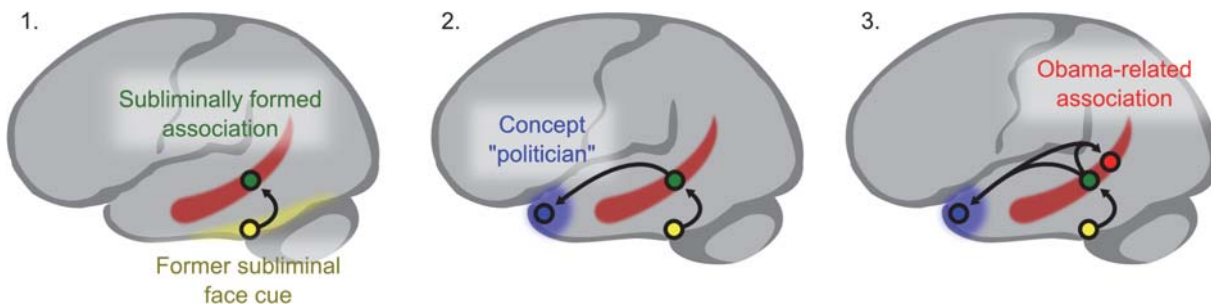
Fig 5. Network model of the assumed unconscious-conscious retrieval interaction. We suggest an intrahippocampal interaction mechanism as cause for the congruence effects: the supraliminal presentation of the former subliminal face elicits unconscious face recognition activating the fusiform gyrus. The fusiform signal triggers the hippocampal reactivation of the face-associated occupation (e.g., politician) (**1**), which in turn activates occupational knowledge (politician) in the lateral temporal lobe (**2**). The activated hippocampal relational engram coactivates other overlapping engrams; e.g., memories of other politicians. This intra-hippocampal preactivation facilitates the retrieval of a celebrity's occupation (**3**). If a portrait of Obama were presented in the congruent condition, hippocampal and lateral temporal activity would be reduced compared to the baseline condition, where a hippocampal ab-initio activation would build up. Incongruence costs are not to be expected because the preceding hippocampal retrieval of a professional with another occupation (actor) would leave non-overlapping politician-related hippocampal memories unaffected. This scenario would support the view that consciously and unconsciously acquired memories are organized in a single, cohesive hippocampal-neocortical memory space with memories organized relative to their contents. Overlapping memories are linked, which supports pattern completion, abstraction and anticipation.

signal. Because the congruent condition was contrasted with the baseline condition, the activation level was relatively reduced. This mechanistic explanation of the unconscious-conscious interaction is, however, flawed by the absence of an inhibitory unconscious-conscious interaction in the incongruent condition. If conceptual priming was the crucial mechanism, we should have observed negative priming [46] in the incongruent condition (i.e. slower reactions). Because no incongruence costs occurred in the incongruent condition, conceptual priming is probably not the only mechanism underlying the unconscious-conscious interaction. A further reason why conceptual priming is unlikely to be the only mechanism at work is evidence in amnesic patients that the hippocampus is necessary for the relational encoding and retrieval of subliminal item pairs [4]. Accordingly, it can be assumed that both neocortex and hippocampus were involved in the unconscious and conscious retrieval of face-occupation associations.

2) Another mechanism seems therefore more likely, which assumes an intrahippocampal interaction as an additional cause for the congruence effects. According to this scenario, the supraliminal presentation of the former subliminal face elicited unconscious face recognition activating the fusiform gyrus. The fusiform signal triggered the hippocampal reactivation of the face-associated occupation (e.g., politician) (Fig 5.1.), which in turn activated occupational knowledge (politician) in the lateral temporal lobe (Fig 5.2.). The activated hippocampal relational engram coactivated other overlapping engrams [47], e.g., memories of other politicians. This intra-hippocampal preactivation facilitated the retrieval of the presented celebrity's occupation (Fig 5.3.). This second scenario is in line with known characteristics of the hippocampal memory system: the hippocampal memory system forms relational networks of memory traces that share aspects. This organizational structure permits an activated memory trace to trigger the activation of memory traces that share aspects and hence overlap [47]. E.g., the presentation of Obama's portrait in the congruent condition would be accompanied by reduced activity in the hippocampus and lateral temporal lobe due to the semantic overlap of unconscious and conscious memory traces. The preactivation of the overlapping neural populations in hippocampus during unconscious retrieval allows for a more sparing activation during conscious retrieval. In the baseline condition, no unconscious relational memories are formed that could be reactivated at test. Thus, in the baseline condition the hippocampal search process builds up fully. Incongruence costs are not to be expected because the preceding hippocampal retrieval

of, say, an actor would leave non-overlapping politician-related hippocampal memories unaffected. This scenario is analogous to retrieval-induced forgetting, where a partial retrieval of information can impair the subsequent retrieval of the remaining information, if the remembered and forgotten information comes from the same semantic category [48]. This second scenario is also likely in view of earlier findings of a hippocampal role in unconscious relational encoding/retrieval [6]. If this interpretation is correct, the finding suggests that consciously and unconsciously acquired memories are organized in a single, cohesive hippocampal-neocortical memory space. There is evidence that memories are organized topologically within the hippocampus relative to their contents, with more closely related engrams represented increasingly overlapping neural populations [49]. Linked overlapping memories support pattern completion, abstraction and anticipation [47] and newly encoded information is readily integrated into pre-existing relational networks [50]. The degree of representation from consciously accessible to inaccessible memories is presumably orthogonal to the content-based organization of hippocampal memories [4].

A synergistic unconscious-conscious interaction may be counterintuitive when considering previous reports of competing interactions between implicit and explicit memories [51,52]. In these earlier studies, however, declarative memory was compared to either procedural memory or priming, managed by hippocampus, basal ganglia and neocortex, respectively. Consequently, competing memory interactions may have occurred because unconscious and conscious learning mechanisms did not share the same memory system. Conversely, the interaction in the current study was harmonious because both unconscious and conscious relational memories were supported by the hippocampus.

The current study design has its limitations. It does not allow the isolation of neural activity underlying unconscious versus conscious retrieval because the rapid succession (500 ms) of the non-famous face cue for unconscious reactivation and the famous face cue for conscious retrieval results in a blurring of signals. Therefore, we can only speculate about the mechanisms underlying the facilitatory unconscious-conscious interactions. A further limitation is that our portraits of celebrities might tap semantic information [14] besides episodic memories. Hence, the probed memory system cannot be determined beyond doubt. Yet, during the test of knowledge of celebrities it became clear that our participants were not overly familiar with many of the used famous faces and had to draw on their episodic memory. Furthermore, if semantic person knowledge was sufficient to recall occupations, unconscious-conscious interactions would likely not have modulated hippocampal signals but neocortical signals alone [14].

The classic view of the hippocampal memory system holds that consciousness is required for episodic memory formation [1–3]. However, the unconscious-conscious retrieval interaction reported here suggests that conceptually overlapping unconscious and conscious memories are stored in close association within hippocampus. An intertwined store of consciously accessible and consciously inaccessible relational hippocampal memories is compatible with the processing based memory model [5]. Also, a single, cohesive hippocampal memory space for any level of representation—unconscious to conscious—is evolutionarily sensible. As pointed out earlier, episodic memories may shift from a conscious to an unconscious representation and vice versa over time [9–11]. In both these cases, a cohesive memory space provides for a stable organizational structure of hippocampal memories. Such representational shifts appear difficult if one assumes a strict division between memory systems based on conscious access. It is more economical to assume one hippocampal memory system that serves one computational goal, namely rapidly establishing new flexible associations, irrespective of conscious access [5].

Author Contributions

Conceived and designed the experiments: TPR PV KH. Performed the experiments: PC SR. Analyzed the data: MAZ. Contributed reagents/materials/analysis tools: TPR MH. Wrote the paper: MAZ TPR PV SR KH.

References

1. Tulving E. How many memory systems are there? Am Psychol. 1985; 40(4):385–98.

2. Squire LR, Zola SM. Structure and function of declarative and nondeclarative memory systems. Proc Natl Acad Sci. 1996 Nov 26; 93(24):13515–22. PMID: 8942965

3. Moscovitch M. The hippocampus as a "stupid," domain-specific module: Implications for theories of recent and remote memory, and of imagination. Can J Exp Psychol Rev Can Psychol Expérimentale. 2008 Mar; 62(1):62–79.

4. Duss SB, Reber TP, Hänggi J, Schwab S, Wiest R, Müri RM, et al. Unconscious relational encoding depends on hippocampus. Brain. 2014 Sep 27;3355–70.

5. Henke K. A model for memory systems based on processing modes rather than consciousness. Nat Rev Neurosci. 2010; 11(7):523–32. doi: 10.1038/nrn2850 PMID: 20531422

6. Degonda N, Mondadori CRA, Bosshardt S, Schmidt CF, Boesiger P, Nitsch RM, et al. Implicit Associative Learning Engages the Hippocampus and Interacts with Explicit Associative Learning. Neuron. 2005 May; 46(3):505–20. PMID: 15882649

7. Henke K, Mondadori CRA, Treyer V, Nitsch RM, Buck A, Hock C. Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. Neuropsychologia. 2003 Jan; 41(8):863–76. PMID: 12667523

8. Reber TP, Luechinger R, Boesiger P, Henke K. Unconscious Relational Inference Recruits the Hippocampus. J Neurosci. 2012 May 2; 32(18):6138–48. doi: 10.1523/JNEUROSCI.5639-11.2012 PMID: 22553020

9. Drosopoulos S, Harrer D, Born J. Sleep and awareness about presence of regularity speed the transition from implicit to explicit knowledge. Biol Psychol. 2011 Mar; 86(3):168–73. doi: 10.1016/j.biopsycho.2010.11.005 PMID: 21163328

10. Wagner U, Gais S, Haider H, Verleger R, Born J. Sleep inspires insight. Nature. 2004 Jan 22; 427(6972):352–5. PMID: 14737168

11. Yordanova J, Kolev V, Wagner U, Born J, Verleger R. Increased Alpha (8–12 Hz) Activity during Slow Wave Sleep as a Marker for the Transition from Implicit Knowledge to Explicit Insight. J Cogn Neurosci. 2011 Aug 3; 24(1):119–32. doi: 10.1162/jocn_a_00097 PMID: 21812555

12. Anderson MC, Green C. Suppressing unwanted memories by executive control. Nature. 2001 Mar 15; 410(6826):366–9. PMID: 11268212

13. Saletin JM, Goldstein AN, Walker MP. The Role of Sleep in Directed Forgetting and Remembering of Human Memories. Cereb Cortex N Y NY. 2011 Nov; 21(11):2534–41. doi: 10.1093/cercor/bhr034 PMID: 21459838

14. Harand C, Bertran F, La Joie R, Landeau B, Mézenge F, Desgranges B, et al. The Hippocampus Remains Activated over the Long Term for the Retrieval of Truly Episodic Memories. Johnson EC, editor. PLoS ONE. 2012 Aug 24; 7(8):e43495. doi: 10.1371/journal.pone.0043495 PMID: 22937055

15. Phillips PJ, Wechsler H, Huang J, Rauss PJ. The FERET database and evaluation procedure for face-recognition algorithms. Image Vis Comput. 1998; 16(5):295–306.

16. Binder JR, Desai RH, Graves WW, Conant LL. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. Cereb Cortex. 2009 Dec 1; 19(12):2767–96. doi: 10.1093/cercor/bhp055 PMID: 19329570

17. Tsukiura T, Mochizuki-Kawai H, Fujii T. Dissociable roles of the bilateral anterior temporal lobe in face–name associations: An event-related fMRI study. NeuroImage. 2006 Apr; 30(2):617–26. PMID: 16275140

18. Chapman LJ, Chapman JP. The measurement of handedness. Brain Cogn. 1987 Apr; 6(2):175–83. PMID: 3593557

19. Henke K, Treyer V, Nagy ET, Kneifel S, Dürsteler M, Nitsch RM, et al. Active hippocampus during nonconscious memories. Conscious Cogn. 2003 Mar; 12(1):31–48. PMID: 12617861

20. Snodgrass M, Shevrin H. Unconscious inhibition and facilitation at the objective detection threshold: Replicable and qualitatively different unconscious perceptual effects. Cognition. 2006 Aug; 101(1):43–79. PMID: 16289068

21. Reingold EM, Merikle PM. Using direct and indirect measures to study perception without awareness. Percept Psychophys. 1988 Dec; 44(6):563–75. PMID: 3200674

22. Li Y-O, Adalı T, Calhoun VD. Estimating the number of independent components for functional magnetic resonance imaging data. Hum Brain Mapp. 2007; 28(11):1251–66. PMID: 17274023

23. Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 1995 Nov; 7(6):1129–59. PMID: 7584893

24. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995 Jan 1; 57(1):289–300.

25. Aggleton JP, Brown MW. Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. Behav Brain Sci. 1999 Jun; 22(3):425–44; discussion 444–89. PMID: 11301518

26. Todorov A. The role of the amygdala in face perception and evaluation. Motiv Emot. 2012 Mar; 36 (1):16–26. PMID: 22448077

27. Henke K, Turi Nagy E, Berthold T, Weber B, Buck A. H2 15O PET imaging of conscious and unconscious encoding/retrieval of semantic associations. Society for Neuroscience Abstracts. 1999. p. 647.

28. Martin A. Automatic activation of the medial temporal lobe during encoding: lateralized influences of meaning and novelty. Hippocampus. 1999; 9(1):62–70. PMID: 10088901

29. Stark CEL, Squire LR. When zero is not zero: The problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci. 2001 Oct 23; 98(22):12760–6. PMID: 11592989

30. Kovács G, Vogels R, Orban GA. Cortical correlate of pattern backward masking. Proc Natl Acad Sci. 1995 Jun 6; 92(12):5587–91. PMID: 7777553

31. Rolls ET, Tovee MJ. Processing speed in the cerebral cortex and the neurophysiology of visual masking. Proc Biol Sci. 1994 Jul 22; 257(1348):9–15. PMID: 8090795

32. Fiez JA, Petersen SE. Neuroimaging studies of word reading. Proc Natl Acad Sci. 1998 Feb 3; 95 (3):914–21. PMID: 9448259

33. Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann N Y Acad Sci. 2010 Mar 1; 1191(1):62–88. doi: 10.1111/j.1749-6632.2010.05441.x PMID: 20392272

34. Gainotti G. Unconscious processing of emotions and the right hemisphere. Neuropsychologia. 2012 Jan; 50(2):205–18. doi: 10.1016/j.neuropsychologia.2011.12.005 PMID: 22197572

35. Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends Cogn Sci. 2000 Jun 1; 4(6):223–33. PMID: 10827445

36. Frith U, Frith CD. Development and neurophysiology of mentalizing. Philos Trans R Soc B Biol Sci. 2003 Mar 29; 358(1431):459–73.

37. Duss SB, Oggier S, Reber TP, Henke K. Formation of semantic associations between subliminally presented face-word pairs. Conscious Cogn. 2011 Sep; 20(3):928–35. doi: 10.1016/j.concog.2011.03.018 PMID: 21481607

38. Reber TP, Henke K. Rapid formation and flexible expression of memories of subliminal word pairs. Conscious Res. 2011; 2:343. PMID: 21520509

39. Reber TP, Henke K. Integrating unseen events over time. Conscious Cogn. 2012 Jun; 21(2):953–60. doi: 10.1016/j.concog.2012.02.013 PMID: 22425537

40. Reber TP, Luechinger R, Boesiger P, Henke K. Detecting analogies unconsciously. Front Behav Neurosci. 2014; 8:9. doi: 10.3389/fnbeh.2014.00009 PMID: 24478656

41. Grill-Spector K, Kushnir T, Hendler T, Malach R. The dynamics of object-selective activation correlate with recognition performance in humans. Nat Neurosci. 2000 Aug; 3(8):837–43. PMID: 10903579

42. Gaillard R, Dehaene S, Adam C, Clémenceau S, Hasboun D, Baulac M, et al. Converging Intracranial Markers of Conscious Access. PLoS Biol. 2009 Mar 17; 7(3):e1000061. doi: 10.1371/journal.pbio.1000061 PMID: 19296722

43. Dehaene S, Charles L, King J-R, Marti S. Toward a computational theory of conscious processing. Curr Opin Neurobiol. 2014 Apr; 25:76–84. doi: 10.1016/j.conb.2013.12.005 PMID: 24709604

44. Wiggs CL, Martin A. Properties and mechanisms of perceptual priming. Curr Opin Neurobiol. 1998 Apr; 8(2):227–33. PMID: 9635206

45. Grill-Spector K, Henson R, Martin A. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci. 2006 Jan; 10(1):14–23. PMID: 16321563

46. Tipper SP. The negative priming effect: inhibitory priming by ignored objects. Q J Exp Psychol A. 1985 Nov; 37(4):571–90. PMID: 4081101

47. Eichenbaum H. Hippocampus: Cognitive Processes and Neural Representations that Underlie Declarative Memory. Neuron. 2004 Sep 30; 44(1):109–20. PMID: 15450164

Page 18

48. Ciranni MA, Shimamura AP. Retrieval-induced forgetting in episodic memory. J Exp Psychol Learn Mem Cogn. 1999; 25(6):1403. PMID: 10605828

49. McKenzie S, Frank AJ, Kinsky NR, Porter B, Rivière PD, Eichenbaum H. Hippocampal Representation of Related and Opposing Memories Develop within Distinct, Hierarchically Organized Neural Schemas. Neuron. 2014 Jul 2; 83(1):202–15. doi: 10.1016/j.neuron.2014.05.019 PMID: 24910078

50. McKenzie S, Robinson NTM, Herrera L, Churchill JC, Eichenbaum H. Learning Causes Reorganization of Neuronal Firing Patterns to Represent Related Experiences within a Hippocampal Schema. J Neurosci. 2013 Jun 19; 33(25):10243–56. doi: 10.1523/JNEUROSCI.0879-13.2013 PMID: 23785140

51. Wagner AD, Maril A, Schacter DL. Interactions Between Forms of Memory: When Priming Hinders New Episodic Learning. J Cogn Neurosci. 2000 Nov 1; 12(supplement 2):52–60. PMID: 11506647

52. Poldrack RA, Clark J, Paré-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, et al. Interactive memory systems in the human brain. Nature. 2001 Nov 29; 414(6863):546–50. PMID: 11734855

Subliminal messages exert long-term effects on decision-making

Simon Ruch*, Marc Alain Züst, and Katharina Henke

Department of Psychology, University of Bern, Switzerland and Center for Cognition, Learning and Memory, University of Bern, Switzerland

*Correspondence address. Department of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland. Tel: +41 (0)31 631 40 11; E-mail: simon.ruch@psy.unibe.ch

Abstract

Subliminal manipulation is often considered harmless because its effects typically decay within a second. So far, subliminal long-term effects on behavior were only observed in studies which repeatedly presented highly familiar information such as single words. These studies suggest that subliminal messages are only slowly stored and might not be stored at all if they provide novel, unfamiliar information. We speculated that subliminal messages might affect delayed decision-making especially if messages contain several pieces of novel information that must be relationally bound in long-term memory. Relational binding engages the hippocampal memory system, which can rapidly encode and durably store novel relations. Here, we hypothesized that subliminally presented stimulus pairs would be relationally processed influencing the direction of delayed conscious decisions. In experiment 1, subliminal face–occupation pairs affected conscious decisions about the income of these individuals almost half an hour later. In experiment 2, subliminal presentation of vocabulary of a foreign language enabled participants to later decide whether these foreign words are presented with correct or incorrect translations. Subliminal influence did not significantly decay if probed after 25 versus 15 min. This is unprecedented evidence of the longevity and impact of subliminal messages on conscious, rational decision-making.

