

# Current Biology

## Larger capacity for unconscious versus conscious episodic memory

### Highlights

- Humans draw inferences when encoding episodes with and without consciousness
- Memory capacity is larger for unconscious than conscious episodic memory
- Hippocampus supports encoding and retrieval with and without consciousness
- Conscious versus unconscious encoding/retrieval recruits additional brain regions

### Authors

Else Schneider, Marc Alain Züst, Sergej Wuethrich, ..., Roland Wiest, Simon Ruch, Katharina Henke

### Correspondence

henke@psy.unibe.ch

### In brief

Humans draw inferences when encoding episodes with and without consciousness. Schneider et al. find a larger memory capacity for unconscious than conscious episodic encoding. Hippocampus supports encoding and retrieval at both consciousness levels. Conscious processing spawned steeper and broader activity increases and recruited additional regions.

## Article

# Larger capacity for unconscious versus conscious episodic memory

Else Schneider,<sup>1,4</sup> Marc Alain Züst,<sup>1,2,4</sup> Sergej Wuethrich,<sup>1</sup> Flavio Schmidig,<sup>1</sup> Stefan Klöppel,<sup>2</sup> Roland Wiest,<sup>3</sup> Simon Ruch,<sup>1</sup> and Katharina Henke<sup>1,5,6,\*</sup>

<sup>1</sup>Institute of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland

<sup>2</sup>University Hospital of Old Age Psychiatry and Psychotherapy, University of Bern, Bolligenstraße 111, 3000 Bern, Switzerland

<sup>3</sup>Institute of Diagnostic and Interventional Neuroradiology, University Hospital Bern, Freiburgstrasse 18, 3010 Bern, Switzerland

<sup>4</sup>These authors contributed equally

<sup>5</sup>Twitter: @HenkeLab

<sup>6</sup>Lead contact

\*Correspondence: [henke@psy.unibe.ch](mailto:henke@psy.unibe.ch)

<https://doi.org/10.1016/j.cub.2021.06.012>

## SUMMARY

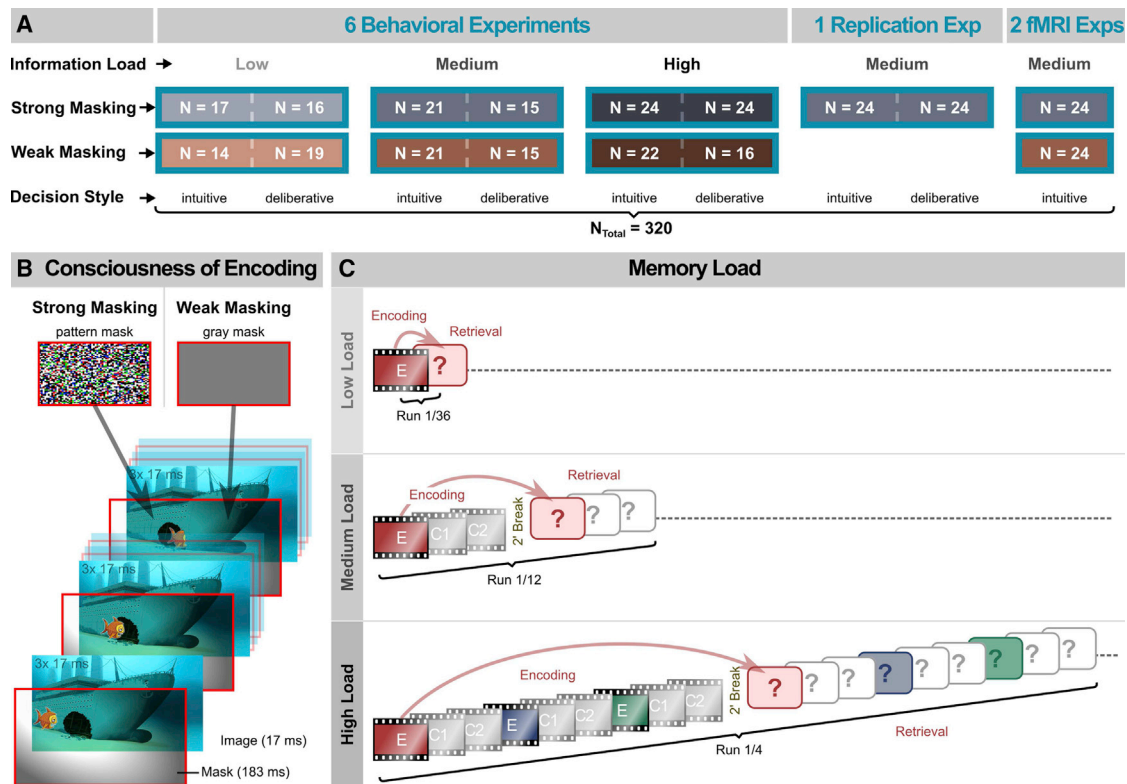
Episodic memory is the memory for experienced events. A peak competence of episodic memory is the mental combination of events to infer commonalities. Inferring commonalities may proceed with and without consciousness of events. Yet what distinguishes conscious from unconscious inference? This question inspired nine experiments that featured strongly and weakly masked cartoon clips presented for unconscious and conscious inference. Each clip featured a scene with a visually impenetrable hiding place. Five animals crossed the scene one-by-one consecutively. One animal trajectory represented one event. The animals moved through the hiding place, where they might linger or not. The participants' task was to observe the animals' entrances and exits to maintain a mental record of which animals hid simultaneously. We manipulated information load to explore capacity limits. Memory of inferences was tested immediately, 3.5 or 6 min following encoding. The participants retrieved inferences well when encoding was conscious. When encoding was unconscious, the participants needed to respond intuitively. Only habitually intuitive decision makers exhibited a significant delayed retrieval of inferences drawn unconsciously. Their unconscious retrieval performance did not drop significantly with increasing information load, while conscious retrieval performance dropped significantly. A working memory network, including hippocampus, was activated during both conscious and unconscious inference and correlated with retrieval success. An episodic retrieval network, including hippocampus, was activated during both conscious and unconscious retrieval of inferences and correlated with retrieval success. Only conscious encoding/retrieval recruited additional brain regions outside these networks. Hence, levels of consciousness influenced the memories' behavioral impact, memory capacity, and the neural representational code.

## INTRODUCTION

To make sense of the world, we need to remember what happened where and when. We also need to combine experiences using inferential reasoning<sup>1</sup> to detect commonalities or relations between them. Remembering experienced episodes and inferential reasoning require episodic memory and the hippocampus.<sup>2</sup> For example, we can use episodic memory while sitting outside a cave to infer which tourists visit the cave simultaneously, based on our mental record of the tourists' entrance and exit times. The mental computations that allow drawing inferences are (1) one-shot encoding (a person enters/exits the cave only once), (2) forming individual pairwise associations (tourist A is the first to enter the cave, tourist B is the second ...), and (3) identifying temporal overlaps of cave visitors (noticing which tourists lingered in the cave simultaneously).<sup>3–8</sup> Cohen and Eichenbaum<sup>3</sup> suggested that episodic memory entails both the ability to encode individual associations between items

and the capacity to express memories flexibly through inference, as in the cave example. Findings indicate that drawing inferences by forming associations indirectly (not based on sensory input) requires both episodic memory and working memory.<sup>9–15</sup> Here, we probe the formation and delayed retrieval of inferences to assess the role of consciousness in episodic memory and working memory.

Encoding events and drawing inferences appear associated with consciousness. However, consciousness of event perception is a debated precondition for episodic encoding,<sup>3,4,16–19</sup> inferential reasoning,<sup>20</sup> and working memory.<sup>21</sup> According to textbooks, human episodic memory, and hippocampal processing belong to declarative/explicit memory and depend on consciousness.<sup>16,18,19,22</sup> This notion derives from amnesic patients with hippocampal damage, who exhibited impaired conscious episodic learning but preserved unconscious learning, e.g., skill acquisition, priming, and conditioning.<sup>2,17–19</sup> Importantly, investigators probed episodic memory using consciously accessible



**Figure 1. Experimental designs**

(A) The current research consists of three parts. In part one, information load was manipulated in three levels (low, medium, high). In part two, we replicated the experiment with strong masking and a medium information load. In part three, we performed two fMRI experiments—one with strong masking and one with weak masking—and with a medium information load. N, numbers of participants.

(B) Strong and weak masking. For strong masking, masks were filled with colored pixels. For weak masking, masks were uniform gray.