Key words: subliminal; decision-making; unconscious processing; long-term memory; hippocampus; information integration

Introduction

Subliminal messages exert diverse influences on our thoughts and our behavior (van Gaal *et al.*, 2012; Hassin, 2013). Subliminal stimuli can facilitate conscious processing of related information (Van den Bussche *et al.*, 2009), change our current mood (Monahan *et al.*, 2000), boost our motivation (Aarts *et al.*, 2008), and can even alter our political attitudes and voting intentions (Hassin *et al.*, 2007; Weinberger and Westen, 2008). With such a broad impact, subliminally planted information might have the potential to alter our decisions in everyday situations such as voting.

In order to influence decision-making in real-life situations, subliminal messages must be stored for long-term after only a few exposures, e.g. after a single confrontation with a subliminal TV advert. Furthermore, messages must be stored even if

they contain complex relational information that requires semantic integration, such as "politician X will lower the taxes." For subliminal manipulation to be effective, humans thus have to be able to semantically integrate and rapidly store unconscious pieces of novel information into long-lasting associative memories that can be retrieved if relevant to the context of a later decision.

The processes which allow novel information to shape subsequent decisions are generally thought to depend on consciousness – be it the integration of novel information into abstract mental representations (e.g. Tononi, 2004), rapid encoding of these representations into long-term memory (e.g. Shanks, 2010), or the use of these representations to make informed decisions (e.g. Newell and Shanks, 2014). However,

1

growing evidence indicates that the human unconscious can perform various high-level cognitive functions (van Gaal *et al.*, 2012; Hassin, 2013) that might allow decision processes to benefit from subliminal messages.

Several studies reported that subliminally planted information can be semantically integrated outside conscious awareness (for a detailed review, see Mudrik *et al.*, 2014). Indeed, humans can detect incompatible object-background configurations in subliminal scenes (Mudrik and Koch, 2013), solve subliminally presented arithmetic problems (García-Orza *et al.*, 2009; Van Opstal et al., 2011; Sklar *et al.*, 2012; Karpinski *et al.*, 2016), and draw inferences from subliminal picture sequences and word pairs (Kawakami and Yoshida, 2015 and Reber and Henke, 2012, respectively).

Subliminal stimulation was further found to nonconsciously shape decision-making – at least if masked stimuli consisted of single familiar items that required little integration. For example, priming studies reported that subliminal primes not only facilitate correct responses to related targets in a classification task, but also bias responses in "free choice" tasks in which participants can freely decide between response alternatives (Schlaghecken and Eimer, 2004; Klapp and Haas, 2005; Kiesel et al., 2006; Parkinson and Haggard, 2014; Ocampo, 2015). Similarly, studies on subliminal persuasion suggested that repeated subliminal exposure to brand names (e.g. "Lipton Ice Tea") or goal-relevant words (e.g. "thirst" in thirsty participants) can bias participants' product choices (Karremans *et al.*, 2006; Bermeitinger *et al.*, 2009; Verwijmeren *et al.*, 2011, 2013) or reinforce a certain behavior (e.g. to drink, see Strahan *et al.*, 2002). Thus, there is ample evidence that subliminal messages can be integrated unconsciously and can influence decisions and choices.

Whether subliminally presented information is stored in long-term memory to guide delayed decisions is vastly unknown. So far, studies on information integration and decision-making only assessed immediate influences of subliminal stimulation. Priming studies which assessed the longevity of subliminal influences usually reported that behavioral effects of masked primes decay within 1 s (e.g. Forster *et al.*, 1990; Ferrand, 1996; Greenwald *et al.*, 1996). This suggests that subliminal information is not stored and thus cannot affect delayed decisions. Yet, some studies reported that longer lasting subliminal influences on behavior are possible under certain conditions. In many of these studies, the same subliminal messages were presented multiple times (Lowery et al., 2007; Capa et al., 2011; Levy et al., 2014; Farooqui and Manly, 2015). Furthermore, participants were often informed about the presence of subliminal stimuli and were provided conscious feedback or rewards after each subliminal message (Aarts *et al.*, 2008; Capa *et al.*, 2011; Farooqui and Manly, 2015) or were asked to consciously detect or classify each subliminal event (Gaillard et al., 2007; Chen *et al.*, 2009; Chong *et al.*, 2014). These studies thus suggest that subliminal messages are only slowly stored and are only retained if subjects have the explicit intention to process the hidden events. Importantly, most studies used familiar information such as single words as subliminal stimuli to prime subsequent conscious processing of this information (Gaillard et al., 2007; Chen *et al.*, 2009; Chong *et al.*, 2014) or to prime a specific goal (e.g. to perform well on a test), intention, or stereotype (Lowery et al., 2007; Capa *et al.*, 2011; Levy *et al.*, 2014). This suggests that subliminal long-term effects are achieved only if familiar information is presented but not if novel relational information has to be learned.

We asked if humans can rapidly integrate and store novel relational information (e.g. "person X is a manager", see Fig. 1a) from subliminal messages for later use in a decision-making situation (e.g. "guess the income of X"). We speculate that

subliminal messages should be stored especially if they consist of multiple items that require relational processing. Relational binding calls upon the hippocampal memory system, which can rapidly store novel relations for long term (Henke *et al.*, 1997; Holdstock *et al.*, 2002; Harand *et al.*, 2012). Traditional views hold that hippocampus is only involved in the encoding and retrieval of consciously perceived information (Moscovitch, 1995; Squire and Zola, 1996; Tulving, 2002), and that associative learning outside conscious awareness is unlikely (Shanks, 2010). However, growing evidence suggests that hippocampus operates independently of consciousness and that nonconscious relational learning is humanly feasible (for reviews, see e.g. Reder *et al.*, 2009; Henke, 2010; Dew and Cabeza, 2011; Hannula and Greene, 2012; Olsen *et al.*, 2012). Indeed, hippocampus was found to mediate implicit learning (Chun and Phelps, 1999; Greene *et al.*, 2006; Negash *et al.*, 2015) and retrieval (Greene *et al.*, 2007; Hannula and Ranganath, 2009; Addante, 2015; Reber *et al.*, 2016) of relational information between visible stimuli. Most importantly, hippocampus was also found to be involved in the encoding and retrieval of subliminally presented stimulus pairs (Henke *et al.*, 2003; Reber *et al.*, 2012; Duss *et al.*, 2014; Züst *et al.*, 2015).

We ran two experiments to test whether subliminal stimulus pairs would affect delayed decision-making (Figure 1a). In experiment 1, we tested whether subliminal presentations of face–occupation pairs would guide later conscious decisions about the income of the same faces. We assessed the longevity of subliminal influence by measuring its decay across delays of 15–25 min. In experiment 2, we assessed if new vocabulary of a foreign language is acquired subliminally affecting later lexical-semantic decisions on the same foreign words. Both experiments were composed of an encoding phase, an encoding-test interval that was filled with rest, and a test phase (Fig. 1a). The decision task given in the test phase constituted an indirect (implicit) memory test. Importantly, participants were kept naïve regarding subliminal stimulation until the end of the experiment. This allowed us to assess subliminal influences in the absence of any explicit intention to process the hidden messages.

Method

General procedure

Unless indicated otherwise, experimental procedures were identical for both experiments. We used a well-established paradigm to render stimuli invisible (Duss *et al.*, 2011). Each stimulus was presented 12 times for 17 ms during one unconscious encoding episode of 6 s duration (see Supplementary Fig. 1). Stimuli were preceded and followed by visual noise masks (sandwich masking). In experiment 1, 40 face–word pairs were encoded in 40 unconscious encoding episodes. In experiment 2, each unconscious encoding episode contained a pseudoword–word pair that was presented twice in nonadjacent repetitions, yielding a total of 48 encoding episodes for 24 pseudoword–word pairs. Stimuli and masks were embedded in an attention task which allowed us to direct participants' focal attention to the screen without disclosing the presence of subliminal messages (see Supplementary Fig. 1). The task required participants to respond to target screens that appeared at random times once in every encoding episode (see Supplementary Fig. 1). Mean hit rate to targets was >85% in both experiments. Following the attention task, participants rested for 15, 20, or 25 min (depending on experiment and condition) before performing the decision tasks. The attention task and the decision
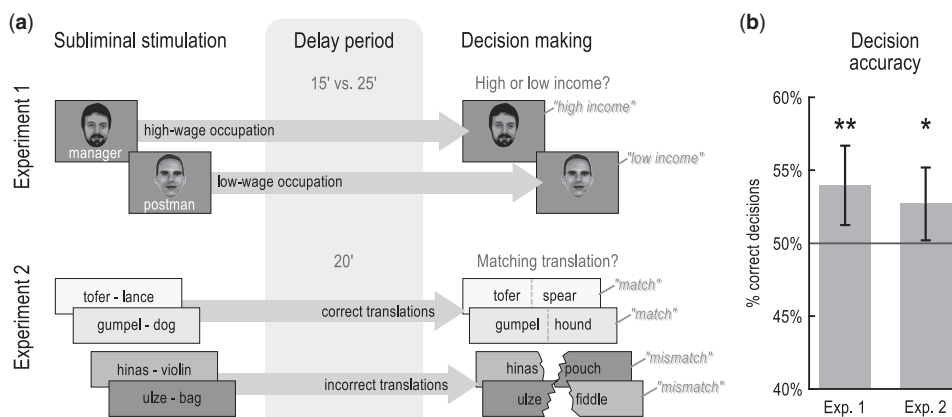
Figure 1. (a) Design of experiments 1 and 2. Subliminal presentation of face–occupation pairs (Experiment 1) or pairs consisting of pseudowords and their translations (Experiment 2) altered participants' decisions when later asked to rate the income of faces or to identify correct (matching) translations of pseudowords. (b) Results. Two-tailed *t*-tests indicated that decision accuracy was significantly above the 50% chance level (mean accuracy with 95% confidence interval). *P < 0.05, **P < 0.01.

task were briefly practiced at the beginning of the experiment using a few stimuli that were not used later on.

Following the main experiments, a structured interview was performed to assess subjective awareness for subliminal information. Using funneled questions, we first asked participants whether they had seen any unexpected stimuli during the attention task or had suspected the presence of any hidden information, and then asked more specifically whether they had seen hidden faces or words. We then informed participants of the subliminal stimulation and administered forced-choice tests to assess awareness of subliminal stimuli objectively. Masking paradigm and attention task in these awareness tests were the same as in the main experiments. However, awareness was tested trial-by-trial: following each encoding episode of a new stimulus pair, the attention task was interrupted and participants were asked to provide information about the subliminal stimuli. We expected participants to perform at chance level in these tests if masking rendered stimuli truly subliminal.

Experiments were approved by the local ethics committee. Written semi-informed consent was obtained from all participants before experimentation. The information about subliminal stimulation was provided following the experiment.

Participants
We recruited participants with normal or corrected-to-normal visual acuity. In experiment 1, two participants were excluded *post hoc* due to errors in data logging. Of the remaining 46 participants (19 men and 27 women, mean age = 36.0), 23 were tested with a 15 min delay between subliminal encoding and decision-making, and 23 with a 25 min delay. In experiment 2, two participants were excluded *post hoc* because they had failed to report their decisions within the response time window. The remaining sample consisted of 23 women and 11 men (mean age = 22.8).

Experiment 1
During the encoding phase, we presented 10 faces combined with written high-wage occupations and 10 faces with low-wage occupations (Fig. 1a). Twenty additional faces were paired with consonant strings and presented subliminally in a control condition (not reported). Each stimulus pair was presented in just one subliminal encoding episode, which comprised 12 stimulus repetitions within 6 s. The encoding-test interval spanned

15 or 25 min. In the test phase, the former subliminal faces were re-presented for 5 s for conscious inspection with the instruction to decide whether an individual would earn a high or low income.

Objective awareness tests
We administered two forced-choice tests to assess awareness of subliminal faces and occupations objectively. In each test, we presented 40 novel face–occupation pairs subliminally using the same attention task as in the main experiment. Immediately following the encoding of a novel pair in a subliminal encoding episode, the attention task was interrupted and participants were interrogated regarding the just presented face or occupation. In the face-awareness test, the just subliminally presented face and a novel face were presented side by side for conscious inspection. Participants were asked to select the face that they thought was just flashed subliminally. In the occupation-awareness test, the text "Income?" was presented which signaled participants to name the income ("high" vs. "low") that they thought would the just subliminally flashed occupation yield. After giving their response, the attention task was continued and a new face–occupation pair was presented subliminally. Participants took either 40 face-awareness trials followed by 40 occupation awareness trials or vice versa.

Material
We used 160 grayscale images of male faces that were given average income ratings in a pilot study. Images were equalized regarding luminance and contrast and were assigned to 16 lists of 10 faces each. Lists were comparable with respect to income, age, facial hair, and emotional facial expression. Lists were counterbalanced over experimental conditions such that each face was presented an equal number of times with a high- and a low-wage occupation and with a consonant string. All faces were also rotated into the two awareness tests. Stimuli were presented an equal number of times in the main experiment, and the awareness tests of faces and occupations. Hence, results in the experiment and the awareness test derived from the same stimulus material (presented to different participants).

We used 10 typical high- and 10 low-wage occupation words that were similar regarding mean logarithmic word frequency

(Leipzig Corpora Collection, http://corpora.uni-leipzig.de/) and character count.

Faces were displayed at the center of the screen in front of a dark gray background. Words were presented below the faces in light gray in a *sans-serif* font. Stimuli were delivered at a visual angle of 15° using a Digital Light Processing (DLP) projector with a 60 Hz refresh rate.

Experiment 2

Participants encoded 24 combinations of written pseudowords (fictitious foreign language) and German words (fictitious translations). We presented these pseudoword–word combinations subliminally for unconscious semantic relational encoding using the same paradigm as in experiment 1. There were two nonadjacent subliminal encoding episodes (each comprising 12 stimulus repetitions within 6 s) per word pair within a randomized sequence of subliminal encoding episodes. The encoding-test interval spanned 20 min. At test, we re-presented the same foreign words for 5 s for conscious inspection. Each foreign word was shown besides a German word that was a synonym to the subliminal German translation of either this or another foreign word. Synonyms were used to test for semantic rather than perceptual relational retrieval. Half of foreign words presented at test were recombined to break the encoded semantic relation (incorrect translations) and half were combined to keep the semantic relation from encoding to test (correct translation). Participants were instructed to decide whether or not a presented foreign word and the German translation word fit together (match/mismatch decision).

Objective awareness test

We administered one forced-choice test to assess awareness of subliminal words objectively. Participants encoded 24 (novel) pseudoword–German word pairs. Following each subliminal encoding episode, a probe word was displayed for conscious inspection for participants to decide whether or not the probe was a synonym to the just presented subliminal German word. In half the trials, the probe was a valid synonym to the subliminally presented German word, in the other half the probe was an unrelated foil. Hence, we assessed the semantic processing of subliminal words, as in the main experiment.