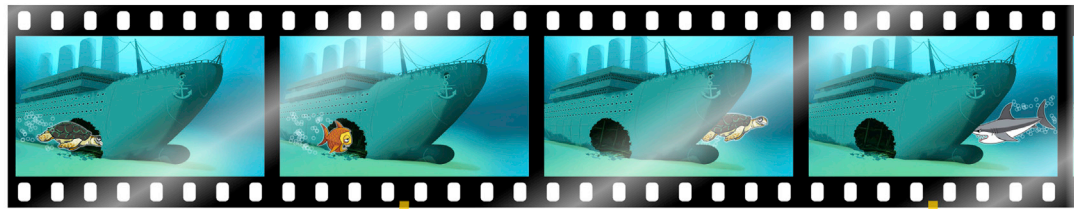
(C) Manipulation of information load and encoding-test interval. The low load condition consisted of 36 short encoding-test runs, the medium load condition consisted of 12 encoding-test runs, and the high load condition consisted of four encoding-test runs. Question marks indicate retrieval testing; each box with a question mark stands for the ten consecutively administered retrieval trials that pertain to one clip. E, experimental condition; C1, control condition 1; C2, control condition 2. See [Video S1](#) for an impression of the masked cartoon clips.

learning material without obviating consciousness to see whether unconscious episodic encoding would also operate in volunteers and would decrease following hippocampal damage. That consciousness is required for episodic and hippocampal encoding is therefore an empirically unfounded belief. Strict tests of this belief necessitate (supra)liminal encoding in participants during deep sleep, coma, or anesthesia or subliminal encoding in wake participants. Experiments using both of these approaches revealed that hippocampal-assisted episodic encoding operates with and without consciousness.<sup>23–28</sup> These findings confirm newer theoretical claims<sup>3,4,8</sup> that the sole premise for episodic encoding through hippocampus is not consciousness but a task that calls upon the core computational competence of the hippocampus, namely, the rapid formation of new and flexible associations. When applying such tasks, unconscious episodic encoding and retrieval was revealed using subliminal word pairs,<sup>23,29,30</sup> subliminal face-word pairs,<sup>24,31–33</sup> and subliminal objects in space.<sup>26</sup> Moreover, participants inferred unconsciously that two ordinarily unrelated words A-C are semantically related based on the preceding subliminal encoding of discontinuously flashed subliminal word pairs A-B and B-C.<sup>34,35</sup> Furthermore, subliminal associative encoding

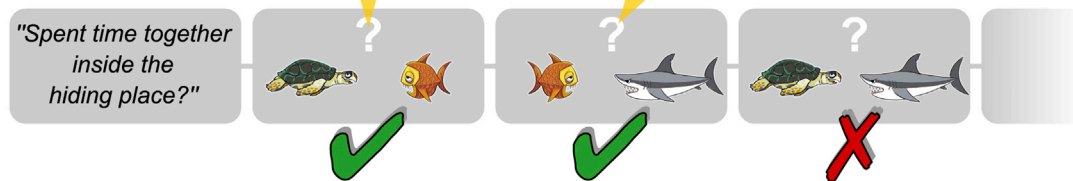
encompassed the encoding of complex visuospatial second-order sequences,<sup>36</sup> the understanding of causal events,<sup>37</sup> and narratives from spatial-temporal patterns.<sup>38</sup> Intriguingly, the hippocampus was not only activated during subliminal associative encoding and unconscious retrieval<sup>24,30,33,35</sup> but hippocampal damage also impeded subliminal associative encoding.<sup>23</sup> Even in the unconscious state of deep sleep did participants form and store into wakefulness lexical-semantic associations unconsciously.<sup>28</sup> These findings question the textbook account of memory<sup>2,17–19</sup> and support the relational memory account<sup>3</sup> and the processing account of memory systems.<sup>4</sup>

The finding of an unconscious besides the known conscious form of episodic memory raises the question of how the two differ. Here, we address this question by contrasting conscious and unconscious encoding/retrieval of cartoon clips. We created a memory task that requires the continuous updating of information and the drawing of temporal relational inferences while watching a clip ([Figure 1](#)). We presented strongly and weakly masked clips for unconscious and conscious encoding, respectively. We informed participants of subliminal clips before experimentation to match instructions between consciousness levels. Each clip presented a scene with a visually impenetrable hiding

## Encoding



## Retrieval



**Figure 2. A decision task served as a retrieval test of previously drawn inferences**

At encoding, the participants' task was to observe the animals' entrances and exits to maintain a mental record of which animals resided inside the hiding place simultaneously. For retrieval testing, participants were presented with the unmasked images of a clip's hiding place and two animals for conscious inspection to decide based on conscious knowledge (following weak masking) or based on intuition (following strong masking) whether the two animals lingered simultaneously inside the hiding place or not. See [Video S1](#) for an impression of the masked cartoon clips.

place and five animals that entered and left the hiding place consecutively ([Video S1](#)). All animals moved through the hiding place and left it immediately or lingered inside potentially meeting other animals. The participants' task was to encode the time points of the animals' entrances and exits to infer which animals resided in the hiding place simultaneously. Drawing inferences while encoding animal trajectories requires Piagetian object permanence<sup>39</sup> and conscious/unconscious working memory.<sup>40,41</sup> Outputs of working memory needed to be stored long term. We manipulated information load by presenting either one single clip for encoding and immediate retrieval, three clips for encoding and delayed retrieval, or nine clips for encoding and very delayed retrieval ([Figure 1](#)).

This experimental design allows testing several hypotheses. Given the alignment of task demands, instructions, and learning materials between consciousness levels, we anticipated brain activations in a common working memory network during conscious and unconscious inferential reasoning while watching clips and in an episodic retrieval network during the conscious and unconscious retrieval of inferences. We hypothesized the hippocampus to mediate both the formation and the retrieval of inferences.<sup>42–45</sup> Because visual input is stronger during supraliminal versus subliminal encoding, we expected steeper local maxima in brain activation and broader activation clusters during conscious versus unconscious processing.<sup>23,35,46–50</sup> We also hypothesized that unconscious retrieval performance would remain stable with increasing information load,<sup>23,24,26,29–35</sup> while conscious retrieval performance would drop. Finally, we measured the participants' habitual decision style—intuitive versus deliberative—to account for the possibility that habitually intuitive decision makers would outperform deliberative decision makers in the unconscious condition that required intuitive decisions at test.<sup>26,51</sup> Directly comparing conscious with

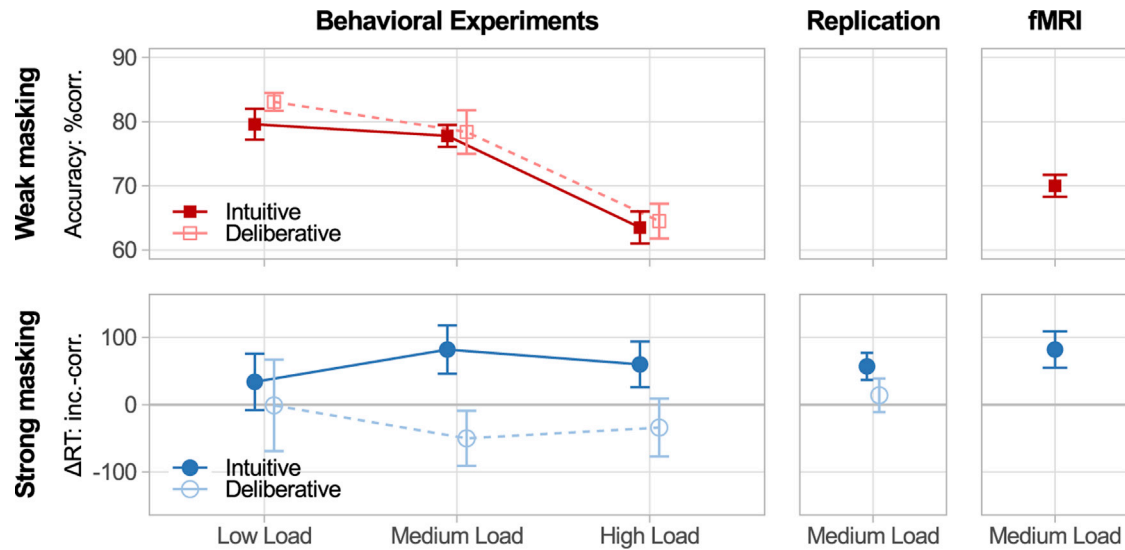
unconscious processing in episodic memory will advance the relational<sup>3</sup> and processing account of memory systems<sup>4</sup> because results reveal how conscious differs from unconscious episodic processing.

## RESULTS

### Overview of results in the three parts of this research

The current research consists of three parts that all include the same experimental task ([Figure 1](#)): drawing temporal relational inferences while watching cartoon clips and retrieving these inferences at test. The first part of this research features six behavioral experiments with manipulated information load, in which cartoon clips were either presented strongly or weakly masked for unconscious or conscious encoding with low/medium/high information load. The second part of this research features a behavioral replication experiment with strong masking and a medium information load. The third part of this research features two experiments with functional magnetic resonance imaging (fMRI) and a medium information load. One experiment included strongly masked cartoon clips and the other weakly masked cartoon clips ([Figure 1](#)). [Table S1](#) gives values for two-tailed one-sample t tests computed against chance level (chance level of accuracy: 0.50; chance level of  $\Delta RT$  correct – incorrect: 0.00) for each condition and each experiment.

Following experiments featuring strong masking, we carried out an objective test of clip awareness to ascertain the level of (un)consciousness of clip perception. In addition to this objective awareness measure, participants rated their conscious awareness of each cartoon clip following its presentation using the perceptual awareness scale (PAS;<sup>52</sup>). PAS results are presented at the end of the results section. The results of the objective awareness tests are presented in the [STAR Methods](#).



**Figure 3. Retrieval performance in all experiments**

Top panel: conscious retrieval following the encoding of weakly masked clips (mean % correct responses; chance level = 50%). Bottom panel: unconscious retrieval following the encoding of strongly masked clips (difference in reaction times,  $\Delta$ RT: incorrect minus correct retrieval responses). Error bars represent one standard error of the mean. See also [Figures S1–S3](#) as well as [Tables S1](#) and [S6](#).

### First part of this research: Six behavioral experiments with manipulated information load

Inferences drawn during clip encoding needed to be retained for retrieval testing. A retrieval test trial consisted of the unmasked presentation of a clip's hiding place plus the unmasked images of two animals that had featured in the respective clip. There were ten retrieval trials per clip. We used a direct, explicit retrieval instruction in both consciousness conditions. The participants' task was to decide whether the two animals had hidden simultaneously or not in the presented hiding place (responses by manual button press). Following strong masking, participants were asked to rely on intuition. The average percentage of retrieval trials that call for a "simultaneous" response was 50%. We expected longer response latencies for incorrect versus correct retrieval responses in both consciousness conditions. This reaction time difference can reveal the presence of unconscious memory traces even when response accuracy is at chance level.<sup>24,31,53</sup> Accuracy (percentage of correct responses) and the reaction time (RT) difference between the median RT of incorrect and correct retrieval responses served as dependent variables. We computed t tests to obtain overall results ([Table S1](#)) and computed analyses of variance (ANOVA) with the between-subjects factors information load (IL; low load, medium load, high load) and decision style (DS; intuitive/deliberative). [Figure 3](#) provides an overview of results and [Figures S1](#) and [S2](#) distributions of retrieval accuracy and RT differences.

### Experiments with strong masking for unconscious encoding

#### Accuracy

The overall t test indicates that participants' retrieval performance did not exceed chance level ( $M_{Acc} = 50.20\%$ ,  $SE_{Acc} = 0.43\%$ ;  $t(116) = 0.463$ ,  $p = 0.644$ ,  $d_z = 0.042$ ). The ANOVA

indicates that IL and DS failed to influence retrieval performance significantly:  $F_{IL}(2, 111) = 0.33$ ,  $p = 0.721$ ;  $F_{DS}(1, 111) = 1.05$ ,  $p = 0.307$ ;  $F_{IL \times DS}(2, 111) = 0.27$ ,  $p = 0.766$ . Bayes analyses confirmed that retrieval performance was at chance level. In all IL conditions and for both DS groups, Bayes factors were  $<0.30$ , which is substantial evidence for the  $H_0$  assumption of chance level performance. Only deliberative decision makers in the high-load condition had a  $BF > 1/3$  ( $BF = 1.26$ ) suggesting that the data could not properly discriminate between  $H_0$  and  $H_1$ . BFs were calculated with the  $R^{54}$  function provided by Baguley<sup>55</sup> using a half-normal prior distribution with a mode of 0% (reflecting chance-level accuracy) and a standard deviation of 5% (expected above chance accuracy based on previous research).<sup>56</sup> Chance-level retrieval performance attests to the results of the objective awareness test (reported in the [STAR Methods](#)) that participants were unable to process the strongly masked clips using conscious awareness ([Tables S1](#) and [S6](#); [Figure S1](#)).

#### Reaction time difference

The overall t test shows that participants' reaction times did not distinguish significantly between correct and incorrect retrieval responses ( $M_{\Delta RT} = 19$  ms,  $SE = 18$  ms;  $t(116) = 1.032$ ,  $p = 0.304$ ,  $d_z = 0.095$ ). The ANOVA indicates that IL had no significant influence on the RT difference ( $F_{IL}(2, 111) < 0.001$ ,  $p = 0.996$ ,  $\eta_p^2 < 0.001$ ) but DS had ( $F_{DS}(1, 111) = 5.75$ ,  $p = 0.018$ ,  $\eta_p^2 = 0.049$ ). Namely, intuitive decision makers gave correct retrieval responses significantly faster than incorrect retrieval responses ( $M_{\Delta RT} = 60$  ms,  $SE = 21$  ms;  $t(61) = 2.869$ ,  $p = 0.006$ ,  $d_z = 0.36$ ). Importantly, this result was not significantly modulated by IL ( $F_{IL}(2, 59) = 0.38$ ,  $p = 0.685$ ) ([Figures 3](#) and [S2](#)). The deliberative decision makers did not give correct retrieval responses significantly faster than incorrect retrieval responses ( $M_{\Delta RT} = -29$  ms,  $SE = 29$  ms;  $t(54) = -0.989$ ,  $p = 0.327$ ,  $d_z = -0.133$ ).

*No association between accuracy and reaction time difference*

Retrieval accuracy was not significantly related to the RT difference between incorrect and correct retrieval responses ( $r(115) = 0.001$ ,  $p = 0.989$ ; both measures were centered by the mean of the respective load condition). Furthermore, neither load nor decision style modulated the relation between accuracy and RT difference (all  $p > 0.18$  in a linear model with RT difference as dependent variable and with the predictors accuracy, load, DS, and all interaction terms; only DS was significant:  $F_{DS}(1,105) = 6.27$ ,  $p = 0.014$ ). Hence, a potential residual stimulus awareness of participants with high accuracy scores did not significantly contribute to the reported RT difference (the measure of implicit memory), neither in intuitive nor in deliberative decision makers.

#### Experiments with weak masking for conscious encoding Accuracy

The overall t test shows that participants' retrieval performance was well above chance level ( $M_{Acc} = 74.12\%$ ,  $SE_{Acc} = 1.21\%$ , chance level 50%;  $t(106) = 19.994$ ,  $p < 0.001$ ,  $d = 5.940$ ). Notably, the ANOVA indicates that IL affected participants' retrieval performance:  $F_{IL}(2, 101) = 30.62$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.378$ . Namely, participants performed significantly worse in the high load (HL) condition ( $M_{Acc} = 63.88\%$ ) compared to the low load (LL) condition ( $M_{Acc} = 81.63\%$ ;  $t(69) = -7.772$ ,  $p < 0.001$ ,  $d_S = -1.849$ ) and also compared to the medium load (ML) condition ( $M_{Acc} = 78.06\%$ ;  $t(72) = -5.671$ ,  $p < 0.001$ ,  $d_S = -1.319$ ) (Figures 3 and S1). The comparison between the ML and the LL conditions yielded no significant result:  $t(67) = 1.622$ ,  $p = 0.110$ ,  $d_S = 0.391$ . The ANOVA suggests that DS had no significant influence on retrieval performance ( $F_{DS}(1,101) = 0.77$ ,  $p = 0.383$ ), and DS did not interact significantly with IL ( $F_{IL \times DS}(2,101) = 0.21$ ,  $p = 0.808$ ).

#### Reaction time difference

The overall t test reveals that participants gave correct retrieval responses significantly faster than incorrect retrieval responses ( $M_{\Delta RT} = 556$  ms,  $SE = 62$  ms;  $t(106) = 8.959$ ,  $p < 0.001$ ,  $d_Z = 0.866$ ). The ANOVA indicates that IL tended to influence the measure of the RT difference:  $F_{IL}(2,101) = 2.99$ ,  $p = 0.055$ ,  $\eta_p^2 = 0.056$ . The greater speed of correct versus incorrect responses diminished when information load was high (LL:  $M_{\Delta RT} = 603$  ms,  $SE = 107$  ms; ML:  $M_{\Delta RT} = 722$  ms,  $SE = 130$  ms; HL:  $M_{\Delta RT} = 358$  ms,  $SE = 74$  ms). The RT difference tended to be smaller in the HL compared to the LL condition ( $t(69) = -1.919$ ,  $p = 0.059$ ,  $d_S = -0.457$ ) and was significantly smaller in the HL compared to the ML condition ( $t(72) = -2.466$ ,  $p = 0.016$ ,  $d_S = -0.574$ ), but not in the ML compared to the LL condition ( $t(67) = 0.697$ ,  $p = 0.488$ ,  $d_S = 0.168$ ) (Figure S2). Decision style exerted no significant influence on the RT difference ( $F_{DS}(1,101) = 0.49$ ,  $p = 0.486$ ) and did not interact significantly with information load ( $F_{IL \times DS}(2,101) = 0.13$ ,  $p = 0.875$ ).

#### Association between accuracy and reaction time difference

Retrieval accuracy predicted significantly reaction time differences between incorrect and correct retrieval responses ( $r(105) = 0.321$ ,  $p < 0.001$ ; both measures were centered by the mean of the respective load condition). The association between accuracy and RT differences was not significantly affected by load or DS (all  $p > 0.20$  in a linear model with RT difference as dependent variable and with the predictors accuracy, load, DS, and all interaction terms; only the main effect accuracy was significant:  $F_{Acc} = 10.98$ ,  $p = 0.001$ ). This suggests that

the magnitude of the difference in reaction times between incorrect and correct responses reliably mirrored individual differences in conscious retrieval performance independently of DS and load.