Material

We created 48 word triplets consisting of a pseudoword, its German translation used for subliminal encoding, and a synonym to the German translation used for the decision task (see Fig. 1). German synonym pairs were gathered from Open Thesaurus (http://www.openthesaurus.de/). Two-syllabic pronounceable pseudowords were created using German and Dutch syllables provided by the Celex database (http://celex.mpi.nl/). The 48 encoding stimuli were assigned to four lists of 12 items each with equal distributions of word lengths, pronounceability, concreteness, animatedness, and logarithmic frequency of appearance (drawn from Leipzig Corpora Collection, http://corpora.uni-leipzig.de/). Assignment of German translations to pseudowords was randomized anew for each list and each participant to reduce potential bias resulting from particular combinations. Two lists were used for the main experiment and the other two lists for the awareness test; this assignment was counterbalanced over participants. Twelve further German words were used as foils in the awareness test. Pseudowords were displayed in the left and German words in the right visual field of participants. Words were presented for subliminal encoding

using the same masking paradigm and psychophysical conditions as applied in experiment 1.

Results

Experiment 1

In experiment 1, we tested whether subliminal presentations of face–occupation pairs would influence later classification of the same faces. We hypothesized that participants would rate individuals who had earlier been flashed with a high-wage or low-wage occupation as high-income or low-income earners, respectively. A two-tailed *t*-test on the mean decision accuracy confirmed this hypothesis. Mean decision accuracy was 53.97% (95% CI [51.25, 56.96]), which exceeded the chance level of 50% ($t(45) = 2.94$, $P = 0.005$; $r = 0.40$) (Fig. 1b). Mean decision accuracy was not significantly smaller after 25 compared to 15 min (53.04% vs. 54.90%; two-tailed $t(44) = 0.68$, $P = 0.50$; $r = 0.10$), which indicates that the effect of subliminal encoding persisted through both retention intervals at equal strength. Hence, new semantic associations were stored for long term affecting the direction of decision-making almost half an hour later.

An interview that was administered after the main experiment to assess subjective awareness of the masked stimuli suggested that none of the participants had seen the subliminal faces or words during encoding. Two objective awareness tests further indicated that masked stimuli could not be consciously perceived. Participants performed at chance level of 50% if asked which one of two faces had just been presented subliminally (mean recognition accuracy = 48.81%, 95% CI [46.36, 51.27], missing data of one participant due to data loss) or if a high-or low-wage occupation had been presented (mean recognition accuracy = 49.06%, 95% CI [46.90, 51.23]). The *t*-tests against chance level were not significant (awareness for faces: $t(43) = -0.98$, $P = 0.34$, $r = 0.15$; awareness for occupations: $t(44) = -0.90$, $P = 0.37$, $r = 0.13$), suggesting that participants were consciously unaware of the subliminal stimuli. However, the nonsignificant tests do not reveal whether participants were truly unaware (i.e. whether H0 can be accepted), or whether our tests were not sensitive enough to detect awareness. We, therefore, calculated Bayes factors for the two *t*-tests following the recommendations provided in Dienes (2008, 2014). Bayes factors (BF) indicate the relative strength of two hypotheses H1 and H0 with a BF(H1/H0) = 1 suggesting that the data are inconclusive and favor neither hypothesis, and BFs > 3 or < 1/3 suggesting substantial evidence for H1 or H0, respectively. Assuming that awareness of subliminal stimuli should yield recognition accuracy that is comparable to performance in the decision task (3.97% above chance level), we chose a half-normal prior distribution with a mode of 0 and a standard deviation of 3.97% to calculate BF (Dienes, 2014). The resulting BF for the face-awareness test (mean accuracy = 1.19% below chance level, SE = 1.21%) and the occupation-awareness test (mean accuracy = 0.94% below chance, SE = 1.04%) were 0.16 and 0.14, respectively. Both factors were below 1/3, indicating substantial evidence for the null hypothesis, i.e. that participants were not aware of the masked stimuli. We further performed a multiple regression analysis to assess whether potential awareness of faces or occupations predicted decision accuracy in the main experiment, and to estimate whether accuracy remained above chance level if the theoretical awareness in both tests is 0 (i.e., whether the intercept is significant). This method was brought forward by Greenwald et al. (1995). Although subject to criticism (e.g. Miller, 2000), this method is used widely to assess whether

subliminal influences on behavior are independent of stimulus awareness. Neither potential awareness of faces ($\beta = 0.16$, $t(41) = 0.88$, $P = 0.383$) nor potential awareness of occupations ($\beta = 0.24$, $t(41) = 1.2$, $P = 0.23$) predicted decision accuracy in the main experiment. Decision accuracy in the main experiment remained above chance level when we controlled for selection accuracy in the two awareness tests (intercept of the regression with two predictors; $t(41) = 2.93$, $P < 0.01$, $r = 0.42$). In a final step, we excluded those participants from the analysis of the data from the main experiment, who tended to perform above or below chance level in either of the two awareness tests (12 participants exhibited an awareness score with a binomial probability of $P < 20\%$). Classification accuracy in the main experiment remained above chance level (54.12%, 95% CI [50.72, 57.51], $t(33) = 2.468$, $P = 0.019$, $r = 0.39$) with these 12 participants removed. We, therefore, conclude that there was no awareness of encoding stimuli in the main experiment and that long-term effects derived from unconscious processing alone.

Experiment 2

In experiment 2, we tested for long-term effects using a subliminal vocabulary acquisition task. We hypothesized that new vocabulary would be encoded subliminally influencing delayed lexical–semantic decisions on the same foreign words when they were presented visibly. At test, participants were instructed to decide whether or not a given foreign word and a German translation word fit together (match/mismatch decision). On average, 52.7% (95% CI [50.1, 55.2]) (Fig. 1b) of synonyms were classified accurately, which exceeded chance level (two-tailed t-test: $t(33) = 2.128$; $P = 0.041$; $r = 0.35$). Hence, the meaning of subliminal German words was decoded, linked to foreign words, and stored to influence lexical–semantic decisions 20 min following subliminal encoding.

An interview administered following the main experiment assessed the subjective awareness of the masked stimuli. Participants' responses indicated that none had consciously perceived subliminal words or letters. An objective awareness was used to corroborate the subjective reports. When asked whether a consciously displayed German word represents a synonym to a just subliminally flashed German word, participants decided correctly in 51.3% of cases (95% CI [48.9, 53.8]). This performance did not significantly exceed chance level of 50% ($t(33) = 1.121$, two-tailed $P = 0.27$; $r = 0.19$). To validate whether performance was truly at chance level, we again calculated the Bayes factor for the t-test as suggested by Dienes (2008, 2014). Assuming that awareness for subliminal words should yield a recognition accuracy comparable to performance in the decision task (2.7% above chance level), we chose a half-normal prior distribution with a mode of 0 and a standard deviation of 2.7% to estimate the BF (Dienes, 2014). The resulting factor for the awareness test (mean accuracy = 1.3% above chance level, SE = 1.16%) was 1.14, which suggests that the test was not sensitive enough to reject stimulus awareness. Whether participants were truly unaware of the subliminal words thus remains elusive. However, further analyses suggested that subliminal long-term influences on decision-making did not benefit from stimulus awareness. First of all, a regression analysis (Greenwald et al., 1995) indicated that accuracy in the awareness test did not predict decision accuracy in the main experiment ($\beta = 0.038$, $t(32) = 0.216$, $P = 0.83$), which was above chance level when selection accuracy on the awareness test was 0 (intercept of regression; $t(31) = 2.052$, $P = 0.045$, $r = 0.34$). Furthermore, even if those two participants, who

tended to perform above or below chance level in the awareness test (binomial probability of $P < 20\%$), were excluded from the analysis of the data from the main experiment, classification accuracy remained above chance level (52.6%, 95% CI [50.1, 55.1], $t(31) = 2.142$, $P = 0.040$, $r = 0.36$). Although the Bayes analysis yielded an inconclusive result regarding stimulus awareness in the awareness test, the regression analysis suggests that long-term effects on decision-making derived from unconscious processes.

Discussion

To summarize, subliminal verbal and nonverbal item pairs influenced participants' conscious, deliberate decisions almost half an hour following subliminal stimulation. In experiment 1, subliminal face–occupation pairs influenced conscious decisions on the income of the same individuals after delays of 15 and 25 min. In experiment 2, subliminal presentation of foreign language vocabulary influenced participants' decisions on correct/incorrect word translations after a delay of 20 min. These findings provide unprecedented evidence of a considerable longevity of subliminal effects on intentional behavior such as decision-making.

So far, most investigators who addressed the longevity of subliminal priming reported that subliminal effects on behavior would fade within 1 s (e.g. Greenwald et al., 1996), which suggests that subliminal messages leave no long-term memory traces and therefore cannot influence delayed decisions. The few studies which reported longer lasting effects used familiar stimuli that were repeatedly presented to participants who were informed about the presence or purpose of subliminal events (e.g. Chong et al., 2014; Levy et al., 2014). These studies suggest that subliminal information may be stored for long term if it is familiar, if many subliminal exposures are provided for a slow, incremental encoding process, and if subjects explicitly intent to process the subliminal information. Whether novel subliminal information can be integrated and stored following a single exposure to naïve subjects remains unclear. Even less clear is whether a subliminal one-time processing has sufficient strength to influence delayed decision-making. We found that participants, who were unaware of the presence and purpose of subliminal information, successfully processed subliminal face–word and nonword–word pairs and formed lasting unconscious semantic relational memories based on only one or two subliminal encoding episodes (each comprising 12 adjacent subliminal stimulus repetitions).

We speculate that the reported rapid encoding and long-term retention of subliminal information owes to the type of stimuli and the memory system these stimuli called upon. As mentioned above, conclusions regarding the longevity of subliminal priming rested on the use of familiar single-item stimuli such as words (e.g. Forster et al., 1990; Ferrand, 1996; Greenwald et al., 1996). It is known that the processing of single items engages neocortex (Henke et al., 1997; Duss et al., 2014), which forms long-term memory traces rather slowly over many learning trials (McClelland et al., 1995). This might explain why subliminal influences were found to be short-lived or to build up slowly. Here, we used multi-item displays that contained both novel (unfamiliar faces or foreign words) and familiar (occupation words or German translation words) pieces of information. Relational encoding of multiple stimuli engages hippocampus (Henke et al., 1997; Duss et al., 2014), which learns rapidly and stores relations for long term due to its exceptional plasticity (McClelland et al., 1995). Subliminal relational encoding and

retrieval have been associated reliably with hippocampal activity changes (Henke *et al.*, 2003; Reber *et al.*, 2012; Züst *et al.*, 2015). Importantly, hippocampal damage abolished both supraliminal (conscious) and subliminal (unconscious) relational binding but left subliminal single-item priming intact (Duss *et al.*, 2014). As encoding-test intervals spanned a few minutes or less in these previous neuroimaging experiments, they were uninformative regarding longer term effects of subliminal stimuli on behavior. Using similar stimuli and the same masking technique, we now demonstrate subliminal effects on decisions delayed by 15–25 min. Because subliminal influences did not noticeably decay from the delay of 15–25 min in experiment 1, we speculate that subliminally planted information might affect decision-making even at longer intervals.

The long-lasting influence of subliminal stimulus pairs evinces nonconscious relational learning, but the experimental paradigms employed in this study do not pin down the exact nature of the unconsciously formed memories. Although we assume that subliminal stimulus pairs yield semantically precise unconscious relational memories ("person X is a manager", "*gumpel* means *dog*") that are later reactivated to support decision-making, our experimental setup cannot rule out the possibility that participants had formed associations between faces/foreign words and broad semantic or affective categories (e.g. "person X as a manager must be wealthy", "I despise *gumpel* because I fear dogs"). Such fuzzy semantic/affective associations could suffice to guide subsequent decisions ("X has a high income because he is wealthy", "*hound* is a valid translation of *gumpel* because I despise both"). But we would like to point out that the results of earlier investigations on subliminal encoding and long-term memory formation demonstrated unequivocally that subliminal words are understood with high semantic precision. In Duss *et al.* (2011), subliminal presentations of face–occupation pairs influenced participants' subsequent conscious classifications of the same faces when various semantic dimensions were offered, namely regularity of income, length of education, and creativity of work. Faces encoded with an artistic instead of an academic occupation (e.g. "actor" vs. "lawyer") were later classified as generating an irregular income, coming from a shorter school education, and performing creative work. These distinct influences of subliminal information on diverse semantic classifications suggest a precise rather than diffuse lexical–semantic word analysis. Further evidence for precise subliminal encoding is provided by our studies on unconscious relational inference (Reber and Henke, 2012; Reber *et al.*, 2012; Henke *et al.*, 2013). In these studies, overlapping subliminal word pairs such as "winter-red" and "red-computer" (A–B, B–C) were presented apart in time and were nevertheless integrated semantically to influence delayed judgments regarding the semantic relatedness of A and C, like "winter-computer". The successful relational integration in this subliminal paradigm cannot be explained by affective or fuzzy semantic word priming.

Doubts have been raised recently as to whether experimenters had adequately assessed stimulus awareness and correctly estimated influences of subliminal stimuli on behavior (Newell and Shanks, 2014; Hesselmann and Moors, 2015). To avert such suspicion, we assessed stimulus awareness following the main experiments using objective awareness tests that had the same statistical power as our decision tasks. The objective awareness tests confirmed that the subliminal stimulation paradigm rendered stimuli largely (experiment 2) or completely (experiment 1) imperceptible to the conscious mind, and that the observed long-term effects of subliminal stimulation

were independent of stimulus awareness. Because we did not assess awareness of each subliminal stimulus immediately following the respective encoding episode in the main experiments, we cannot rule out the possibility that participants were briefly aware of some of the subliminal images. However, none of the participants reported to have noticed the presence of hidden or masked information during the main experiments or during the objective awareness tests. Hence, all available data suggest that the reported long-term influences of subliminal stimuli were independent of consciousness.

The finding that subliminally processed information is rapidly integrated and stored to guide delayed decisions challenges prevailing views of the cognitive function of consciousness. Consciousness is usually considered a precondition for successful information integration (Tononi, 2004; Mudrik *et al.*, 2014), relational learning (Shanks, 2010), and decision-making (Bettman *et al.*, 1998; Simonson, 2005; Newell and Shanks, 2014). However, evidence is accumulating that these notions of consciousness need revision (Dijksterhuis and Nordgren, 2006; Nordgren, 2006; Reder et al., 2009; Henke, 2010; Dew and Cabeza, 2011; Olsen *et al.*, 2012; Hannula and Greene, 2012; van Gaal *et al.*, 2012; Hassin, 2013). Indeed, unconscious integration of different semantic concepts and of temporally or spatially distributed percepts was reported not only for stimuli that were rendered subliminal using visual masking but also for stimuli made invisible using continuous flash suppression (Mudrik *et al.*, 2011; Sklar *et al.*, 2012; Vlassova *et al.*, 2014; Bergström and Eriksson, 2015; Karpinski *et al.*, 2016) or visual crowding (Atas *et al.*, 2013). Even during the unconsciousness of deep sleep, words, sounds, and odors were found to be integrated (Ruby et al., 2008; Daltrozzo, *et al.*, 2012) and stored in long-term memory (Arzi *et al.*, 2012, 2014; Ruch *et al.*, 2014) to modulate behavior following waking. More evidence for the feasibility of unconscious relational integration is provided by social psychology: decisions, which require the consideration, weighing and integration of large amounts of (supraliminally provided) information, were better following unconscious deliberation than conscious reasoning (e.g. Dijksterhuis *et al.*, 2006; Mealor and Dienes, 2012; Abadie *et al.*, 2013). This 'deliberation without attention' effect suggests that our conscious decisions are vitally influenced by nonconscious processes (but see e.g. Newell and Shanks, 2014; Nieuwenstein *et al.*, 2015; and Vadillo *et al.*, 2015 for critical reviews).

In sum, our findings add to a growing body of evidence suggesting that subliminal messages can be used to change our thoughts, attitudes, emotions, and actions (van Gaal *et al.*, 2012; Hassin, 2013). Subliminal stimulation was shown to influence current moods (Monahan *et al.*, 2000), political attitudes (Hassin *et al.*, 2007; Weinberger and Westen, 2008), intentions (Hassin *et al.*, 2007), choices and decisions (Bermeitinger *et al.*, 2009), and cognitive strategies (Lau and Passingham, 2007; Reuss *et al.*, 2011). However, little is known about the longevity of these subliminal influences because these studies were focused mainly on immediate effects of subliminal stimulation. Here, we demonstrate that a few exposures to novel subliminal information is sufficient to influence delayed decision-making. The surprising impact of subliminal messages on rational, intentional, conscious behavior lends subliminal protocols to practical applications, of which advertising is just one example.

Supplementary data

Supplementary data is available at Neuroscience of Consciousness Journal online.

Acknowledgments

## References

Aarts H, Custers R, Marien H. Preparing and motivating behavior outside of awareness. *Science* 2008;**319**:1639. http://doi.org/10.1126/science.1150432.

Abadie M, Waroquier L, Terrier P. Gist memory in the unconscious-thought effect. *Psychol Sci* 2013;**24**:1253–9. http://doi.org/10.1177/0956797612470958.

Addante RJ. A critical role of the human hippocampus in an electrophysiological measure of implicit memory. *NeuroImage* 2015;**109**:515–28. http://doi.org/10.1016/j.neuroimage.2014.12.069.

Arzi A, Holtzman Y, Samnon P, *et al.* Olfactory aversive conditioning during sleep reduces cigarette-smoking behavior. *J Neurosci* 2014;**34**:15382–93. http://doi.org/10.1523/JNEUROSCI.2291-14.2014.