#### Is a high information load associated with more forgetting when masking is weak (conscious processing) versus strong (unconscious processing)?

We contrasted the two consciousness conditions directly regarding the relative drop in retrieval performance from the medium to the high load condition. We limited this analysis to intuitive decision makers because only these participants showed implicit retrieval effects in the strongly masked conditions. To compute a two-factorial ANOVA on the z-standardized retrieval data, we pooled and z-standardized the intuitive decision makers' retrieval data (percentage of correct responses) acquired following weak masking. We did the same for the intuitive decision makers' retrieval data acquired following strong masking (RT difference between incorrect and correct responses). The ANOVA with the z-standardized values as dependent variable and the independent variables masking (strong versus weak) and IL (low, medium, high) yielded a significant interaction between IL and masking:  $F_{IL \times M}(2,113) = 6.88$ ,  $p = 0.002$ ,  $\eta_p^2 = 0.109$ . Retrieval performance dropped significantly from the ML to the HL condition following weak masking ( $t(41) = -2.198$ ,  $p = 0.034$ ,  $d_S = -0.671$ ), but not following strong masking ( $t(43) = -0.441$ ,  $p = 0.661$ ,  $d_S = -0.132$ ).

#### Second part of this research: Replication experiment with medium information load and strong masking for unconscious encoding

Before running fMRI experiments, we wished to replicate the unconscious retrieval effect with medium load in intuitive decision makers. We ran the medium load version of the experiment with strong masking in new participants that had either a deliberative or intuitive decision style.

#### Accuracy

The overall t test suggests that participants' percentage of correct retrieval responses were not significantly above chance level ( $M_{Acc} = 50.10\%$ ,  $SE_{Acc} = 0.61\%$ ;  $t(47) = 0.172$ ,  $p = 0.864$ ,  $d_Z = 0.025$ ). DS did not influence retrieval accuracy significantly:  $t(46) = 0.799$ ,  $p = 0.428$ ,  $d_S = 0.231$  (Figure S1). Bayes factors for both DS groups were  $< 0.23$ , which is substantial evidence for the null assumption of chance-level performance.

#### Reaction time difference

The overall t test revealed significantly shorter reaction times for correct than incorrect retrieval responses ( $M_{\Delta RT} = 35$  ms,  $SE = 16$  ms;  $t(47) = 2.156$ ,  $p = 0.036$ ,  $d_Z = 0.311$ ). DS did not influence the size of the RT difference significantly ( $t(46) = 1.313$ ,  $p = 0.196$ ,  $d_S = 0.379$ ), although intuitive decision makers ( $M_{\Delta RT} = 56$  ms,  $SE = 20$  ms) presented with a larger RT difference than deliberative decision makers ( $M_{\Delta RT} = 14$  ms,  $SE = 25$  ms). The intuitive decision makers' RT difference was larger than zero ( $t(23) = 2.786$ ,  $p = 0.011$ ,  $d_Z = 0.569$ ), while the deliberative decision makers' RT difference did not differ significantly from zero ( $t(23) = 0.551$ ,  $p = 0.587$ ,  $d_Z = 0.112$ ) (Figures 3 and S2). We computed the Bayes factor to quantify the evidence in favor of the hypothesized RT difference, as observed in the first part of this research with a medium information load, versus the null hypothesis of no RT difference. As recommended by Dienes,<sup>57,58</sup>

we chose a half-normal prior distribution with a mode of 0 and a standard deviation of 82 ms, which corresponds to the effect size of the intuitive decision makers in the initial experiment. Using the Bayes factor function for  $R^{54}$  provided by Baguley,<sup>55</sup> the resulting Bayes factor for the performance of intuitive decision makers was 19.10. This factor speaks in favor of an RT difference for intuitive decision makers as observed in the medium load experiment in the first part of this research. The Bayes factor for deliberative decision makers was 0.473, which is inconclusive but tends to favor the null assumption of no difference between RTs for correct versus incorrect retrieval responses.

#### **No association between accuracy and reaction time difference**

Retrieval accuracy was not related significantly to RT differences between incorrect and correct retrieval responses ( $r(46) = -0.058$ ,  $p = 0.694$ ). Furthermore, DS had no significant influence on the relation between accuracy and RT difference (all  $p > 0.185$  in a linear model with the predictors accuracy, DS, and the interaction term accuracy  $\times$  DS). Hence, a potential residual stimulus awareness of participants with high accuracy scores did not contribute significantly to the reported RT difference (the measure of implicit memory), neither in intuitive nor in deliberative decision makers.

#### **Third part of this research: fMRI experiments with strongly and weakly masked cartoon clips and a medium information load**

We recorded the fMRI BOLD signal during both encoding and retrieval. One fMRI experiment included weakly masked clips for conscious encoding and the other fMRI experiment strongly masked clips for unconscious encoding. Based on previous evidence,<sup>24,59</sup> we expected the hippocampus to be activated during both conscious and unconscious encoding and retrieval. We recruited exclusively intuitive decision makers for both experiments.

#### **Behavioral results**

##### *Experiment with strong masking for unconscious encoding*

The percentage of correct retrieval responses was not better than chance ( $M_{\text{Acc}} = 49.51\%$ ,  $SE = 0.91\%$ ;  $t(23) = -0.540$ ,  $p = 0.595$ ,  $d_z = -0.110$ ) (Figure S1). The Bayes factor was 0.12, which is substantial evidence for the null assumption of chance-level performance. However, participants exhibited significantly shorter reaction times for correct versus incorrect retrieval responses ( $M_{\Delta\text{RT}} = 81$  ms,  $SE = 31$  ms;  $t(23) = 2.600$ ,  $p = 0.016$ ,  $d_z = 0.531$ ) (Figures 3 and S2). We computed the Bayes factor to quantify the evidence in favor of a null effect versus an RT difference of the size observed in the medium load experiment of the first part of this research. We used the same half-normal prior distribution as in part two of this research (mode of 0,  $SD = 82$  ms). The Bayes factor for the obtained effect was 13.92, which speaks in favor of an RT difference as observed in the medium load experiment of the first part of this research.

##### *No association between accuracy and reaction time difference for unconscious retrieval*

Retrieval accuracy was not related to RT differences ( $r(22) = 0.129$ ,  $p = 0.547$ ). Hence, a potential residual stimulus awareness of participants with high accuracy scores did not contribute to the reported RT difference.

##### *Experiment with weak masking for conscious encoding*

The percentage of correct retrieval responses was above chance ( $M_{\text{Acc}} = 75.67\%$ ,  $SE_{\text{Acc}} = 1.75\%$ ;  $t(23) = 14.642$ ,  $p < 0.001$ ,  $d_z = 2.989$ ) (Figures 3 and S1). Participants exhibited significantly shorter reaction times for correct versus incorrect retrieval responses ( $M_{\Delta\text{RT}} = 458$  ms,  $SE = 58$  ms;  $t(23) = 7.870$ ,  $p < 0.001$ ,  $d_z = 1.606$ ) (Figure S2).

##### *Association between accuracy and reaction time difference for conscious retrieval*

Retrieval accuracy predicted the reaction time differences ( $r(22) = 0.531$ ,  $p = 0.008$ ).

#### **Note concerning the retrieval-related reaction time differences following the encoding of strongly masked clips**

To legitimate an interpretation of reaction time differences in terms of unconscious inferential reasoning and unconscious retrieval of inferences, we considered alternative explanations for the reaction time differences, such as biases in the stimulus material and response biases. These additional analyses are reported in the STAR Methods. Results indicate that the difference in reaction times can neither be explained by biases in the stimulus material nor by a participant response bias.

#### **fMRI results: Control conditions**

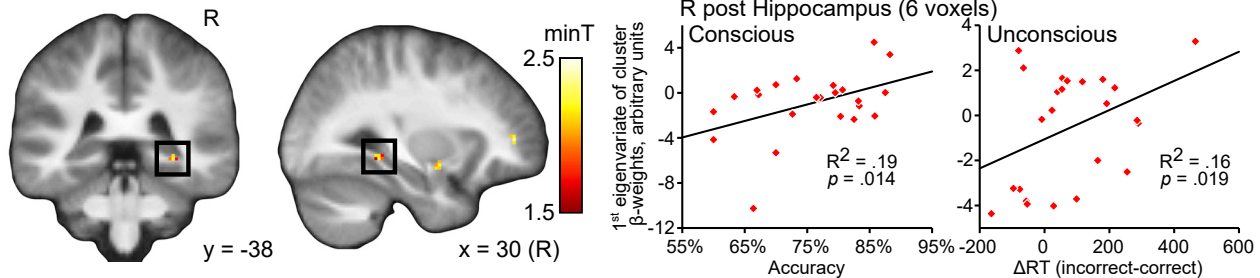
We had implemented two control conditions in the experimental design used in all parts of this research, but these control conditions are only relevant for the analysis of the fMRI data. In control condition 1, five animals crossed the scene one by one moving through the hiding place without lingering inside. This circumstance alleviates the need for temporal relational inference and leaves the encoding of the sequence of the appearance of five animals. In control condition 2, a cartoon clip featured one single animal crossing the scene five times in the same direction, moving straight through the hiding place. The comparison between the experimental condition and control condition 1 allows isolating BOLD signal underlying temporal relational inference. The comparison between control condition 1 and control condition 2 allows isolating BOLD signal underlying the encoding of a temporal sequence.