Arzi A, Shedlesky L, Ben-Shaul M, *et al.* Humans can learn new information during sleep. *Nat Neurosci* 2012;**15**:1460–5. http://doi.org/10.1038/nn.3193

Atas A, Faivre N, Timmermans B, *et al.* Nonconscious learning from crowded sequences. *Psychol Sci* 2013. http://doi.org/10.1177/0956797613499591.

Bergström F, Eriksson J. The conjunction of non-consciously perceived object identity and spatial position can be retained during a visual short-term memory task. *Conscious Res* 2015;1470. http://doi.org/10.3389/fpsyg.2015.01470

Bermeitinger C, Goelz R, Johr N, *et al.* The hidden persuaders break into the tired brain. *J Exp Soc Psychol* 2009;**45**:320–6. http://doi.org/10.1016/j.jesp.2008.10.001

Bettman JR, Luce MF, Payne JW. Constructive consumer choice processes. *J Cons Res* 1998;**25**:187–217. http://doi.org/10.1086/209535

Capa RL, Cleeremans A, Bustin GM, *et al.* Long-lasting effect of subliminal processes on cardiovascular responses and performance. *Int J Psychophysiol* 2011;**81**:22–30. http://doi.org/10.1016/j.ijpsycho.2011.04.001

Chen JCW, Li W, Lui M, *et al.* Left-frontal brain potentials index conceptual implicit memory for words initially viewed subliminally. *Brain Res* 2009;**1285**:135–47. http://doi.org/10.1016/j.brainres.2009.05.085

Chong TTJ, Husain M, Rosenthal CR. Recognizing the unconscious. *Curr Biol* 2014;**24**:R1033–5. http://doi.org/10.1016/j.cub.2014.09.035

Chun MM, Phelps EA. Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nat Neurosci* 1999;**2**:844–7. http://doi.org/10.1038/12222

Daltrozzo J, Claude L, Tillmann B, *et al.* Working memory is partially preserved during sleep. *PLoS ONE* 2012;**7**:e50997. http://doi.org/10.1371/journal.pone.0050997

Dew ITZ, Cabeza R. The porous boundaries between explicit and implicit memory: behavioral and neural evidence. *Ann NY Acad Sci* 2011;**1224**:174–90. http://doi.org/10.1111/j.1749-6632.2010.05946.x

Dienes Z. *Understanding Psychology as a Science: An Introduction to Scientific and Statistical Inference*. New York: Palgrave Macmillan, 2008.

Dienes Z. Using Bayes to get the most out of non-significant results. *Quant Psychol Meas* 2014;**5**:781. http://doi.org/10.3389/fpsyg.2014.00781

Dijksterhuis A, Bos MW, Nordgren LF, *et al.* On making the right choice: the deliberation-without-attention effect. *Science* 2006;**311**:1005–7. http://doi.org/10.1126/science.1121629.

Dijksterhuis A, Nordgren LF. A theory of unconscious thought. *Perspect Psychol Sci* 2006;**1**:95–109. http://doi.org/10.1111/j.1745-6916.2006.00007.x.

Duss SB, Oggier S, Reber TP, *et al.* Formation of semantic associations between subliminally presented face-word pairs. *Conscious Cogn* 2011;**20**:928–35. http://doi.org/10.1016/j.concog.2011.03.018.

Duss SB, Reber TP, Hänggi J, *et al.* Unconscious relational encoding depends on hippocampus. *Brain* 2014;**137**:3355–70. http://doi.org/10.1093/brain/awu270.

Farooqui AA, Manly T. Anticipatory control through associative learning of subliminal relations: invisible may be better than visible. *Psychol Sci* 2015;**26**:325–34. http://doi.org/10.1177/0956797614564191.

Ferrand L. The masked repetition priming effect dissipates when increasing the inter-stimulus interval: evidence from word naming. *Acta Psychologica* 1996;**91**:15–25. http://doi.org/10.1016/0001-6918(95)00010-0.

Forster K, Booker J, Schacter DL, *et al.* Masked repetition priming: lexical activation or novel memory trace? *Bull Psychon Soc* 1990;**28**:341–5. http://doi.org/10.3758/BF03334039.

Gaillard R, Cohen L, Adam C, *et al.* Subliminal words durably affect neuronal activity. *NeuroReport* 2007;**18**:1527–31. http://doi.org/10.1097/WNR.0b013e3282f0b6cd.

García-Orza J, Damas-López J, Matas A, *et al.* "2 x 3" primes naming "6": evidence from masked priming. *Atten Percept Psychophys* 2009;**71**:471–80. http://doi.org/10.3758/APP.71.3.471

Greene AJ, Gross WL, Elsinger CL, *et al.* An fMRI analysis of the human hippocampus: inference, context, and task awareness. *J Cogn Neurosci* 2006;**18**:1156–73. http://doi.org/10.1162/jocn.2006.18.7.1156

Greene AJ, Gross WL, Elsinger CL, *et al.* Hippocampal differentiation without recognition: an fMRI analysis of the contextual cueing task. *Learn Mem* 2007;**14**:548–53. http://doi.org/10.1101/lm.609807.

Greenwald AG, Draine SC, Abrams RL. Three cognitive markers of unconscious semantic activation. *Science* 1996;**273**:1699–702. http://doi.org/10.1126/science.273.5282.1699

Greenwald AG, Klinger MR, Schuh ES. Activation by marginally perceptible ("subliminal") stimuli: dissociation of unconscious from conscious cognition. *J Exp Psychol-Gen* 1995;**124**:22–42.

Hannula DE, Greene AJ. The hippocampus reevaluated in unconscious learning and memory: at a tipping point? *Front Hum Neurosci* 2012;**6**. http://doi.org/10.3389/fnhum.2012.00080.

Hannula DE, Ranganath C. The eyes have it: hippocampal activity predicts expression of memory in eye movements. *Neuron* 2009;**63**:592–9. http://doi.org/10.1016/j.neuron.2009.08.025.

Harand C, Bertran F, La Joie R, *et al.* The hippocampus remains activated over the long term for the retrieval of truly episodic memories. *PLoS ONE* 2012;**7**:e43495. http://doi.org/10.1371/journal.pone.0043495.

Hassin RR. Yes it can: on the functional abilities of the human unconscious. *Perspect Psychol Sci* 2013;**8**:195–207. http://doi.org/10.1177/1745691612460684.

Hassin RR, Ferguson MJ, Shidlovski D, *et al.* Subliminal exposure to national flags affects political thought and behavior. *Proc Natl Acad Sci USA* 2007;**104**:19757–61. http://doi.org/10.1073/pnas.0704679104.

Henke K. A model for memory systems based on processing modes rather than consciousness. *Nat Rev Neurosci* 2010;**11**:523–32. http://doi.org/10.1038/nrn2850.

Henke K, Buck A, Weber B, *et al*. Human hippocampus establishes associations in memory. *Hippocampus* 1997;**7**:249–56. http://doi.org/10.1002/(SICI)1098-1063(1997)7:3<249::AID-HIPO1>3.0.CO;2-G.

Henke K, Mondadori CRA, Treyer V, *et al*. Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. *Neuropsychologia* 2003;**41**:863–76. http://doi.org/10.1016/S0028-3932(03)00035-6.

Henke K, Reber TP, Duss SB. Integrating events across levels of consciousness. *Front Behav Neurosci* 2013;**7**:68. http://doi.org/10.3389/fnbeh.2013.00068.

Hesselmann G, Moors P. Definitely maybe: can unconscious processes perform the same functions as conscious processes? *Conscious Res* 2015;**584**. http://doi.org/10.3389/fpsyg.2015.00584.

Holdstock JS, Mayes AR, Isaac CL, *et al*. Differential involvement of the hippocampus and temporal lobe cortices in rapid and slow learning of new semantic information. *Neuropsychologia* 2002;**40**:748–68. http://doi.org/10.1016/S0028-3932(01)00192-0.

Karpinski A, Yale M, Briggs JC. Unconscious arithmetic processing: a direct replication. *Eur J Soc Psychol* 2016; n/a-n/a. http://doi.org/10.1002/ejsp.2175.

Karremans JC, Stroebe W, Claus J. Beyond Vicary's fantasies: the impact of subliminal priming and brand choice. *J Exp Soc Psychol* 2006;**42**:792–8. http://doi.org/10.1016/j.jesp.2005.12.002.

Kawakami N, Yoshida F. Perceiving a story outside of conscious awareness: when we infer narrative attributes from subliminal sequential stimuli. *Conscious Cogn* 2015;**33**:53–66. http://doi.org/10.1016/j.concog.2014.12.001.

Kiesel A, Wagener A, Kunde W, *et al*. Unconscious manipulation of free choice in humans. *Conscious Cogn* 2006;**15**:397–408. http://doi.org/10.1016/j.concog.2005.10.002.

Klapp ST, Haas BW. Nonconscious influence of masked stimuli on response selection is limited to concrete stimulus-response associations. *J Exp Psychol-Hum Percept Perform* 2005;**31**:193–209. http://doi.org/10.1037/0096-1523.31.1.193.

Lau HC, Passingham RE. Unconscious activation of the cognitive control system in the human prefrontal cortex. *J Neurosci* 2007;**27**:5805–11. http://doi.org/10.1523/JNEUROSCI.4335-06.2007.

Levy BR, Pilver C, Chung PH, *et al*. Subliminal strengthening: improving older individuals' physical function over time with an implicit-age-stereotype intervention. *Psychol Sci* 2014;**2525**:2127–35. http://doi.org/10.1177/0956797614551970.

Lowery BS, Eisenberger NI, Hardin CD, *et al*. Long-term effects of subliminal priming on academic performance. *BASP* 2007;**29**:151–7. http://doi.org/10.1080/01973530701331718.

McClelland JL, McNaughton BL, O'reilly RC. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 1995;**102**:419–57.

Mealor AD, Dienes Z. Conscious and unconscious thought in artificial grammar learning. *Conscious Cogn* 2012;**21**:865–74. http://doi.org/10.1016/j.concog.2012.03.001.

Miller J. Measurement error in subliminal perception experiments: simulation analyses of two regression methods. *J Exp Psychol-Hum Percept Perform* 2000;**26**:1461–77.

Monahan JL, Murphy ST, Zajonc RB. Subliminal mere exposure: specific, general, and diffuse effects. *Psychol Sci* 2000;**11**:462–6. http://doi.org/10.1111/1467-9280.00289.

Moscovitch M. Recovered consciousness: a hypothesis concerning modularity and episodic memory. *J Clin Exp Neuropsychol* 1995;**17**:276–90. http://doi.org/10.1080/01688639508405123.

Mudrik L, Breska A, Lamy D, *et al*. Integration without awareness: expanding the limits of unconscious processing. *Psychol Sci* 2011;**22**:764–70. http://doi.org/10.1177/0956797611408736.

Mudrik L, Faivre N, Koch C. Information integration without awareness. *Trends Cogn Sci* 2014;**18**:488–6. http://doi.org/10.1016/j.tics.2014.04.009.

Mudrik L, Koch C. Differential processing of invisible congruent and incongruent scenes: a case for unconscious integration. *J Vis* 2013;**13**:24. http://doi.org/10.1167/13.13.24.

Negash S, Kliot D, Howard DV, *et al*. Relationship of contextual cueing and hippocampal volume in amnestic mild cognitive impairment patients and cognitively normal older adults. *J Int Neuropsychol Soc* 2015;**21**:285–96. http://doi.org/10.1017/S1355617715000223.

Newell BR, Shanks DR. Unconscious influences on decision making: a critical review. *Behav Brain Sci* 2014;1–19. http://doi.org/10.1017/S0140525X12003214.

Nieuwenstein MR, Wierenga T, Morey RD, *et al*. On making the right choice: a meta-analysis and large-scale replication attempt of the unconscious thought advantage. *Judgm Decis Mak* 2015;**10**:1–17.

Ocampo B. Unconscious manipulation of free choice by novel primes. *Conscious Cogn* 2015;**34**:4–9. http://doi.org/10.1016/j.concog.2015.03.007.

Olsen RK, Moses SN, Riggs L, *et al*. The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front Hum Neurosci* 2012;**6**:146. http://doi.org/10.3389/fnhum.2012.00146.

Parkinson J, Haggard P. Subliminal priming of intentional inhibition. *Cognition* 2014;**130**:255–65. http://doi.org/10.1016/j.cognition.2013.11.005.

Reber TP, DoLam A, Elger CE, *et al*. Intracranial EEG correlates of implicit relational inference within the hippocampus. *Hippocampus* 2016;**26**:54–66. http://doi.org/10.1002/hipo.22490.

Reber TP, Henke K. Integrating unseen events over time. *Conscious Cogn* 2012;**21**:953–60. http://doi.org/10.1016/j.concog.2012.02.013

Reber TP, Luechinger R, Boesiger P, *et al*. Unconscious relational inference recruits the hippocampus. *J Neurosci* 2012;**32**:6138–48. http://doi.org/10.1523/JNEUROSCI.5639-11.2012.

Reder LM, Park H, Kieffaber PD. Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychol Bull* 2009;**135**:23–49. http://doi.org/10.1037/a0013974.

Reuss H, Kiesel A, Kunde W, *et al*. Unconscious activation of task sets. *Conscious Cogn* 2011;**20**:556–67. http://doi.org/10.1016/j.concog.2011.02.014.

Ruby P, Caclin A, Boulet S, *et al*. Odd sound processing in the sleeping brain. *J Cognitive Neurosci* 2008;**20**:296–311. http://doi.org/10.1162/jocn.2008.20023.

Ruch S, Koenig T, Mathis J, *et al*. Word encoding during sleep is suggested by correlations between word-evoked up-states and post-sleep semantic priming. *Front Psychol* 2014;**5**:1319. http://doi.org/10.3389/fpsyg.2014.01319.

Schlaghecken F, Eimer M. Masked prime stimuli can bias "free" choices between response alternatives. *Psychon Bull Rev* 2004;**11**:463–8. http://doi.org/10.3758/BF03196596.

Shanks DR. Learning: from association to cognition. *Annu Rev Psychol* 2010;**61**:273–301. http://doi.org/10.1146/annurev.psych.093008.100519.

Simonson I. In defense of consciousness: the role of conscious and unconscious inputs in consumer choice. *J Cons Psychol* 2005;**15**:211–17. http://doi.org/10.1207/s15327663jcp1503_5.

Sklar AY, Levy N, Goldstein A, *et al*. Reading and doing arithmetic nonconsciously. *Proc Natl Acad Sci USA* 2012. http://doi.org/10.1073/pnas.1211645109.

Squire LR, Zola SM. Structure and function of declarative and nondeclarative memory systems. *Proc Natl Acad Sci USA* 1996;**93**:13515–22.

Strahan EJ, Spencer SJ, Zanna MP. Subliminal priming and persuasion: striking while the iron is hot. *J Exp Soc Psychol* 2002;**38**:556–68. http://doi.org/10.1016/S0022-1031(02)00502-4

Tononi G. An information integration theory of consciousness. *BMC Neurosci* 2004;**5**:42. http://doi.org/10.1186/1471-2202-5-42.

Tulving E. Episodic memory: from mind to brain. *Annu Rev Psychol* 2002;**53**:1–25. http://doi.org/Article

Vadillo MA, Kostopoulou O, Shanks DR. A critical review and meta-analysis of the unconscious thought effect in medical decision making. *Conscious Res* 2015;636. http://doi.org/10.3389/fpsyg.2015.00636.

Van den Bussche E, Van den Noortgate W, Reynvoet B. Mechanisms of masked priming: a meta-analysis. *Psychol Bull* 2009;**135**:452–77. http://doi.org/10.1037/a0015329.

van Gaal S, De Lange FP, Cohen MX. The role of consciousness in cognitive control and decision making. *Front Hum Neurosci* 2012;**6**:121. http://doi.org/10.3389/fnhum.2012.00121.

Van Opstal F, de Lange FP, Dehaene S. Rapid parallel semantic processing of numbers without awareness. *Cognition* 2011;**120**:136–47. http://doi.org/10.1016/j.cognition.2011.03.005.

Verwijmeren T, Karremans JC, Bernritter SF, *et al*. Warning: You are being primed! The effect of a warning on the impact of subliminal ads. *J Exp Soc Psychol* 2013;**49**:1124–9. http://doi.org/10.1016/j.jesp.2013.06.010.

Verwijmeren T, Karremans JC, Stroebe W, *et al*. The workings and limits of subliminal advertising: the role of habits. *J Cons Psychol* 2011;**21**:206–13. http://doi.org/10.1016/j.jcps.2010.11.004.

Vlassova A, Donkin C, Pearson J. Unconscious information changes decision accuracy but not confidence. *Proc Natl Acad Sci USA* 2014;**111**:16214–8. http://doi.org/10.1073/pnas.1403619111.

Weinberger J, Westen D. RATS, we should have used Clinton: subliminal priming in political campaigns. *Polit Psychol* 2008;**29**:631–51. http://doi.org/10.1111/j.1467-9221.2008.00658.x.

Züst MA, Colella P, Reber TP, *et al*. Hippocampus is place of interaction between unconscious and conscious memories. *PLoS ONE* 2015;**10**:e0122459. http://doi.org/10.1371/journal.pone.0122459.

Supplementary Figures
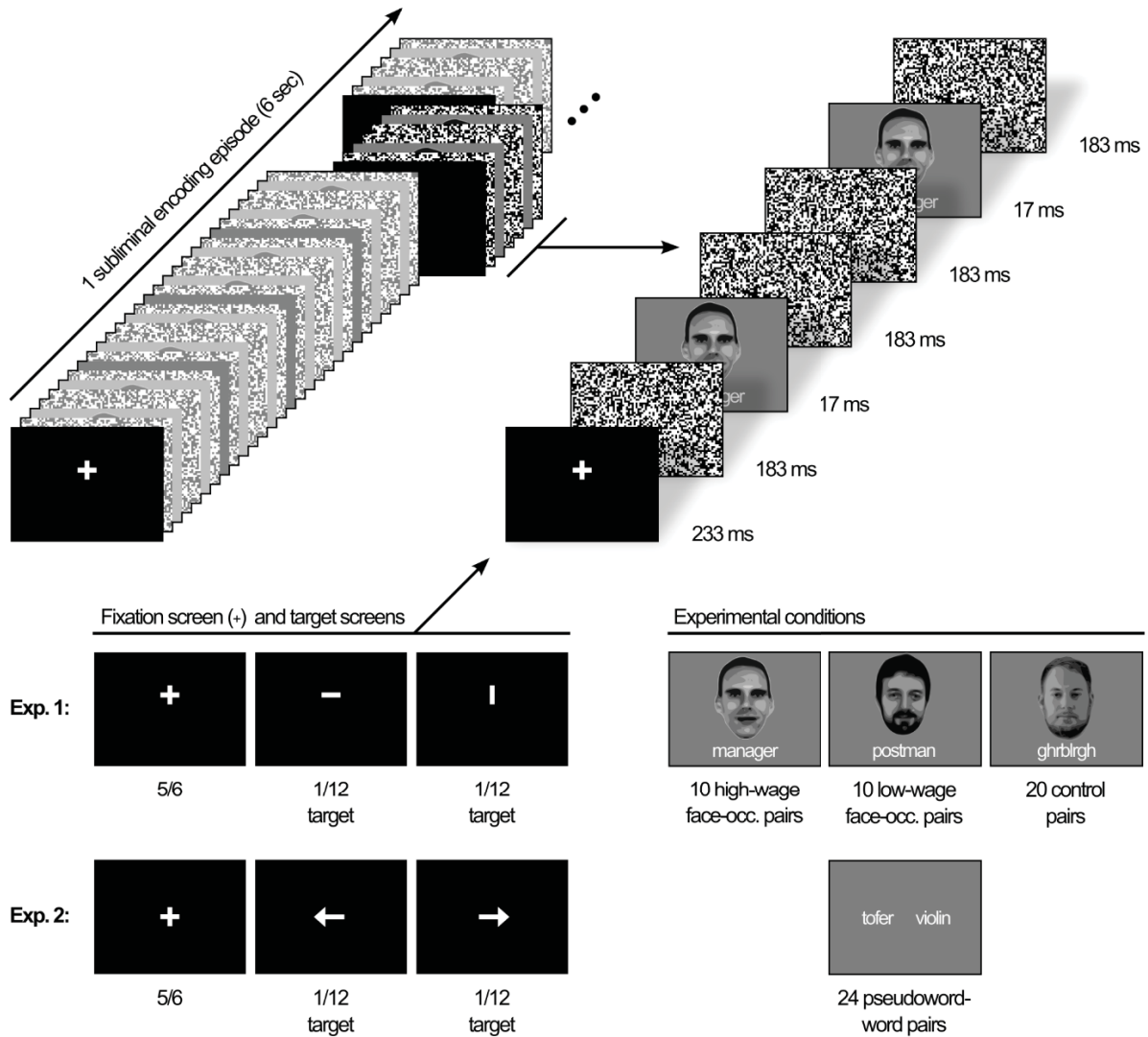
Supplementary Figure 1.

Subliminal encoding episode. One episode consisted of 5 supraliminal fixation screens and one randomly positioned target screen that were displayed for 233 ms each. Participants had to respond to target screens by pressing the left or right mouse button (left for "–" or "←" and right for "|" or "→" in experiment 1 or 2 respectively). In one encoding episode, one subliminal stimulus was presented 12 times for 17 ms each. Stimuli either consisted of face-occupation or face-nonword pairs (experiment 1), or of pseudoword-word pairs (experiment 2). Each stimulus was preceded and followed by a visual noise mask (183 ms).

Vocabulary acquisition during sleep

Marc Alain Züst*†[1], Simon Ruch†[1], Roland Wiest[2] & Katharina Henke[1].

[1] Department of Psychology and Center for Cognition, Learning and Memory, University of Bern, Switzerland.

[2] Support Center for Advanced Neuroimaging (SCAN) and Institute of Diagnostic and Interventional Neuroradiology, University Hospital of Bern, Switzerland.

* Corresponding author

† These authors contributed equally to the manuscript

Abstract

Recent evidence suggests verbal comprehension and simple nonverbal associative learning is possible during sleep. We investigated vocabulary learning during sleep in healthy, young participants hypothesizing that new semantic association between foreign and German words can be established in long-term memory during deep sleep. We presented novel vocabulary during a midday nap and assessed the memory of sleep-played word pairs with an indirect cued-retrieval test following waking. Performance indicated that participants had acquired new vocabulary during slow-wave sleep. Electrophysiological recordings revealed that word pairs were only encoded and stored long-term if the second word of a pair was presented into an ongoing slow wave peak. Functional magnetic resonance imaging during implicit retrieval following sleep revealed that a successful associative retrieval was mediated by hippocampus, as well as cortical language areas and semantic storage sites. These findings suggest that hippocampal-dependent associative encoding and long-term storage are feasible during the unconsciousness of deep sleep.

Introduction

Learning during sleep is an ancient dream of mankind, but current theories on the function of sleep suggest it is not feasible. Sleep is thought of as a period of rest during which the brain ceases to process information from the outside world[1]. Rather, information processing is focused inward to replay memories[2] of the preceding waking period. This replay allows hippocampus, which initially encodes and stores new information during wakefulness, to transfer these memories to neocortex for long-term storage[3,4]. Besides enabling this consolidation process, sleep is also thought to compensate for learning-induced synaptic growth by down-scaling or erasing unused synaptic connections. This process helps maintain synaptic homeostasis and makes room for new learning[5]. Both processes – memory consolidation and synaptic down-scaling – not only benefit from the reduced sensory input that defines sleep, but might actually be incompatible with learning of new information presented during sleep[6–8]. Recent evidence suggests, however, that sleeping humans can recognize their own names[9], can distinguish sensible form nonsensical sentences[10], and can encode words for more efficient processing after waking[11]. Moreover, humans can learn new tone-odor associations during sleep via trace conditioning[12,13], a variation of classical conditioning that is thought to depend on hippocampus[14].

Given that semantic processing of verbal messages[9–11] and formation of new nonverbal associations[12,13] are possible during sleep, we hypothesized that humans should be able to acquire new *verbal associations* – i.e. learn new vocabulary during sleep. We assumed that the unconsciousness of sleep would prevent formation of conscious memories that can be explicitly reported following waking. Sleep sleep-learning was expected to induce implicit knowledge that would guide behavior unconsciously[11–13]. To study vocabulary learning during sleep, we presented pairs of foreign words and German object

words, like "*tofer = Haus (house)*" to participants who were in deep sleep during an afternoon nap (**Fig. 1a**). We targeted deep sleep because of its involvement in memory consolidation processes[3,4], and because deep sleep has been shown to be receptive to external triggering of consolidation[15] and to learning[12]. Most importantly, deep sleep, also called slow-wave sleep, is dominated by high amplitude, low-frequency (< 4 Hz) electrical brain activity as recorded with electroencephalogram (EEG)[3,4,16]. These eponymous slow waves demark states of increased cortical excitability during peaks (corresponding to depolarized "up" states on neuronal level) and states of cortical silence during troughs (corresponding to hyperpolarized "down" states on neuronal level)[4,16,17]. During slow-wave peaks, neuronal network properties resemble fragments of a wake-like state, facilitating information transfer between brain regions[18–20]. We therefore hypothesized that slow wave peaks might provide windows of opportunity for associative learning during deep sleep. To utilize the sleeping brain's excitability during slow-wave peaks, we set the onset asynchrony of the two words of each German-foreign word pair to the expected peak-to-peak interval of slow-waves. Acoustic stimulation at that interval is known to induce and enhance slow-wave activity[21] and should therefore benefit sleep-learning. In the subsequent waking period, we tested whether participants had associated semantic aspects of German object words to foreign words by asking about physical size occupied by the object the foreign words stood for – an object property that can correctly be identified only if semantic associations between German and foreign words had been formed during sleep. For instance, we expected participants to classify "*tofer*" as a large object if it had successfully been associated with "*house*" during sleep. As vocabulary acquisition is a hippocampus-dependent form of relational learning[22,23], we expected retrieval of sleep-learned word associations to recruit hippocampus and associated neocortical lexical-semantic storage sites. To test this hypothesis, a subsample of our participants underwent functional magnetic resonance imaging (fMRI) while their memory for sleep-played vocabulary was tested.

Results

Vocabulary presentation did not wake but entrained slow waves and sleep spindles
The 41 participants included in this study listened to an average of 45.10 word pairs (SD=5.28) while in EEG-defined deep sleep. Words were presented at a quiet volume, superimposed on constantly playing Brownian noise to reduce salience. Each pair was played four times in sequence, yielding ~187 stimulations per participant. Stimulation was interrupted upon detection of arousals or waking responses. Word pairs associated with arousals or waking (3.44% of pairs per participant, SD = 7.54%) were post-hoc excluded from analysis. Most of the vocabulary was played during deep sleep (M = 90.73%, SD = 13.61%).

Electroencephalographic recordings during sleep suggested that auditory stimulation did not wake but induced a train of frontal slow-wave peaks that were accompanied by sleep spindles (Fig 2). Presentation of word pairs evoked three positive voltage peaks in frontal electrode clusters at 300-600 ms (cluster-based p-value < .001), 900-1500ms (p = .003), and 2000-2300 ms (p = .022) after stimulation onset. Peaks coincided with increased spectral power in the spindle frequency band (12-16 Hz) in wide-spread fronto-central electrode clusters (cluster-based p < .01 for all three time-windows). The 1st and 2nd voltage peaks were followed by troughs (500-900ms and 1500-2000ms) during which theta activity (4-8Hz) was significantly increased in large electrode clusters centered around the vertex (both p < .001). Vocabulary presentation thus evoked voltage-peaks and troughs that emerged at intervals (~1 s) representative of slow-waves (~1 Hz) and that orchestrated activity in the spindle and theta frequency range. Auditory stimulation during sleep is known to entrain slow-waves at ~1 Hz[11,21]. Furthermore, the

nesting of spindles in peaks and theta bursts in troughs has been observed for both spontaneous and induced sleep slow-waves[11,15,16]. This suggests that stimulation did not wake but entrained oscillations that are hallmarks of deep sleep.

Humans can acquire new vocabulary during sleep

Following waking, each sleep-played foreign word was presented aurally and visually for participants to rate whether it stands for a small or large object. This test assessed implicit memory for sleep-learned semantic associations. Sleep-learning was successful, as participants correctly classified the size of sleep-played foreign object words at a level significantly above what would be expected by chance ($52.7 \pm 1.1\%$ accuracy [$M \pm SEM$], $F(1,38) = 5.51$, $p = .024$, Cohen's $d = .36$; chance level = 50%, **Fig. 1b**). Classification accuracy did not significantly vary between subgroups of participants who did and did not undergo fMRI at test ($F(1,38) = 0.35$, $p = .56$) and significantly exceeded chance-level in both groups (fMRI: $52.1 \pm 1.8\%$, $F(1,13) = 9.13$, $p = .01$, $d = .29$; no fMRI: $53.1 \pm 1.2\%$, $F(1,24) = 5.63$, $p = .026$, $d = .49$). Importantly, participants were unable to name any of the sleep-played words and denied having noticed word presentation during sleep. Hence, sleep-learning induced unconscious representations of new vocabulary.
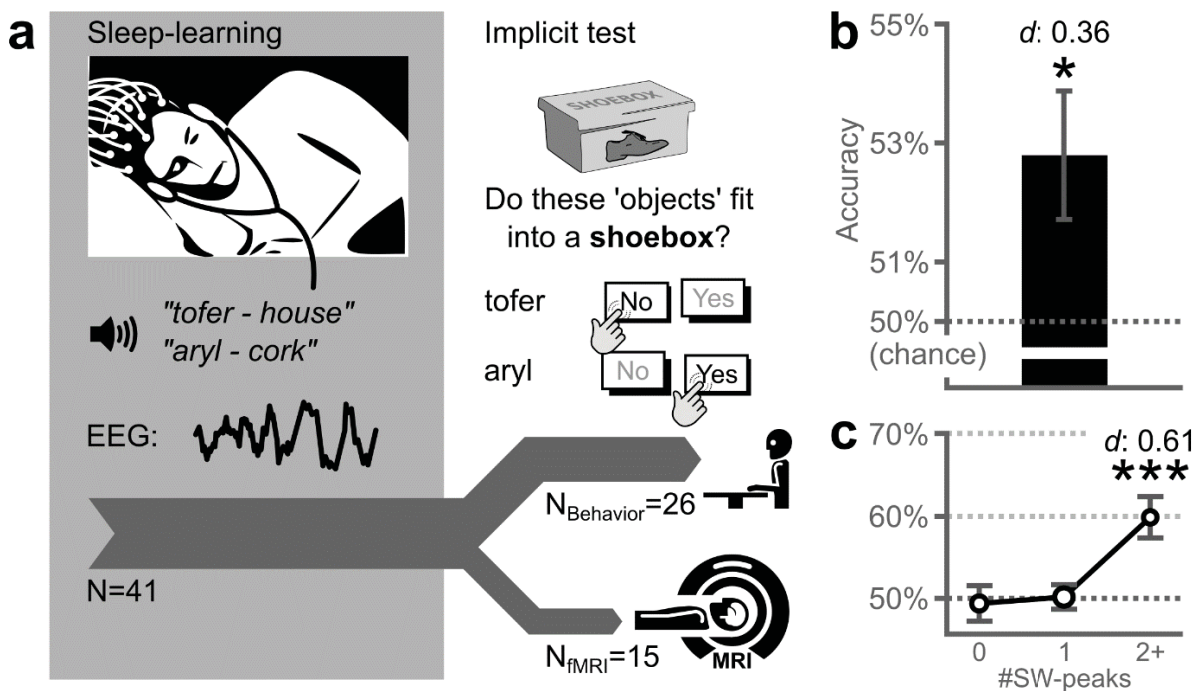

Figure 1 Experimental design and behavioral results. **(a)** We repeatedly presented novel vocabulary consisting of German-foreign word pairs to forty-one participants who were in deep sleep during an afternoon nap. Deep sleep is characterized by high-amplitude slow waves (~1 Hz) visible in the EEG. Stimulus onset asynchrony between German and foreign word was 1075 ms, which corresponds to the expected frequency of slow waves. After waking, participants completed an implicit test for semantic associations, where they intuitively judged foreign words according to their size. If sleep-learning was successful, we expect participants to judge foreign words' size in accordance with their sleep-played translation. Participants indicated by keyboard-press whether they thought objects described by foreign words would fit into a shoebox or not. A subgroup of participants ($N_{fMRI} = 15$)

completed this task while their brain activity was recorded with functional magnetic resonance imaging. **(b)** Despite being unaware of sleep-learning, participants (N = 41) correctly identified size of foreign object words above chance level (52.7% vs. 50%, Cohen's *d* = .38), indicating implicit memory for sleep-played vocabulary. **(c)** Sleep-learning depended on the number of times the second word in a pair was played into an ongoing slow-wave peak (#SW-peaks). Pairs in which the second stimulus never or only once hit an ongoing peak could not be retrieved afterwards (50% accuracy). However, pairs for which two or more repetitions hit a peak reached 60% classification accuracy (circle area indicates relative number of pairs with no, one, and two or more slow-wave peaks); *p = .024, ***p < .001; error bars represent the *SEM*.

Slow wave peaks mediate sleep-learning

Assuming that slow-waves would contribute to memory formation during sleep, we assessed whether the evoked potential (ERP) over frontal electrodes – where slow-waves originate[24] – could predict retention of sleep-played vocabulary. We found that the latency of the 2nd evoked slow-wave peak was related to performance at test (**Fig. 2a**). An early 2nd peak around 800-1100 ms, just *before* the onset of the second word, was associated with correct responses in the shoebox task. Conversely, a late 2nd slow wave peak around 1200-1500 ms, just *after* onset of the second word, was associated with incorrect responses (**Fig. 2a,** green vs. red trace). This was confirmed by the significant time*accuracy interaction on mean voltage in the ERP ($F(1,40)$ = 8.23, p = .006, d = .91). Presumably, presenting the second word of a pair into an already ongoing slow-wave peak provided optimal conditions for relational encoding. The neuronal excitability and cortical connectivity provided by this slow-wave peak[19,20] probably allowed the first word to be kept in working memory[10] and to be bound to the immediately following second word.