Brain areas exhibiting encoding- and retrieval-related activity increases that are common to conscious and unconscious processing

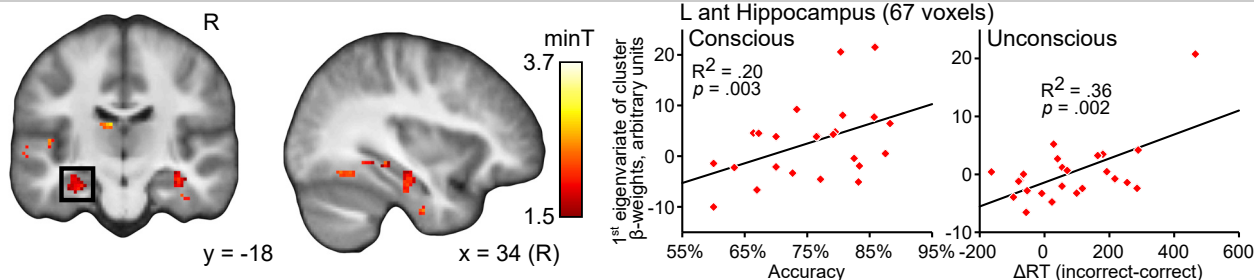
We computed conjunction analyses of positive brain-behavior correlations that had been computed for each experiment (see STAR Methods for results and Tables S2 and S3). In the encoding fMRI time series, we used a block design (one block corresponds to one clip) and contrasted the experimental condition with control condition 1. We then correlated this contrast with retrieval performance (strongly masked clips: RT difference; weakly masked clips: retrieval accuracy) between subjects to reveal signal associated with temporal relational inference and retrieval success. In the retrieval fMRI time series, we used a rapid event-related design that allowed contrasting correct versus incorrect retrieval responses in the experimental condition. We computed the contrast of correct versus incorrect retrieval responses and correlated this contrast with retrieval performance (strongly masked clips: RT difference; weakly masked clips: retrieval accuracy) between subjects.

The right posterior hippocampus was the only brain region that featured both in the encoding-related conjunction (6 voxels,

### Encoding Conjunction: Conscious $\wedge$ Unconscious



### Retrieval Conjunction: Conscious $\wedge$ Unconscious



**Figure 4. Brain areas exhibiting encoding- and retrieval-related activity increases that are common to conscious and unconscious processing (conjunction analyses)**

The top panel presents positive brain-behavior correlations in the right posterior hippocampus pertaining to the unconscious and conscious drawing of inferences while watching strongly and weakly masked clips. The bottom panel presents positive brain-behavior correlations in the left anterior hippocampus (also visible are right anterior and right posterior hippocampal results) pertaining to the unconscious and conscious retrieval of the formed inferences. Conjunction results are visualized on brain slices. Scatterplots show brain-behavior correlations at the hippocampal location highlighted by a black rectangle on the brain slices. See also [Tables S2–S4](#).

peak at MNI 30,  $-38$ ,  $-4$ ;  $T = 2.10$ ;  $p_{\text{uncor}} < 0.005$ ) and in the retrieval-related conjunction (15 voxels, peak at MNI = 32,  $-34$ ,  $-6$ ;  $T = 2.77$ ;  $p_{\text{uncor}} < 0.005$ ) ([Figure 4](#); [Table S4](#)).

The encoding-related conjunction analysis yielded further regions of commonly increased activity that correlated with retrieval success, namely, regions in right prefrontal cortex. The significant clusters were located in the right superior frontal gyrus (BA8; 16 voxels, peak at 8, 30, 34;  $T = 2.25$ ;  $p_{\text{uncor}} < 0.001$ ), the right middle frontal gyrus (BA10; 24 voxels, peak at 28, 48, 6;  $T = 2.46$ ;  $p_{\text{uncor}} < 0.001$ ), the right anterior cingulate gyrus (BA32; 20 voxels, peak at 12, 36, 18;  $T = 2.38$ ;  $p_{\text{uncor}} < 0.001$ ), and the right claustrum (13 voxels, peak at 30, 0,  $-10$ ;  $T = 2.20$ ;  $p_{\text{uncor}} < 0.001$ ) ([Table S4](#)). The conjunction analysis of the inverse correlations yielded no significant results.

The retrieval-related conjunction analysis yielded many further (apart from the right posterior hippocampus) regions of commonly increased activity that correlated with retrieval success. The significant clusters were located in the following regions ([Table S4](#)): bilateral anterior hippocampus (left side: 67 voxels, peak at MNI  $-28$ ,  $-22$ ,  $-20$ ;  $T = 3.04$ ; right side: 41 voxels, peak at MNI 34,  $-22$ ,  $-18$ ;  $T = 2.36$ ;  $p_{\text{uncor}} < 0.005$ ) ([Figure 4](#)), bilateral parahippocampal gyri, bilateral lingual and fusiform gyri, bilateral middle temporal gyri, left medial, middle and inferior frontal gyri, left supramarginal gyrus, right precuneus, bilateral cuneus, and bilateral occipital gyri. The conjunction analysis of the inverse correlations yielded no significant results.

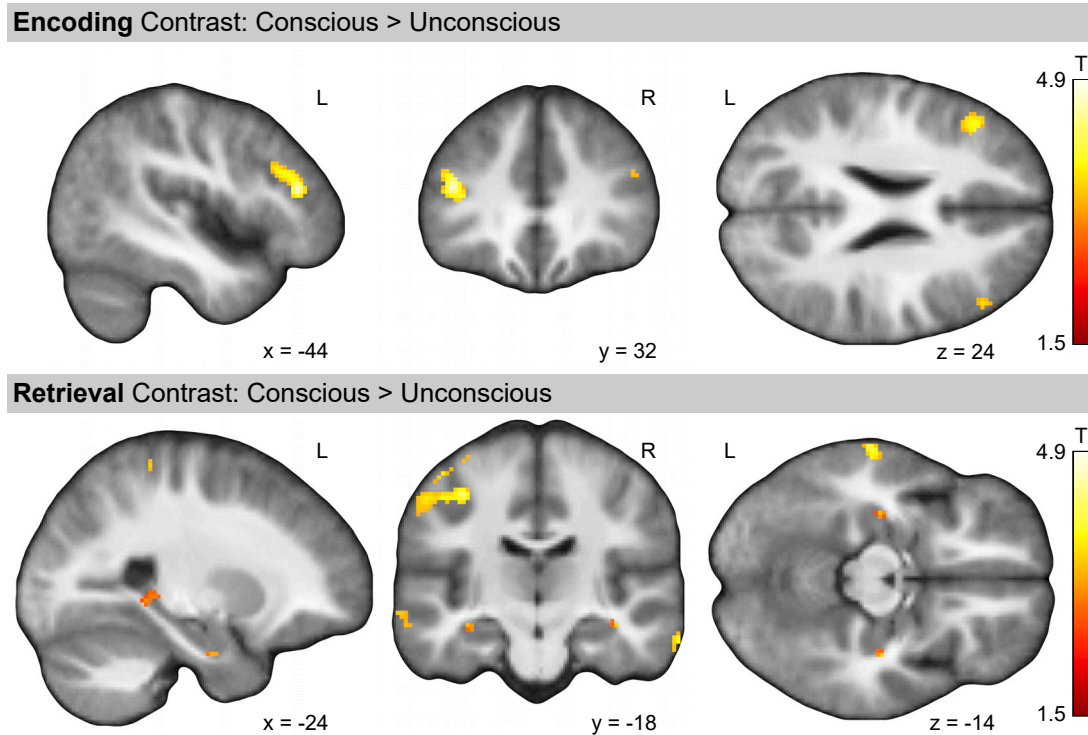
The strength of the correlation results illustrated in [Figure 4](#) did not differ significantly between consciousness levels. This was true for both encoding (right posterior hippocampus, comparison between  $r_{\text{conscious}} = 0.44$  and  $r_{\text{unconscious}} = 0.40$ ; Fisher's  $z = 0.175$ ,  $p = 0.86$ , Zou's 95% CI:  $[-0.45-0.54]$ ) and retrieval (left anterior hippocampus, comparison between  $r_{\text{conscious}} = 0.44$  and  $r_{\text{unconscious}} = 0.60$ ; Fisher's  $z = -0.683$ ,  $p = 0.49$ , Zou's 95% CI:  $[-0.60-0.28]$ ). Hence, the relationship between hippocampal activity and retrieval performance did not differ significantly between conscious and unconscious processing. When removing outliers from the brain-behavior correlations ([STAR Methods](#)), the hippocampal clusters depicted in [Figure 4](#) remained statistically significant with the exception of the retrieval-related cluster situated in the right anterior hippocampus with its peak at MNI = 34,  $-22$ ,  $-18$ .

Brain areas exhibiting stronger, broader, or additional encoding- and retrieval-related activity increases during conscious versus unconscious processing

We had expected conscious versus unconscious processing to be paralleled by enhanced fMRI signal increases. To reveal signal differences between consciousness conditions, we computed contrasts at the second level between encoding- or retrieval-related contrasts at the first level (see [STAR Methods](#) for first-level results).

Conscious versus unconscious temporal relational inference yielded signal increases in brain areas that were not commonly activated, namely in Broca's area in the left inferior frontal gyrus





**Figure 5. Enhanced signal increases during conscious versus unconscious processing**

Top panel: conscious versus unconscious temporal relational inference at encoding was paralleled by bilateral signal increases in the inferior frontal gyrus. Bottom panel: the conscious versus unconscious retrieval of the formed inferences was paralleled by signal increases in bilateral anterior hippocampus and the left posterior hippocampus. See also [Table S5](#).

(BA45; 203 voxels, peak at MNI  $-42, 34, 18$ ;  $T = 4.36$ ;  $p_{\text{uncor}} < 0.001$ ) and its homotopic area in the right inferior frontal gyrus (BA45; 17 voxels, peak at MNI  $48, 36, 22$ ;  $T = 3.64$ ;  $p_{\text{uncor}} < 0.001$ ) ([Figure 5](#)). A further cluster was located in the right thalamus (29 voxels, peak at MNI  $4, -10, -4$ ;  $T = 4.14$ ;  $p_{\text{uncor}} < 0.001$ ) ([Table S5](#)). The inverse contrast yielded no significant result. That is, no brain area exhibited increased signal during unconscious versus conscious temporal relation inference.

The conscious versus unconscious retrieval of the formed inferences yielded signal increases in (1) regions that were commonly activated and (2) in regions adjacent to commonly activated regions, and (3) in regions that were outside the commonly activated regions ([Table S5](#)). (1) The enhanced signal in commonly activated regions was located in the right middle temporal gyrus (BA21) and the left anterior hippocampus ([Figure 5](#); [Table S5](#)). (2) Enhanced signal in regions adjacent to commonly activated regions was located in bilateral middle frontal gyri (BA10), left middle frontal gyrus (BA46), bilateral middle temporal gyri (BA21), bilateral superior temporal gyri (BA22), and bilateral anterior hippocampi ([Figure 5](#)). (3) Enhanced signal in regions outside the commonly activated regions was located in the right anterior cingulate gyrus (BA24 and 32), right middle (BA10) and inferior frontal gyrus (BA47), right amygdala and bilateral entorhinal cortex, right claustrum, left pre- and postcentral gyrus, right supramarginal gyrus (BA40), left superior parietal lobule (BA7), and the left posterior hippocampus ([Figure 5](#); [Table S5](#)). The inverse contrast yielded no significant result. That is, no

brain area exhibited increased signal during unconscious versus conscious retrieval ([Table S5](#)).

### Any conscious awareness of the strongly masked cartoon clips?

Participants rated their conscious awareness of clips following the presentation of each clip using the 4-point perceptual awareness scale (PAS;<sup>52</sup>), where (1) means no awareness, (2) means a vague feeling that something was present, (3) means an impression of a scene or animal, and (4) means a clear percept of a scene and animals. A rating of 1 was given to 93.5% of all strongly masked clips. Most participants (68.3%) rated all of the strongly masked clips with a 1. [Table S6](#) provides PAS ratings. The qualitative pattern of the behavioral results remained unchanged if strongly masked trials that received a PAS rating  $>1$  were excluded from data analysis. Hence, the above-reported behavioral results originate from genuinely unconscious clip processing. In addition, results obtained on the objective awareness tests also indicate that participants had processed the strongly masked clips unconsciously ([STAR Methods](#)).

### DISCUSSION

Making sense of the world presumes that we combine distinct events to infer commonalities and relations between events through inferential reasoning, which depends on episodic and working memory and involves the hippocampus.<sup>1</sup> Here, we

observed successful conscious and unconscious temporal relational inference and a successful delayed retrieval of inferences. Delayed retrieval performance was strong in the conscious but weak in the unconscious condition. Unconscious retrieval performance remained significant when information load increased, while conscious retrieval performance dropped. Conscious and unconscious inference were paralleled by activity increases in the right prefrontal cortex and hippocampus, which correlated with retrieval success. The delayed conscious and unconscious retrieval of inferences was paralleled by activity increases in an episodic memory retrieval network, which also correlated with retrieval success. Conscious versus unconscious processing spawned steeper activity increases and broader activation clusters within commonly activated regions and beyond.

A notorious concern with subliminal experiments pertains to the efficacy of the masking protocol. Therefore, we had applied both an online subjective and a delayed objective measure of clip awareness. Participants gave a subjective awareness rating on a perceptual awareness scale<sup>52</sup> following each clip presentation. The objective measure of clip awareness consisted in a forced-choice test given in a trial-by-trial encoding-immediate-test procedure applied following experimentation. Data collected with both methods suggest that the psychophysics in all subliminal experiments prohibited participants from gaining conscious access to clips. Moreover, the absence of accuracy effects and the absence of correlations between the implicit and explicit retrieval measure in the subliminal experiments also suggest unconscious clip processing.

We informed participants of subliminal clips before experimentation to match instructions between consciousness levels. Yet, by laying open the subliminal stimulation before encoding and then asking participants to retrieve animal trajectories that they could not see, we put participants in an unsettling situation. Therefore, experimenters usually keep participants naive about subliminal presentations and give them indirect retrieval instructions that introduce the memory test as a memory-unrelated new task to obviate participants' thinking back to the encoding situation. This procedure allows participants to adopt a relaxed, nonintentional attitude, which improves subliminal stimulus processing.<sup>51,60</sup> When, as here, participants are informed of subliminal stimuli and retrieval instructions are direct instead of indirect, participants adopt a scrutinizing, intentional encoding, and retrieval mode. Our participants have sometimes relied on common sense regarding which animals would normally choose spending time together rather than on unconscious knowledge. Although all participants may have had this inclination, previous research revealed that particularly deliberative decision makers prefer relying on consciously accessible knowledge.<sup>61</sup> They lower their motivation when asked to rely on intuition.<sup>62</sup> Concordantly, only habitually intuitive decision makers exhibited retrieval effects in the present study. This result replicates the previous result<sup>26</sup> that only intuitive, but not deliberative, decision makers draw from their unconscious object-in-space knowledge at test.<sup>26</sup> Involuntary eye movements recorded at test<sup>26</sup> revealed, however, that the deliberative decision makers could spontaneously access their unconscious knowledge, although they failed letting this knowledge guide their deliberate retrieval responses. We cannot determine whether our deliberative decision makers had also encoded the clips but then failed letting this

knowledge guide their retrieval responses or whether they simply failed to encode the clips in the first place. Anyway, the limitation of memory effects to intuition-experienced participants corroborates the general finding that intuition is positively related to implicit learning from supraliminal<sup>63,64</sup> and subliminal stimuli<sup>26</sup> and to dealing with a high information load.<sup>65–67</sup>

Implicit retrieval effects in the present study were limited to the soft memory measure, the reaction time difference between correct and incorrect retrieval responses. The reaction time difference is a more subtle memory measure than response accuracy and emerges not only in tests of unconscious memory<sup>24,34,53</sup> but also in tests of conscious memory.<sup>68–71</sup> Correspondingly, incorrect retrieval responses took longer than correct retrieval responses in both consciousness conditions. This reaction time difference may reflect a swift memory reactivation in correct responses subsequent to successful memory formation, although this reactivation had not sufficed to impact on the direction of retrieval responses in the unconscious condition. The reaction time difference may also result from slow incorrect response due to long search times, low confidence,<sup>72</sup> and idling because of poor memory formation. In any event, reaction times systematically separated between the two retrieval categories “correct” versus “incorrect,” which pertain to the relevant dimension: animals that lingered simultaneously inside the hiding place or not. With reaction times differentiating between the realities of animal encounters, we must conclude that participants had drawn temporal relational inferences unconsciously. The weakness of the current compared to previous unconscious memory effects<sup>23,31,35</sup> likely owes to task difficulty and to overt and direct instead of covert and indirect instructions. The implicit effect occurred in four experiments (experiments with medium and high load, replication experiment with medium load, fMRI experiment with medium load), which gives rise to the reality of an unconscious what-where-when encoding and delayed retrieval.