To further assess the role of slow waves in sleep-learning, we analyzed the phase of evoked frontal oscillations at a frequency of 0.8 Hz, the prominent frequency of human slow-waves[25]. Studies using auditory cues to trigger reactivation and consolidation of memories during sleep suggested that cueing is successful only if stimuli are presented during a specific phase of these waves[26]. If a specific phase at a specific time during or after word presentation contributes to the encoding or consolidation of sleep-played stimuli, retention of German-foreign pairs presented in sync with that phase should be enhanced. To test this hypothesis, we extracted the 0.8 Hz phase-angle at each time point for each presentation of a German-foreign pair and computed for each time-phase bin the percentage of presentations that led to correct classifications in the shoebox task. A phase angle resembling a slow wave peak (0 rad) between 800 to 1100 ms was associated with increased accuracy in the shoebox task (**Fig. 2b**). A cluster-based permutation test suggested that this association between time, phase, and accuracy was significant under the null-assumption that accuracy be independent of slow-wave phase and therefore uniformly distributed across all phase angles (p = .045). The phase analysis thus corroborated our observation that presenting the second stimulus of a pair into an ongoing slow-wave peak enhanced encoding.

To quantify the beneficial effect of slow-wave peaks in sleep-learning, we assessed whether retention improved with repeated presentation of German-foreign word pairs into ongoing up-states. To this aim, we counted for each pair the number of repetitions during which the phase angle was close to 0 rad between 800 and 1100 ms. Pairs for which the second word was never or only once presented during an ongoing peak (36.6% and 41.9% pairs per participant) yielded chance-level performance in the shoebox task (49.4% and 50.2% respectively, all p > .78). However, those 26.5% of pairs in which the second word repeatedly hit an ongoing peak yielded 59.9% correct classifications in the shoebox task (SEM = 2.5%; t(40) = 3.92, p < .001, d = 0.62). Hence, only pairs that were repeatedly played in optimal

synchronization with slow-wave peaks were retained, but these pairs yielded performance that exceeded chance level by ~10%.

Low theta-band power mediates sleep-learning

We assessed whether induced activity in theta (5-7 Hz) and fast spindle (13-15 Hz) frequency bands would predict successful retention of sleep-played vocabulary. Theta bursts and fast spindles are associated with memory consolidation during sleep[3,4,15,27,28] and might thus also contribute to sleep-learning. Theta but not fast spindle band activity was related to learning success, with later correct German-foreign pairs tonically inducing less theta band power than later incorrect pairs. This difference was significant in a frontal cluster of electrodes between 800-1100 ms (cluster-based p < .001 **Fig. 2c**), i.e. at the time of the encoding-related slow-wave peak. This result suggests that high theta-activity during onset of the second pair in a word is detrimental for sleep-learning, potentially because theta is a marker for ongoing consolidation[15,28]. Ongoing consolidation processes could prevent the binding of the upcoming word to the just presented stimulus, thereby impairing relational encoding.
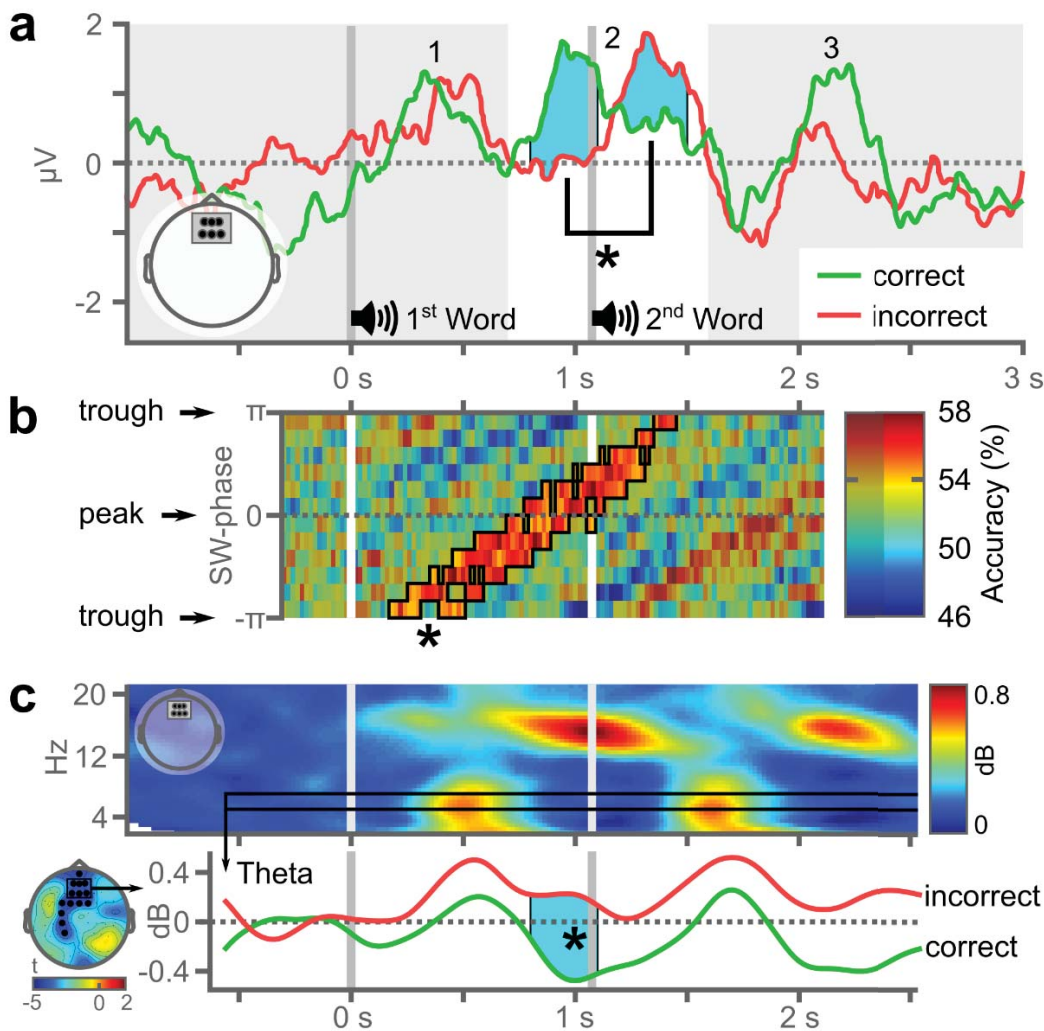

Figure 2 Electrophysiological brain responses predicts sleep-learning success. Plots depict average responses across all trials and participants, time-locked to the presentation of the first word in pairs, over frontal electrodes

(F1, F2, Fz, FC1, FC2, FCz). * *p* < .05, N = 41. **(a)** Event-related potentials (ERPs) for later correct (green) and incorrect (red) responses. Sleep-playing vocabulary led to three entrained slow-wave (SW) peaks (numbered 1-3), with different latency of the 2nd SW peak for later correct vs. incorrect trials (white highlight). An early 2nd SW peak with a maximum just *before* the second word was being played was beneficial for encoding (*p < .001). **(b)** Time-phase-accuracy plot of the SW-frequency (0.8 Hz), color-coded according to mean retrieval performance for trials in time-phase bins. Every 20 ms column contains all trials of all participants. The beneficial early 2nd SW peak manifests as a significant cluster (*p* = .045, outlined and starred) of increased performance for trials following a SW-phase response with a trough at around 400 ms and a peak just before presentation of the second word at around 800 ms, closely resembling the ERP for correct trials in (a). Bins not belonging to this cluster have been darkened for better visibility; separate color bars are provided. **(c)** Analysis of event-related spectral power suggests that stimulation evoked increases in theta (5-7 Hz) and fast spindle (13-15 Hz) activity (top panel). Evoked theta power was tonically lower for later correct vs. incorrect responses (bottom panel). This difference was significant around presentation of the second word, which coincides with the SW-phase early-shift (*cluster-based *p* < .001, bottom panel, blue highlight). This effect manifested in a frontal electrode cluster (bottom left panel).

Retrieval of sleep-learned vocabulary recruits hippocampus

A subgroup of participants underwent fMRI while their memory for sleep-played vocabulary was tested. We contrasted the resulting functional brain images between correct and incorrect responses in the shoebox task to find brain areas associated with successful retrieval of sleep-learned content. To determine which responses truly mirrored the presence vs. absence of sleep-formed memories, we informed the fMRI analysis with the mean slow-wave phase-angle each German-foreign word pair had evoked at 800-1100 ms after onset during sleep. Phase-angles indicating a slow-wave peak in this time window were found beneficial for associative encoding during sleep. Hence, correct responses to foreign words associated with a slow-wave peak and incorrect responses to words associated with a trough were most likely to indicate presence vs. absence of sleep-formed memories. Taking slow-wave phase into account thus allowed us to more precisely estimate the neural signature of successful retrieval of sleep-learned memories.

Retrieval of sleep-learned vocabulary recruited the right hippocampus in a cluster of of 17 voxels (peak at 28, -16, -10, MNI, p_{peak} = .039, family-wise error corrected for hippocampal subvolume (FWE$_{SV}$); **Fig. 3a**). Contrast estimates at the peak voxel revealed that processing of retrieved German-foreign word pairs activated hippocampus (*M±SEM*: 0.26 ± 0.08, arbitrary units), whereas processing of forgotten pairs led to a deactivation (*M±SEM*: -0.21±0.09). Although we did not record fMRI during sleep, the observed hippocampal recruitment during retrieval of sleep-formed vocabulary suggests that associative encoding during sleep depended on hippocampus.

Successful retrieval activates language regions

Along with hippocampus, several cortical areas were activated during successful retrieval, indicating enhanced perceptual processing of foreign words (**Fig. 3a**). A prominent cluster encompassing 1417 voxels peaked in the right lateral premotor (Brodmann area (BA) 6) and sensorimotor (BA 3/4) cortex. This activity was mirrored in the left cortex (191 voxels). Because participants performed the same right-hand motor tasks for retrieved and forgotten words, sonsorimotor activity can not be due to differences in motor activity. Given its role in phonetic processing, activation of sensorimotor cortex rather indicates enhanced phonetic analysis of sleep-learned foreign words[29]. Note that successful retrieval was also associated with bilateral activity in lingual gyri, wich are thought to contribute to processing of written word forms[30]. We speculate that successful sleep-learning facilitated audio-visual integration of spoken foreign words with their written form, yielding enhanced coactivation of areas implicated in the processing of phonemes and written words.

Importantly, successful retrieval activated language comprehension sites and semantic storage sites – albeit with a right-hemisphere bias. In fact, the large right-hemispheric sensorimotor cluster extended into inferior parietal cortex around the supramarginal gyrus (BA 40), into the anterior insula (BA 13), and the inferior frontal gyrus, pars opercularis (BA 44). Inferior parietal cortex activations as well as insular activity (cluster of 29 voxels) were mirrored in the left hemisphere. Successful retrieval was further associated with enhanced activity in right inferior temporal sulcus (BA 20/21, 17 voxels). These regions play vital roles in language comprehension[31] and semantic storage[32]. Hence, hippocampus-mediated retrieval of sleep-learned vocabulary recruited semantic networks.

The right lateralized activation of language homologues could be due to the the unconscious nature of sleep-learned memories. Studies using visual masking to investigate unconscious processing of written words suggested that the right hemisphere contributes to unconscious but not conscious reading[33] and that encoding of unconscious word associations activates right-lateralized language homologues[34]. The strong right-hemispheric activity could further mirror the semantic network and the spatial computations on which the implicit memory test demands. Participants had to estimate the size of objects for which the foreign words stood. Such rather coarse semantic analyses typically recruit right hemispere[35,36]. Participants further had to decide whether objects fit into a shoebox. This decision arguably depends on spatial imagery and mental rotation - computations that are biased to the right hemisphere[37]. In sum, successful retrieval of sleep-learned foreign words recruited a network that indicated enhanced relational processing and semantic analysis its German translation.

Hippocampus and temporal pole activity linearly relates to sleep-learning success
Activity in the right hippocampus was linearly related to retrieval performance across participants: The better the performance in the shoebox task, the more strongly the hippocampus was activated. This result was suggested by a between-subject correlation of the unmodulated functional brain signal with average accuracy in the shoebox task. Three clusters in the right hippocampus were correlated with performance (r = .75-.81, p_{peak} < .001, uncorrected, one cluster surviving FWE$_{SV}$-correction at p_{peak} = .018; **Fig. 3b**). Additionally, activity in the left temporal pole (19 voxels, peak at -36, 18, -36) correlated with performance at test. The left temporal pole has been associated with speech comprehension[38], semantic processing of concrete objects[39], and linkage of semantic features[40], like in this case, the size of an object linked to a new word. The result that temporopolar activity scales with retrieval success for sleep-learned vocabulary alongside the hippocampus suggests that relational encoding and storage of German-foreign word pairs depends on close cross-talk between hippocampus and temporopolar storage sites.
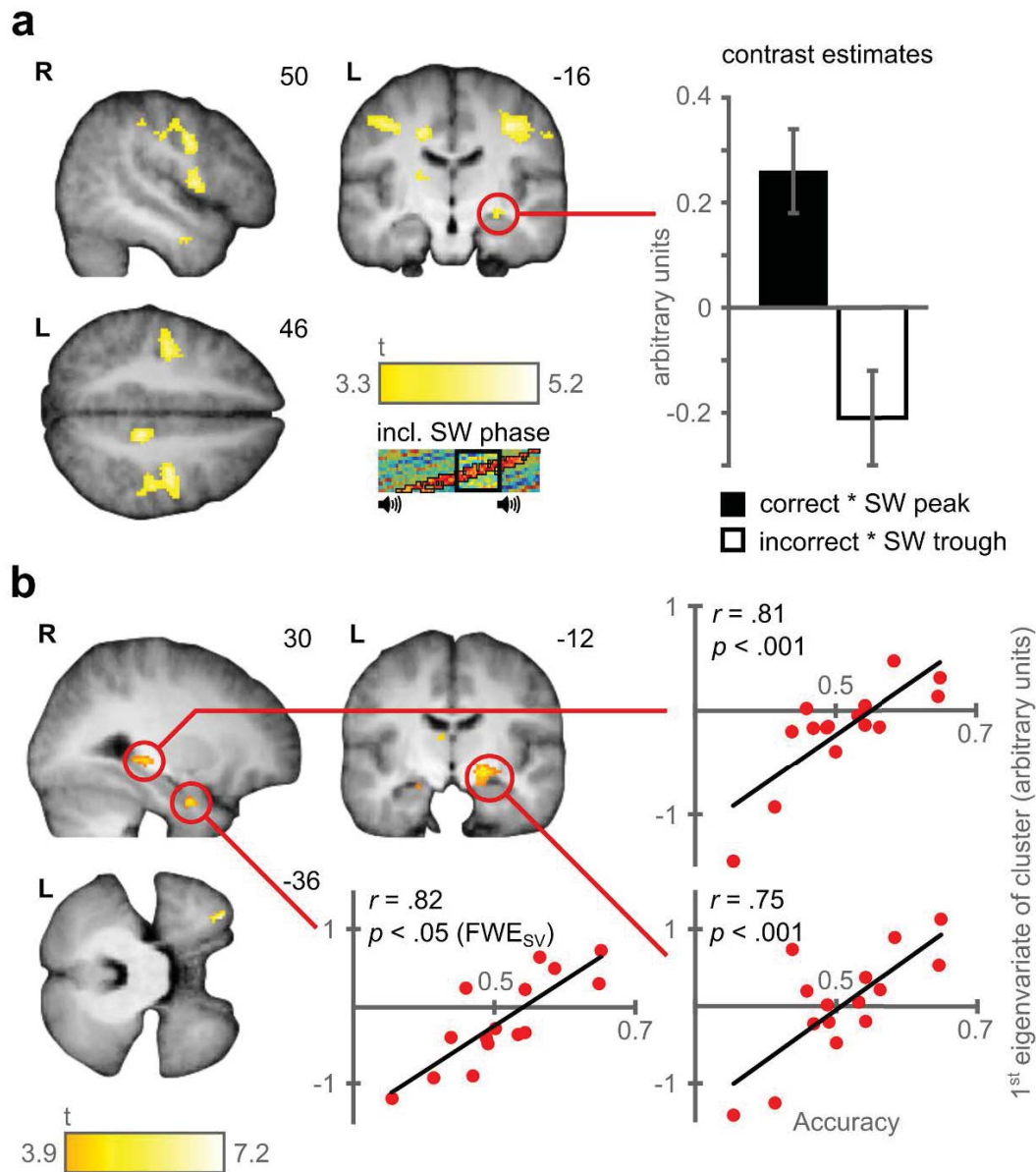
Figure 3 Brain areas associated with successful retrieval. Activity overlaid on the averaged anatomical image of participants. Colors range from critical to maximal values. Hippocampus is assessed at $p < .05$, voxel-wise, using small-volume family-wise-error correction (FWE$_{SV}$), whole-brain results are assessed using $p < .001$, uncorrected voxel-wise threshold, and cluster size ≥ 10. N = 15. **(a)** Slow wave (SW) phase-informed contrast of correct > incorrect responses to sleep-played foreign words. The contrast shows areas that are more strongly activated for truly remembered versus truly forgotten items under the assumption that slow wave phase during sleep-learning had been essential for successful encoding. This contrast yielded extended activity in parietal, frontal and temporal areas associated with vocabulary acquisition, and the right hippocampus. The hippocampal cluster is highlighted and parameter estimates are plotted at its peak voxel indicating hippocampus both activates for truly remembered and deactivates for truly forgotten items. Error bars represent *SEM*. **(b)** Brain behavior correlation (contrast correct > incorrect) $\propto$ (mean decision accuracy). Three hippocampal clusters of activity are linearly associated with retrieval success of sleep-played vocabulary, as visualized in scatterplots with Pearson's correlation coefficient (*r*). One hippocampal cluster survives at $p < .05$ FWE$_{SV}$ voxel-wise. The other two hippocampal clusters are reported at $p < .001$, voxel-wise unc., and cluster size ≥ 10. The cluster displayed in the coronal section (y=-12) overlaps with

the cluster found in the phase-informed contrast in (a). Hippocampal clusters are displayed at *p* < .005 for easy visibility. An additional cluster in the left temporal pole scales activity with retrieval performance, displayed in the transversal section (z=-36).

Discussion

We found that auditory presentation of foreign words and their translation during deep NREM sleep helped participants to semantically categorize the sleep-played foreign words following waking, i.e. to decide whether foreign words designate large or small objects. Accuracy of these semantic decisions exceeded chance level despite participants being unable to consciously recall the event or the content of stimulation during sleep. This is evidence that humans can acquire implicit relational knowledge about foreign vocabulary that is played during the unconsciousness of deep sleep and can apply this knowledge during ensuing wakefulness.

Vocabulary learning during sleep was mediated by the phase of slow-waves. Although electrically induced slow-waves during wakeful learning can improve memory encoding[41], slow-waves during sleep have mainly been associated with memory consolidation[4,42]. Auditory cues played to trigger memory reactivation and consolidation during sleep were found to enhance memory if cues were presented into the transition to a slow-wave trough[43] or if they were followed by a trough associated with increased theta activity[15]. We observed enhanced encoding of sleep-played German-foreign word pairs if the second word of the pair – either the foreign word or its translation – was presented into an ongoing peak of a slow-wave. This peak was accompanied by reduced theta activity. This suggests distinct contributions of slow-wave troughs/enhanced theta to memory *consolidation* and slow-wave peaks/reduced theta to memory *formation*. During consolidation, hippocampus transfers information to neocortex[3,4], whereas during encoding, hippocampus encodes novel information from neocortex to bind aspects of events or stimuli into associative memory traces. These distinct cortico-hippocampal interactions call upon different cerebral states, i.e. peaks vs. troughs. It has been suggested that slow-wave peaks represent brief phases of wake-like neocortical excitability and connectivity[18–20] that benefit processing of sensory input[19]. We speculate that peaks preceding the second word of a German-foreign pair during sleep provided optimal conditions for both words to be activated in neocortex, allowing hippocampus to encode these word pairs as new vocabulary.

Implicit retrieval of sleep-learned vocabulary recruited hippocampus along with neocortical areas for language processing and semantic storage. This was suggested by fMRI data obtained at test. Although we did not record hippocampal activity during sleep, its involvement in retrieval of sleep-formed memories suggests that hippocampus contributed to vocabulary learning during sleep. Direct evidence for hippocampal ability to form new memories during sleep has previously only been obtained in animals, using intracranial EEG during sleep-learning[44]. Studies suggesting sleep-learning in humans only indirectly inferred hippocampal involvement through use of nonverbal trace-conditioning[12,13]. Our study confirms and extends existing evidence of sleep-learning by showing directly that sleep-formed associative memory depends on hippocampus.

The involvement of hippocampus in relational learning during sleep corroborates the emerging view that hippocampal processing does not depend on consciousness[45,46]. Traditional theories of human memory posit that hippocampal learning – which includes vocabulary acquisition – requires consciousness during encoding[47,48]. However, hippocampus-dependent memory formation has been

documented under the exclusion of conscious awareness in the waking state[49,50]. Evidence for vocabulary learning during sleep extends these findings by showing that hippocampus encodes new relational information not only if it is masked from consciousness but also if it is presented during a state of unconsciousness.

Evidence for learning during sleep further informs theories about the function of sleep, which are consolidation of new memories and downscaling of unused synaptic connections. Deep sleep is a brain state that provides the neurochemical milieu[6] and the functional connectivity[7] for hippocampus to reactivate and consolidate memories. However, this state is thought to compromise hippocampus' ability to encode new information[4,51]. Slow-wave sleep is further thought to favor synaptic depression over long-term potentiation in order to compensate for the synaptic growth that resulted from wakeful learning[5,8]. This reduced expression of long-term potentiation is incompatible with learning. Nevertheless, our evidence for vocabulary learning during sleep suggests that the sleeping brain does not abandon its capability to learn. This should not be surprising, as it is known that memories are not only strengthened during sleep but are restructured to provide new insights[52]. We argue that memory restructuring leading to new insights is a form of learning that depends on hippocampal plasticity. Recent findings further suggest that synaptic depression or down-scaling might be achieved during rapid eye-movement (REM) sleep rather than slow-wave sleep[53,54]. This is supported by the observation that non-verbal learning is more robust during slow-wave sleep than REM sleep[12]. In sum, our findings suggest that memory consolidation does not preclude learning of new information, and that expression of synaptic potentiation is preserved during deep sleep.

Our findings have several limitations. First of all, the longevity of sleep-learned memories remains unknown because we tested memory within the first hour after waking. Furthermore, it is unclear if vocabulary learning is feasible during night-sleep because we studied sleep-learning during an afternoon nap. We further only report correlational, not causal evidence for a possible role of slow-wave peaks in relational encoding during sleep. Finally, our study provides no information about potential adverse effects of learning during sleep, such as impaired recuperation or diminished memory consolidation.

Despite these limitations, our findings demonstrate the feasibility of hippocampus-dependent vocabulary learning during sleep. Future studies will have to test whether sleep-learning can be improved by synchronizing vocabulary presentation with the optimal phase of slow-waves, and whether sleep-formed unconscious memories facilitate conscious vocabulary learning or improve the ability to derive the meaning of foreign words if presented in the context of a sentence.

Author contributions
K.H. and M.A.Z. conceived the idea. M.A.Z., S.R. and K.H. designed the experiments. M.A.Z. and S.R. carried out the experiments. M.A.Z. and S.R. analyzed the data. R.W. contributed fMRI analysis methods and tools. M.A.Z., S.R. and K.H. wrote the manuscript. All authors approved the final version of the manuscript.

Competing financial interests

The authors have no competing financial interests to declare.

References

1. Czisch, M. *et al.* Altered Processing of Acoustic Stimuli during Sleep: Reduced Auditory Activation and Visual Deactivation Detected by a Combined fMRI/EEG Study. *NeuroImage* **16,** 251–258 (2002).
2. Lee, A. K. & Wilson, M. A. Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep. *Neuron* **36,** 1183–1194 (2002).
3. Diekelmann, S. & Born, J. The memory function of sleep. *Nat. Rev. Neurosci.* **11,** 114–126 (2010).
4. Rasch, B. & Born, J. About Sleep's Role in Memory. *Physiol. Rev.* **93,** 681–766 (2013).
5. Tononi, G. & Cirelli, C. Sleep and the Price of Plasticity: From Synaptic and Cellular Homeostasis to Memory Consolidation and Integration. *Neuron* **81,** 12–34 (2014).
6. Rasch, B. H., Born, J. & Gais, S. Combined Blockade of Cholinergic Receptors Shifts the Brain from Stimulus Encoding to Memory Consolidation. *J. Cogn. Neurosci.* **18,** 793–802 (2006).
7. Mitra, A. *et al.* Human cortical–hippocampal dialogue in wake and slow-wave sleep. *Proc. Natl. Acad. Sci.* **113,** E6868–E6876 (2016).
8. Vyazovskiy, V. V., Cirelli, C., Pfister-Genskow, M., Faraguna, U. & Tononi, G. Molecular and electrophysiological evidence for net synaptic potentiation in wake and depression in sleep. *Nat. Neurosci.* **11,** 200–208 (2008).
9. Perrin, F., García-Larrea, L., Mauguière, F. & Bastuji, H. A differential brain response to the subject's own name persists during sleep. *Clin. Neurophysiol.* **110,** 2153–2164 (1999).
10. Daltrozzo, J., Claude, L., Tillmann, B., Bastuji, H. & Perrin, F. Working Memory Is Partially Preserved during Sleep. *PLoS ONE* **7,** e50997 (2012).
11. Ruch, S., Koenig, T., Mathis, J., Roth, C. & Henke, K. Word encoding during sleep is suggested by correlations between word-evoked up-states and post-sleep semantic priming. *Front. Psychol.* **5,** (2014).
12. Arzi, A. *et al.* Humans can learn new information during sleep. *Nat. Neurosci.* **15,** 1460–1465 (2012).
13. Arzi, A. *et al.* Olfactory Aversive Conditioning during Sleep Reduces Cigarette-Smoking Behavior. *J. Neurosci.* **34,** 15382–15393 (2014).
14. Clark, R. E., Manns, J. R. & Squire, L. R. Classical conditioning, awareness, and brain systems. *Trends Cogn. Sci.* **6,** 524–531 (2002).
15. Schreiner, T. & Rasch, B. Boosting vocabulary learning by verbal cueing during sleep. *Cereb. Cortex* **25,** 4169–4179 (2015).
16. Mölle, M., Marshall, L., Gais, S. & Born, J. Grouping of Spindle Activity during Slow Oscillations in Human Non-Rapid Eye Movement Sleep. *J. Neurosci.* **22,** 10941–10947 (2002).
17. Steriade, M., Nunez, A. & Amzica, F. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *J. Neurosci.* **13,** 3252–3265 (1993).
18. Destexhe, A., Hughes, S. W., Rudolph, M. & Crunelli, V. Are corticothalamic UP states fragments of wakefulness? *Trends Neurosci.* **30,** 334–342 (2007).
19. Schabus, M. *et al.* The fate of incoming stimuli during NREM sleep is determined by spindles and the phase of the slow oscillation. *Front. Sleep Chronobiol.* **3,** 40 (2012).
20. Cox, R., Driel, J. van, Boer, M. de & Talamini, L. M. Slow Oscillations during Sleep Coordinate Interregional Communication in Cortical Networks. *J. Neurosci.* **34,** 16890–16901 (2014).
21. Ngo, H.-V. V., Claussen, J. C., Born, J. & Mölle, M. Induction of slow oscillations by rhythmic acoustic stimulation. *J. Sleep Res.* **22,** 22–31 (2013).
22. Warren, D. E. & Duff, M. C. Not so fast: Hippocampal amnesia slows word learning despite successful fast mapping. *Hippocampus* **24,** 920–933 (2014).
23. Breitenstein, C. *et al.* Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage* **25,** 958–968 (2005).
24. Massimini, M., Huber, R., Ferrarelli, F., Hill, S. & Tononi, G. The Sleep Slow Oscillation as a Traveling Wave. *J. Neurosci.* **24,** 6862–6870 (2004).
25. Achermann, P. & Borbély, A. A. Low-frequency (<1 Hz) oscillations in the human sleep electroencephalogram. *Neuroscience* **81,** 213–222 (1997).

26. Batterink, L. J., Creery, J. D. & Paller, K. A. Phase of Spontaneous Slow Oscillations during Sleep Influences Memory-Related Processing of Auditory Cues. *J. Neurosci.* **36,** 1401–1409 (2016).

27. Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J. & Gaskell, M. G. Sleep Spindle Activity is Associated with the Integration of New Memories and Existing Knowledge. *J. Neurosci.* **30,** 14356–14360 (2010).

28. Schreiner, T., Lehmann, M. & Rasch, B. Auditory feedback blocks memory benefits of cueing during sleep. *Nat. Commun.* **6,** 8729 (2015).

29. Schomers, M. R. & Pulvermüller, F. Is the Sensorimotor Cortex Relevant for Speech Perception and Understanding? An Integrative Review. *Front. Hum. Neurosci.* 435 (2016). doi:10.3389/fnhum.2016.00435

30. Mechelli, A., Humphreys, G. W., Mayall, K., Olson, A. & Price, C. J. Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. *Proc. R. Soc. B Biol. Sci.* **267,** 1909–1913 (2000).

31. Friederici, A. D. The Brain Basis of Language Processing: From Structure to Function. *Physiol. Rev.* **91,** 1357–1392 (2011).

32. Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cereb. Cortex* **19,** 2767–2796 (2009).

33. Fecteau, J. H., Kingstone, A. & Enns, J. T. Hemisphere differences in conscious and unconscious word reading. *Conscious. Cogn.* **13,** 550–564 (2004).

34. Reber, T. P., Luechinger, R., Boesiger, P. & Henke, K. Unconscious relational inference recruits the hippocampus. *J. Neurosci.* **32,** 6138–6148 (2012).

35. Jung-Beeman, M. Bilateral brain processes for comprehending natural language. *Trends Cogn. Sci.* **9,** 512–518 (2005).

36. Yang, J. The role of the right hemisphere in metaphor comprehension: A meta-analysis of functional magnetic resonance imaging studies. *Hum. Brain Mapp.* **35,** 107–122 (2014).

37. Corballis, M. C. Mental Rotation and the Right Hemisphere. *Brain Lang.* **57,** 100–121 (1997).

38. Giraud, A. L. *et al.* Contributions of sensory input, auditory search and verbal comprehension to cortical activity during speech processing. *Cereb. Cortex N. Y. N 1991* **14,** 247–255 (2004).

39. Noppeney, U. & Price, C. J. Retrieval of Visual, Auditory, and Abstract Semantics. *NeuroImage* **15,** 917–926 (2002).

40. Noppeney, U. & Price, C. J. A PET Study of Stimulus- and Task-Induced Semantic Processing. *NeuroImage* **15,** 927–935 (2002).

41. Kirov, R., Weiss, C., Siebner, H. R., Born, J. & Marshall, L. Slow oscillation electrical brain stimulation during waking promotes EEG theta activity and memory encoding. *Proc. Natl. Acad. Sci.* **106,** 15460–15465 (2009).

42. Marshall, L., Helgadóttir, H., Mölle, M. & Born, J. Boosting slow oscillations during sleep potentiates memory. *Nature* **444,** 610–613 (2006).

43. Batterink, L. J., Creery, J. D. & Paller, K. A. Phase of Spontaneous Slow Oscillations during Sleep Influences Memory-Related Processing of Auditory Cues. *J. Neurosci.* **36,** 1401–1409 (2016).

44. de Lavilléon, G., Lacroix, M. M., Rondi-Reig, L. & Benchenane, K. Explicit memory creation during sleep demonstrates a causal role of place cells in navigation. *Nat. Neurosci.* **18,** 493–495 (2015).

45. Henke, K. A model for memory systems based on processing modes rather than consciousness. *Nat. Rev. Neurosci.* **11,** 523–532 (2010).

46. Reder, L. M., Park, H. & Kieffaber, P. D. Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychol. Bull.* **135,** 23–49 (2009).

47. Tulving, E. How many memory systems are there? *Am. Psychol.* **40,** 385–398 (1985).

48. Squire, L. R. & Zola, S. M. Structure and function of declarative and nondeclarative memory systems. *Proc. Natl. Acad. Sci.* **93,** 13515–13522 (1996).

49. Züst, M. A. *et al.* Hippocampus Is Place of Interaction between Unconscious and Conscious Memories. *PLoS ONE* **10,** e0122459 (2015).

50. Duss, S. B. *et al.* Unconscious relational encoding depends on hippocampus. *Brain* 3355–3370 (2014). doi:10.1093/brain/awu270

51. Hasselmo, M. E. Neuromodulation: acetylcholine and memory consolidation. *Trends Cogn. Sci.* **3,** 351–359 (1999).

52. Wagner, U., Gais, S., Haider, H., Verleger, R. & Born, J. Sleep inspires insight. *Nature* **427,** 352–355 (2004).

53. Born, J. & Feld, G. B. Sleep to Upscale, Sleep to Downscale: Balancing Homeostasis and Plasticity. *Neuron* **75,** 933–935 (2012).

54. Grosmark, A. D., Mizuseki, K., Pastalkova, E., Diba, K. & Buzsáki, G. REM Sleep Reorganizes Hippocampal Excitability. *Neuron* **75,** 1001–1007 (2012).
55. Ruch, S., Züst, M. A. & Henke, K. Subliminal messages exert long-term effects on decision-making. *Neurosci. Conscious.* **2016,** niw013 (2016).
56. Reber, T. P. & Henke, K. Rapid formation and flexible expression of memories of subliminal word pairs. *Conscious. Res.* **2,** 343 (2011).
57. Iber, C., Ancoli-Israel, S., Chesson, A. & Quan, S. F. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. (American Academy of Sleep Medicine, 2007).
58. Ngo, H.-V. V., Martinetz, T., Born, J. & Mölle, M. Auditory Closed-Loop Stimulation of the Sleep Slow Oscillation Enhances Memory. *Neuron* **78,** 545–553 (2013).
59. Renard, Y. *et al.* OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain–Computer Interfaces in Real and Virtual Environments. *Presence Teleoperators Virtual Environ.* **19,** 35–53 (2010).
60. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci.* **2011,** e156869 (2010).
61. Wang, J. *et al.* Arterial transit time imaging with flow encoding arterial spin tagging (FEAST). *Magn. Reson. Med.* **50,** 599–607 (2003).
62. Aguirre, G. K., Detre, J. A., Zarahn, E. & Alsop, D. C. Experimental Design and the Relative Sensitivity of BOLD and Perfusion fMRI. *NeuroImage* **15,** 488–500 (2002).
63. Tzourio-Mazoyer, N. *et al.* Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *NeuroImage* **15,** 273–289 (2002).