The instructed task during encoding was to watch five animals entering and leaving a hiding place to determine which animals lingered simultaneously inside the hiding place. Note that the tested information—animals that lingered simultaneously inside—was never displayed but needed to be inferred. Therefore, visual priming and visual familiarity cannot account for the storage of animal encounters. The necessity of drawing temporal relational inferences called upon mental imagery and working memory because drawing inferences implicates the monitoring of the time points of each animal's entrance and exit and the constant mental updating of the current number and appearance of animals inside the hiding place. This constant updating must have invoked the episodic buffer of working memory that depends on the hippocampus.<sup>40,73</sup> Indeed, activity rose within the right posterior hippocampus during both the conscious and unconscious drawing of inferences and this activity correlated with retrieval success. In addition, regions of the right prefrontal cortex were activated during conscious and unconscious inference, and this activation correlated with retrieval success. These regions were located in the superior and middle frontal gyrus, anterior cingulate gyrus, and the right claustrum, which had been previously associated with visual-spatial working memory.<sup>74,75</sup> The claustrum is strongly linked to prefrontal cortex, helps in attending to salient sensory events, and facilitates executive functions that control posterior cortices.<sup>76,77</sup> Such exigent

unconscious working memory computations and corresponding prefrontal activations exceed previous findings regarding an unconscious form of working memory<sup>78–80</sup> and inform theories that firmly associate working memory with conscious stimulus processing.<sup>40,41,81</sup> These findings also challenge theories of consciousness that consider conscious scene perception necessary for the understanding of an unfolding event, for the encoding of what-where-when associations, and for inferential reasoning.<sup>81–86</sup>

Participants knew that they needed to store their inferences for later retrieval. This requirement applied to the medium and high load condition, where the encoding-test interval spanned 3.5 and 6 min, respectively. Both the conscious and unconscious retrieval of inferences was accompanied by transcortical activity increases that correlated with retrieval success. Results were located in bilateral regions of the anterior and posterior hippocampus, parahippocampal gyrus, lingual and fusiform gyrus, middle temporal gyrus, as well as unilateral left regions in the medial, middle, and inferior frontal gyrus and the supramarginal gyrus. These brain regions correspond to areas that had previously been associated with retrieval success for visual-spatial episodic memories.<sup>87,88</sup> This retrieval network comprises critical hubs of the episodic retrieval network including bilateral hippocampus, which underscores our theoretical claim<sup>4</sup> that the episodic memory system and the hippocampus operate task oriented and independently of conscious awareness. Reiterating the above argument for a role of prefrontal cortices in higher cognitive functions that run unconsciously, we point out that left prefrontal cortices increased their activation level in a retrieval-success-related mode along with many down-stream cortical areas. Hence, specialized cortical processors of the human brain serve environmental demands irrespective of the organism's conscious awareness of the environment with higher cognitive functions and prefrontal processors making no exception to this rule.<sup>4,89,90</sup>

Although sharing the same memory systems, unconscious differed from conscious memory formation: memories formed from subliminal clips were weak and failed to direct retrieval responses, while memories formed from supraliminal clips were strong and influenced retrieval responses. This difference in the memories' behavioral impact dovetails with underlying differences in brain activation. Conscious versus unconscious processing spawned steeper activity increases and broader activation clusters within commonly activated networks including the hippocampus. This difference in strength likely originated from differences in the visual signal emitted by weakly versus strongly masked clips. Weak masking provides for a stronger visual input with stronger visual cortex activation that increases neural propagation strength and long-range coherence.<sup>47,91</sup> Yet, conscious versus unconscious processing was also associated with activations in additional brain regions, namely, bilateral Broca's area during inferential reasoning and right anterior cingulate gyrus, right middle and inferior frontal gyrus, right supramarginal gyrus, and left superior parietal lobule during retrieval. The associated mental processes might correspond to a verbal, in addition to the nonverbal, coding of animals and strategic- and effort-related processes applied when retrieval is conscious. A stronger recruitment of the visual system and the episodic memory system, including the hippocampus,

during conscious versus unconscious encoding/retrieval does not derogate the importance of these networks for unconscious processing. In fact, the strength of encoding- and retrieval-related brain-behavior correlations in the hippocampus (Figure 4) did not differ significantly between the two consciousness conditions.

The downside of the large behavioral impact of consciously formed memories was a significant decline of retrieval accuracy, when information load was high, while no significant decline was observed for memories formed unconsciously. This result is reminiscent of a hypothesis in computational neuroscience postulating a larger memory capacity but poorer retrieval fidelity for sparsely (perhaps unconsciously) versus thickly (perhaps consciously) coded memories.<sup>92–95</sup> Although fMRI does not have the spatial resolution needed to track activity in single neurons, we speculate that the neural memory traces underlying individual unconscious (versus conscious) memories are less susceptible to interference because they are sparse with distinct neurons coding for distinct memories. In contrast, neural memory traces underlying individual conscious memories involve activity in large neural assemblies with neurons coding for several memories, which produces overlaps between memory traces and breeds interference and forgetting.<sup>96–99</sup> In conclusion, this work suggests that unconscious episodic memory provides for weak behavioral effects but comes with a large memory capacity, while conscious episodic memory provides for strong behavioral effects but comes with a small memory capacity.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Participants of the first part of this research
  - Participants of the second part of this research
  - Participants of the third part of this research
- **METHOD DETAILS**
  - Experimental conditions
  - Clip presentation
  - Technical setup
  - Betsch personality inventory
  - First part of this research: Six behavioral experiments with manipulated information load
  - Second part of this research: Replication experiment with medium information load and strong masking for unconscious encoding
  - Third part of this research: fMRI experiments with strongly and weakly masked cartoon clips and a medium information load
  - Experimental procedure
  - MRI data acquisition
  - Subjective and objective measures of clip awareness
  - Results of the objective awareness test

● **QUANTIFICATION AND STATISTICAL ANALYSIS**

- Statistical analyses of behavioral data
- Control analyses with the behavioral data
- Biased stimulus material?
- Participant-inherent response bias?
- Statistical analyses of fMRI data
- Additional fMRI results obtained in the individual experiments

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2021.06.012>.

**ACKNOWLEDGMENTS**

This work was supported by the “Gottfried und Julia Bangerter-Rhyner-Stiftung” and a Center for Cognition, Learning and Memory (CCLM) Ph.D. stipend to S.W. We thank Timothée Zünd, Tanja Lazarevic, and Marina Wunderlin for their help with data collection. The Swiss artists Tom Kuenzli and Nicolas Kristen drew the beautiful artwork of the images that constituted the cartoon clips, which were used for the first time in the present experiments.

**AUTHOR CONTRIBUTIONS**

Conceptualization, E.S., S.R., and K.H.; methodology, E.S., S.W., M.A.Z., S.R., S.K., R.W., and K.H.; data collection, E.S., S.W., and F.S.; data analysis, E.S., S.W., S.R., F.S., and M.A.Z.; writing—original draft, E.S., and K.H.; writing—review & editing, K.H., S.R., and M.A.Z.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: June 30, 2020