Online Methods

Participants

Seventy-six physically and mentally healthy native German speakers with unimpaired hearing ability and no prior experience with our research were recruited. For female participants undergoing fMRI, pregnancy was excluded using a urine-based test. Participants were kept naïve with regard to auditory stimulation during sleep to prevent potentially sleep-harming anticipation of upcoming stimuli and to foster incidental learning during sleep. Written and verbal semi-informed consent was obtained. Participants were fully debriefed after test. Participation was rewarded with 100 CHF (≈100 USD). The study protocol was approved by the cantonal ethics committee Bern. Thirty-five participants had to be excluded post-hoc due to failure to reach sufficiently stable deep sleep, technical problems with data acquisition or stimulus presentation, or because of clinical emergencies at the MR-center. Fourty-one participants (age 19-32, M ± SD = 23.3 ± 3.5; 30 (73%) female) were analyzed, 15 of which (age 19-31, M ± SD = 23.9 ± 3.6; 12 (80%) female) were measured with fMRI at test (=fMRI subgroup). The other 26 participants (age 19-32, M ± SD = 23.0 ± 3.4; 18 (69%) female) were not measured with fMRI (=behavioral subgroup). A power analysis using an expected effect of 4-5% and a SD of 10% (values based on prior research on unconscious associative memory[55,56]) at a power of .8 revealed a reasonable sample size be 32-49.

Stimuli

Foreign words consisted of ninety-six pronounceable two-syllabic pseudowords that we created using German and Dutch syllables. Pseudowords were divided into 4 lists of 24 items. Lists were balanced for word length, pronounceability, perceived concreteness and perceived animateness as rated by five raters. Seventy-two German object words were divided into 3 lists of 24. Lists were balanced for word length, pronounceability, concreteness, and animateness (obtained from 5 raters), as well as log word frequency (collected from http://corpora.uni-leipzig.de/). Half the object words in each list were smaller than a shoebox, the other half distinctively larger. Audio recordings of all stimuli were spoken by a

female speaker in a calm and neutral voice and were recorded in an acoustically isolated room. Recordings were manually processed to attenuate salient features (plosive and sibilant sounds) and were compressed and normalized. Foreign-to-German word translations were randomly selected for each participant to avoid systematic stereotypical translations. This randomization was carried out list-wise, e.g. German list 1 was always associated with foreign list 1, but exact translations varied within list. Of the thus created list-pairs 1-3, two were selected to be presented during sleep, counterbalanced across participants. At test, the two sleep-played foreign lists served as cued-recall items, while the other two foreign lists were presented as new items. This allowed us to collect baseline size judgments (=dependent variable) for each foreign word when presented as new word. These judgments would not be based on sleep-learning, but solely on stimulus properties. We used these ratings to exclude the two stereotypically largest and the two stereotypically smallest foreign words of each list from analysis. This resulted in a total maximum of 40 sleep-played stimuli being analyzed per participant. The fourth foreign list was never presented during sleep and had no German translations. It served as filler-list at test to ensure participants encountered an equal number of sleep-played and new items. Stimulus presentation and response logging was carried out using the software Presentation® (v17.2, Neurobehavioral Systems, http://www.neurobs.com).

Procedure

Participants were asked to keep a regular sleep schedule the 3rd- and 2nd-to-last nights before the experiment. In the night before testing, sleep was restricted to four hours to increase sleep propensity during the experiment. Compliance to this regimen was controlled by sleep diary and direct contact. On the day of the experiment, participants arrived at the lab between noon and 1 pm, gave written informed consent and were equipped with EEG and in-ear headphones.

We then administered an auditory word-identification task to determine the optimal individual intensity for stimulation during sleep. The task required participants to identify spoken German number-words "one" through "four" that were presented with varying intensity while the same constant Brownian noise was playing that was used to reduce stimulus salience during sleep. Intensity of number-words varied randomly between seven levels ranging from clearly audible (signal-to-noise ratio SNR = 0.25) to clearly inaudible (SNR = 0.001). The optimal individual word identification threshold was defined as the lowest SNR at which participants could recognize 50% of stimuli. This threshold was chosen as target SNR (SNR_{target}) for word presentation during sleep (median SNR_{target} = .02). At this intensity, vocabulary was comprehensible but unobtrusive.

Participants were then asked to take a nap in an electrically and acoustically shielded EEG-cabin. Brownian noise was quietly presented at < 50 dB(A) SPL and faded to 65-74 dB(A) SPL as soon as deep slow wave sleep (SWS) was imminent, i.e. late intermediate sleep (S2) with visible delta activity not yet reaching the criterion for SWS of peak-to-peak amplitude > 75 µV[57]. Concurrently, rhythmic presentation of randomly paired spoken numbers "one" through "four" was initiated to habituate participants to stimulation. Initial stimulus intensity was SNR = 0.001 (inaudible) and was increased to the predefined target intensity (SNR_{target}) within about one minute. Once intensity was at SNR_{target}, presentation of German-foreign pairs was initiated. Each pair was repeated four times in sequence. The foreign word was presented first in the first iteration, and switched position with the German word with each repetition. Stimulus onset asynchrony (SOA) for single words in a pair was 1075 ms, corresponding to the expected peak-to-peak interval of slow waves. Rhythmic auditory stimulation with this interval can entrain slow waves which benefits memory consolidation[58]. The SOA for word-pairs was 4 × 1075 ms

which allowed for a 3 s time window of interest to analyze electrical brain response to stimulation with a 1s pre-stimulus baseline. The presentation schema was as follows: F_1-G_1-[ ]-[ ]-G_1-F_1-[ ]-[ ]-F_1-G_1-[ ]-[ ]-G_1-F_1-[ ]-[ ]-F_2-G_2-[ ]-[ ]- … -G_{48}-F_{48}-[ ]-[ ]; where F_n is the nth foreign word, G_n is the nth German word, and [ ] are 1075 ms pauses. Order of occurrence of German-foreign pairs was randomized for each participant. Presentation of all vocabulary took less than 15 minutes in total. Whenever the EEG showed signs of arousal or waking, stimulus presentation was interrupted. If participants regained deep sleep, stimulation was resumed following the same habituation procedure as described above. If participants were unable to enter deep sleep after 2 h, participants were wakened and were tested on their respective subset of sleep-played stimuli. Subjects with less than 20 valid sleep-played items were excluded from analysis. In the reported sample (N = 41), 36.2 ± 5.1 (*M ± SD*) stimuli were analyzed (taking into account exclusion of stereotypical stimuli, cf. section "stimuli", and stimuli associated with arousals).

Once all stimuli were presented, participants slept for another 15 minutes and were then woken up. Participants were given 15-30 minutes to recover from sleep inertia before their memory for sleep-played memory was tested. In the fMRI subgroup, participants were escorted to the MR center. Participants in the behavioral subgroup stayed in the EEG-cabin where they had slept.

The shoebox task used to assess implicit memory for sleep-learned vocabulary was introduced as test that "measures intuitive understanding of a foreign language". Participants were told that foreign words stood for objects of varying size and were asked to intuitively decide whether each word describes a small or large object. The criterion was a shoebox, i.e. participants indicated by key-press whether they felt foreign words described an object that fits into a shoebox or not. Each word was simultaneously presented via headphones and written on screen. Responses to sleep-played foreign words were scored as correct if they matched the size of the sleep-learned German translation (e.g. classifying "tofer" as object that does not fit into a shoebox if it was translated with "house" during sleep). Ninety-six foreign words were presented, up to 48 of which were sleep-played, ordered randomly. In the fMRI subgroup, we additionally presented 48 baseline trials consisting of numbers "one" and "two". Participants had to indicate whether the smaller ("one") or larger ("two") number was presented. This condition is not discussed here. The presentation schema was: *fixation* (250 ms) – *foreign word* (self-paced, max 6 s for behavioral subgroup; fixed & jittered 5-7 s for fMRI subgroup) – *blank* (1 s). This constitutes a rapid-event-related design in the fMRI subgroup.

At the end of the experiment, participants were asked whether they noticed anything unusual during sleep, and then specifically whether they heard any numbers or words during sleep. No participant in the analyzed sample reported having noticed stimulus presentation during sleep. Finally, participants were fully debriefed and compensated.

Stimulus inclusion/exclusion criteria
Only German-foreign word pairs that were either presented during deep sleep or late stage S2 (with visible slow waves below criterion for deep sleep of peak-to-peak amplitude of 75 μV) were analyzed. Furthermore, pairs were excluded from analysis if they contained stereotypical foreign words (see stimuli), or if participants failed to respond in the shoebox task. Participants with less than 20 pairs remaining were excluded from analysis. Applying all inclusion and exclusion criteria, 427 (18.92%) of 1913 sleep-played pairs were excluded in the reported sample, 64 of which due to inappropriate sleep stage.

EEG/Polysomnography

EEG during sleep was recorded with 64 channels in a customized 10-20 montage using BrainCap MR BP-03010MR with "Fast'n Easy" electrodes and two BrainAmp DC, MR plus 32 channel amplifiers by Brain Products (http://www.brainproducts.com). Two channels were used to assess eye movements and one to assess muscle tonus at the chin, the remaining electrodes were used to measure brain responses. Sampling rate was 500 Hz, impedances were kept below 20 kΩ. Recordings were made with BrainVision Recorder (http://www.brainproducts.com). On-line polysomnographic visualization according to American Academy of Sleep Medicine (AASM) guidelines[57] was achieved by a custom software interface created in the OpenVIBE environment[59]. Two raters who were blinded to stimulus presentation performed offline sleep-scorings according to AASM guidelines. Where raters disagreed, the more conservative (i.e. less deep) score was used. EEG data processing was achieved with EEGLAB (http://sccn.ucsd.edu/eeglab/). Data were low-pass filtered at 100 Hz. Artifacts were manually rejected.

Neuroimaging

MRI was achieved with a Siemens MAGNETOM Trio (3 Tesla) equipped with a standard 12-channel head coil. Anatomical reference images were acquired using a T_1–weighted magnetization-prepared rapid acquisition gradient echo (MP-RAGE) sequence: 176 sagittal slices (no inter-slice gaps), 256 × 256 acquisition matrix, isotropic voxel size = 1 mm^3, repetition time (TR) = 2530ms, echo time (TE) = 2.2ms, flip angle (FA) = 9°. Functional volumes were acquired using a T_2*–weighted single-shot echo-planar (EPI) sequence recording 580 volumes of 32 interleaved (even first) transversal slices, longitudinally aligned with the hippocampal formation, with a distance factor of 15% and phase oversampling of 20%, 64 × 64 acquisition matrix, isotropic voxel size = 3 mm^3, TR = 2060 ms, TE = 30 ms, FA = 90°. Additionally, a pseudo-continuous arterial spin-labelling (pCASL) sequence was acquired: 120 volumes of 20 transversal slices oriented perpendicular to the carotid arteries, 128 × 64 acquisition matrix, voxel size = 3.6 × 1.8 × 6mm, TR = 4000 ms, TE = 18 ms, FA = 25°, post label delay = 1250 ms. This sequence was included to gain a mean value of global cerebral blood flow rate of each participant to be used as nuisance covariate in statistical models. Data processing and statistics were calculated with SPM8 (Wellcome Department of Cognitive Neurology, London, UK). Functional volumes were slice-time corrected to slice 16/32, realigned to the first volume, and co-registered to the anatomical reference image. The anatomical reference image of each participant was normalized to the T_1 template provided by SPM8 and the resulting deformation field was applied to the functional volumes. Validity of results was manually verified. Functional volumes were 3-dimensionally smoothed with an isotropic Gaussian kernel with a full width at half maximum of 8 mm.

Statistical Analyses

Mean accuracy scores (shoebox task) were zero-aligned by subtracting chance level (-50%) and were tested as intercept in a one-way ANOVA with subgroup as factor and stimulus intensity (SNR_{target}) as covariate. This test is equivalent to a one-sample t-test but corrects for nuisance variables.

To assess electroencephalographic responses to stimulation during sleep, we extracted event-related potentials (ERPs) and spectral perturbations (ERSPs) time-locked to the first word of each presentation of a word-pair. Analyses were performed with the Matlab® toolbox FieldTrip[60]. ERPs were analyzed for 3 s following stimulus onset and were baseline-corrected by subtracting the mean of the 1 s pre-stimulus window. ERSPs were obtained by performing Morlet wavelet transforms with 3 cycles in length at the lowest frequency of 2 Hz, increasing linearly with frequency to 12.7 cycles at the highest frequency of 21.5 Hz. ERSPs were averaged across trials and were baseline-corrected (dB) for each subject at each

frequency using a 1 s pre-stimulus as baseline. To quantify the general impact of stimulation during sleep, we contrasted the average ERPs and ERSPs with the pre-stimulus-baseline. We averaged mean voltage and mean spectral power across time-windows and time-frequency-windows of interest. Electrodes showing significant deviation from baseline in these windows were identified using cluster-based permutation tests (5000 permutations) to correct for multiple comparisons[60].

To assess the EEG signature of successful encoding during sleep, we contrasted mean sleep ERPs and ERSPs for word-pairs associated with correct vs. incorrect decisions in the shoebox task. We extracted mean ERP voltage for each subject for the early (800 - 1100 ms) and late (1200 - 1500 ms) induced 2nd up-state from electrodes F1, F2, Fz, FC1, FC2, FCz. Voltage was analyzed using a repeated measures ANOVA with the factors time-window (early vs. late peak) and accuracy (correct vs. incorrect). We further extracted mean spectral power in the theta (5-7 Hz[28]) and fast spindle (13-15 Hz[4]) frequency bands around 800-1100 ms. We used cluster-based permutation statistics (5000 permutations) to identify electrodes that distinguished word-pairs associated with correct vs. incorrect decisions in these frequency bands.

Phase analysis was performed by extracting phase angles at the slow wave frequency of 0.8 Hz[16] over frontal electrodes for each single trial, i.e. each of the 5708 presentations of a word-pair. We averaged the raw signal over electrodes F1, F2, Fz, FC1, FC2, FCz, down-sampled to 100 Hz, applied a low-pass filter at 4 Hz, and performed continuous Morlet wavelet transformations with 2 cycles in length at 0.8 Hz. This yielded phase angles between $-\pi$ and $+\pi$, where 0 reflects a slow wave peak and $+/-\pi$ reflects a slow wave trough. Phase angles were extracted at 20 ms intervals from 0.3 s before to 2.1 s after stimulus onset. For each time-point of the resulting time-phase decomposition, we grouped all trials into twelve bins (30° in width) according to their current phase. We calculated the percentage of trials in each time-phase-bin that led to a correct response in the shoebox task. If retention of German-foreign word pairs was associated with the presence of a specific slow-wave phase at a specific time during encoding, percentage of trials leading to correct shoebox-decisions should be elevated in the respective bins. We therefore searched for large clusters of adjacent time-phase bins in which more than 54% trials lead to a correct decision in the shoebox task. The null hypothesis H_0 assumes that phase is not associated with shoebox accuracy. Under H_0, percentage of correct trials should thus randomly vary around the average of 52.42%, yielding small clusters. An empirical cluster was accepted as significant if its size (i.e. the number of adjacent time-phase bins) exceeded the size of 95% of maximal clusters found in randomized data sets. A total of 5000 random data sets were drawn by shuffling the information whether a trial was associated with a correct or an incorrect response in the shoebox task.

To count how many repetitions of each sleep-played word pair were associated with a slow-wave peak between 800-1000 ms following stimulus onset, we assessed for each repetition whether the extracted 0.8 Hz phase reached an angle close 0 rad (reflecting a wave-peak) in this time window. A repetition was said to hit a slow-wave-peak if the phase was within +/-2.5% around 0 rad (within +/-9°) between 800 - 1000 ms. For each participant, we computed the mean accuracy in the shoebox task separately for sleep-played trials that were never, once, and repeatedly associated with a peak.

To analyze rapid event-related fMRI data on subject level, brain responses to correct and incorrect trials in the shoebox task were modelled with stick functions convolved with a canonical hemodynamic response function (HRF) provided by SPM8. Six movement parameters estimated during realignment of functional volumes were entered as nuisance covariates to account for error variance induced by

moving voxel positions. For each sleep-played German-foreign pair, the mean evoked slow-wave phase angle between 700 and 1100 ms, averaged across the 4 repetitions during sleep, was included in the model as parametric modulator of the responses to correct and incorrect trials. Note that mean phase was significantly closer to 0 rad (i.e. a peak) for correct vs. incorrect trials ($t(40) = 3.32$, $p = .002$). This parametric modulator was scaled in a way that the expected phase was 1 (i.e. a slow wave peak for correct pairs and a slow wave trough for incorrect pairs) and the inverse phase -1 (i.e. a slow wave trough for correct pairs and a slow wave peak for incorrect pairs). Group level analyses were computed based on subject level contrasts using a within-subject one-way ANOVA. All group level models included mean global cerebral blood flow as nuisance covariate[61,62]. To find brain areas associated with successful retrieval of sleep-played translations, we contrasted the slow wave phase-informed correct vs. incorrect trials, i.e. the interaction terms of the parametric modulation analysis: (correct * slow wave peak) > (incorrect * slow wave trough). For a between subject correlational analysis we regressed mean accuracy in the shoebox task onto the unmodulated contrast (omitting phase information) comparing correct vs. incorrect trials to find neuronal activity which scaled with individual performance. The primary area of interest was the hippocampal formation. Therefore, we masked brain activity with the automated anatomical labelling (AAL) template[63] for the bilateral hippocampus and applied a small volume correction within that region accepting voxels as significant at $p < .05$, family-wise error (FWE) corrected. Inference on whole brain level was carried out on an uncorrected voxel-wise height threshold of $p < .001$ with a cluster extent threshold of 10 voxels.