Revised: January 29, 2021

Accepted: June 3, 2021

Published: July 12, 2021

**REFERENCES**

1. Eichenbaum, H., and Fortin, N.J. (2009). The neurobiology of memory based predictions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *364*, 1183–1191.
2. Tulving, E. (1972). Episodic and Semantic Memory. In *Organization of Memory*, E. Tulving, and W. Donaldson, eds. (Academic Press), pp. 381–403.
3. Cohen, N.J., and Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System* (MIT Press).
4. Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nat. Rev. Neurosci.* *11*, 523–532.
5. McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). 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.* *102*, 419–457.
6. Moses, S.N., and Ryan, J.D. (2006). A comparison and evaluation of the predictions of relational and conjunctive accounts of hippocampal function. *Hippocampus* *16*, 43–65.
7. Norman, K.A., and O'Reilly, R.C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol. Rev.* *110*, 611–646.
8. O'Reilly, R.C., and Rudy, J.W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus* *10*, 389–397.
9. Koscik, T.R., and Tranel, D. (2012). The human ventromedial prefrontal cortex is critical for transitive inference. *J. Cogn. Neurosci.* *24*, 1191–1204.
10. Ryan, J.D., D'Angelo, M.C., Kamino, D., Ostreicher, M., Moses, S.N., and Rosenbaum, R.S. (2016). Relational learning and transitive expression in aging and amnesia. *Hippocampus* *26*, 170–184.
11. D'Angelo, M.C., Rosenbaum, R.S., and Ryan, J.D. (2016). Impaired inference in a case of developmental amnesia. *Hippocampus* *26*, 1291–1302.
12. FitzGerald, T.H.B., Hämmerer, D., Friston, K.J., Li, S.-C., and Dolan, R.J. (2017). Sequential inference as a mode of cognition and its correlates in fronto-parietal and hippocampal brain regions. *PLoS Comput. Biol.* *13*, e1005418.
13. Kumaran, D. (2012). What representations and computations underpin the contribution of the hippocampus to generalization and inference? *Front. Hum. Neurosci.* *6*, 157.
14. Reber, T.P., Do Lam, A.T., Axmacher, N., Elger, C.E., Helmstaedter, C., Henke, K., and Fell, J. (2016). Intracranial EEG correlates of implicit relational inference within the hippocampus. *Hippocampus* *26*, 54–66.
15. Zeithamova, D., Schlichting, M.L., and Preston, A.R. (2012). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Front. Hum. Neurosci.* *6*, 70.
16. Gabrieli, J.D.E. (1998). Cognitive neuroscience of human memory. *Annu. Rev. Psychol.* *49*, 87–115.
17. Moscovitch, M. (2008). The hippocampus as a “stupid,” domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Can. J. Exp. Psychol.* *62*, 62–79.
18. Squire, L.R., and Zola-Morgan, M. (1991). Memory and brain. *Oxford Rev. Biol. Psychol.* *12*, 7–65.
19. Tulving, E. (2002). Episodic memory: from mind to brain. *Annu. Rev. Psychol.* *53*, 1–25.
20. Smith, C., and Squire, L.R. (2005). Declarative memory, awareness, and transitive inference. *J. Neurosci.* *25*, 10138–10146.
21. Jacobs, C., and Silvanto, J. (2015). How is working memory content consciously experienced? The ‘conscious copy’ model of WM introspection. *Neurosci. Biobehav. Rev.* *55*, 510–519.
22. Schacter, D.L. (1998). Memory and awareness. *Science* *280*, 59–60.
23. Duss, S.B., Reber, T.P., Hänggi, J., Schwab, S., Wiest, R., Müri, R.M., Brugger, P., Gutbrod, K., and Henke, K. (2014). Unconscious relational encoding depends on hippocampus. *Brain* *137*, 3355–3370.
24. Henke, K., Mondadori, C.R.A., Treyer, V., Nitsch, R.M., Buck, A., and Hock, C. (2003). Nonconscious formation and reactivation of semantic associations by way of the medial temporal lobe. *Neuropsychologia* *41*, 863–876.
25. Henke, K., Reber, T.P., and Duss, S.B. (2013). Integrating events across levels of consciousness. *Front. Behav. Neurosci.* *7*, 68.
26. Wuethrich, S., Hannula, D.E., Mast, F.W., and Henke, K. (2018). Subliminal encoding and flexible retrieval of objects in scenes. *Hippocampus* *28*, 633–643.
27. Züst, M.A., Colella, P., Reber, T.P., Vuilleumier, P., Hauf, M., Ruch, S., and Henke, K. (2015). Hippocampus is place of interaction between unconscious and conscious memories. *PLoS ONE* *10*, e0122459.
28. Züst, M.A., Ruch, S., Wiest, R., and Henke, K. (2019). Implicit vocabulary learning during sleep is bound to slow-wave peaks. *Curr. Biol.* *29*, 541–553.e7.
29. Reber, T.P., and Henke, K. (2011). Rapid formation and flexible expression of memories of subliminal word pairs. *Front. Psychol.* *2*, 343.
30. Reber, T.P., Luechinger, R., Boesiger, P., and Henke, K. (2014). Detecting analogies unconsciously. *Front. Behav. Neurosci.* *8*, 9.
31. Duss, S.B., Oggier, S., Reber, T.P., and Henke, K. (2011). Formation of semantic associations between subliminally presented face-word pairs. *Conscious. Cogn.* *20*, 928–935.
32. Ruch, S., Herbert, E., and Henke, K. (2017). Subliminally and supraliminally acquired long-term memories jointly bias delayed decisions. *Front. Psychol.* *8*, 1542.

33. Henke, K., Treyer, V., Nagy, E.T., Kneifel, S., Dürsteler, M., Nitsch, R.M., and Buck, A. (2003). Active hippocampus during nonconscious memories. *Conscious. Cogn.* *12*, 31–48.
34. Reber, T.P., and Henke, K. (2012). Integrating unseen events over time. *Conscious. Cogn.* *21*, 953–960.
35. Reber, T.P., Luechinger, R., Boesiger, P., and Henke, K. (2012). Unconscious relational inference recruits the hippocampus. *J. Neurosci.* *32*, 6138–6148.
36. Rosenthal, C.R., Andrews, S.K., Antoniadou, C.A., Kennard, C., and Soto, D. (2016). Learning and recognition of a non-conscious sequence of events in human primary visual cortex. *Curr. Biol.* *26*, 834–841.
37. Moors, P., Wagemans, J., and de-Wit, L. (2017). Causal events enter awareness faster than non-causal events. *PeerJ* *5*, e2932.
38. Kawakami, N., and Yoshida, F. (2015). Perceiving a story outside of conscious awareness: when we infer narrative attributes from subliminal sequential stimuli. *Conscious. Cogn.* *33*, 53–66.
39. Gómez, J.-C. (2005). Species comparative studies and cognitive development. *Trends Cogn. Sci.* *9*, 118–125.
40. Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* *4*, 417–423.
41. Velichkovsky, B.B. (2017). Consciousness and working memory: current trends and research perspectives. *Conscious. Cogn.* *55*, 35–45.
42. Howard, M.W., and Eichenbaum, H. (2015). Time and space in the hippocampus. *Brain Res.* *1621*, 345–354.
43. MacDonald, C.J., Carrow, S., Place, R., and Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *J. Neurosci.* *33*, 14607–14616.
44. Palombo, D.J., Reid, A.G., Thavabalasingam, S., Hunsberger, R., Lee, A.C.H., and Verfaellie, M. (2020). The human medial temporal lobe is necessary for remembering durations within a sequence of events but not durations of individual events. *J. Cogn. Neurosci.* *32*, 497–507.
45. Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science* *321*, 1322–1327.
46. Dehaene, S., Naccache, L., Cohen, L., Bihan, D.L., Mangin, J.-F., Poline, J.-B., and Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci.* *4*, 752–758.
47. Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., Cohen, L., and Naccache, L. (2009). Converging intracranial markers of conscious access. *PLoS Biol.* *7*, e61.
48. Grill-Spector, K., Kushnir, T., Hendler, T., and Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nat. Neurosci.* *3*, 837–843.
49. Kouider, S., and Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *362*, 857–875.
50. Reber, T.P., Faber, J., Niediek, J., Boström, J., Elger, C.E., and Mormann, F. (2017). Single-neuron correlates of conscious perception in the human medial temporal lobe. *Curr. Biol.* *27*, 2991–2998.e2.
51. Snodgrass, M., and Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: replicable and qualitatively different unconscious perceptual effects. *Cognition* *101*, 43–79.
52. Overgaard, M., Rote, J., Mouridsen, K., and Ramsøy, T.Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Conscious. Cogn.* *15*, 700–708.
53. Savage, G.R., Saling, M.M., Davis, C.W., and Berkovic, S.F. (2002). Direct and indirect measures of verbal relational memory following anterior temporal lobectomy. *Neuropsychologia* *40*, 302–316.
54. R Core Team (2019). R: A language and environment for statistical computing.
55. Baguley, T. (2010). Book Review: understanding psychology as a science. *Br. J. Math. Stat. Psychol.* *63*, 695–698.
56. Ruch, S., Züst, M.A., and Henke, K. (2016). Subliminal messages exert long-term effects on decision-making. *Neurosci. Conscious.* *2016*, niw013.
57. Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Front. Psychol.* *5*, 781.
58. Dienes, Z. (2008). *Understanding Psychology as a Science: An Introduction to Scientific and Statistical Inference* (Palgrave Macmillan).
59. Degonda, N., Mondadori, C.R.A., Bosshardt, S., Schmidt, C.F., Boesiger, P., Nitsch, R.M., Hock, C., and Henke, K. (2005). Implicit associative learning engages the hippocampus and interacts with explicit associative learning. *Neuron* *46*, 505–520.
60. Verwijmeren, T., Karremans, J.C., Bernitter, S.F., Stroebe, W., and Wigboldus, D.H.J. (2013). Warning: You are being primed! The effect of a warning on the impact of subliminal ads. *J. Exp. Soc. Psychol.* *49*, 1124–1129.
61. Voss, J.L., and Paller, K.A. (2010). What makes recognition without awareness appear to be elusive? Strategic factors that influence the accuracy of guesses. *Learn. Mem.* *17*, 460–468.
62. Milyavsky, M., Hassin, R.R., and Schul, Y. (2012). Guess what? Implicit motivation boosts the influence of subliminal information on choice. *Conscious. Cogn.* *21*, 1232–1241.
63. Kaufman, S.B., Deyoung, C.G., Gray, J.R., Jiménez, L., Brown, J., and Mackintosh, N. (2010). Implicit learning as an ability. *Cognition* *116*, 321–340.
64. Woolhouse, L.S., and Bayne, R. (2000). Personality and the use of intuition: individual differences in strategy and performance on an implicit learning task. *Eur. J. Pers.* *14*, 157–169.
65. Betsch, T., Plessner, H., Schwieren, C., and Gütig, R. (2001). I like it but I don't know why: a value-account approach to implicit attitude formation. *Pers. Soc. Psychol. Bull.* *27*, 242–253.
66. Usher, M., Russo, Z., Weyers, M., Brauner, R., and Zakay, D. (2011). The impact of the mode of thought in complex decisions: intuitive decisions are better. *Front. Psychol.* *2*, 37.
67. Wilson, T.D., and Schooler, J.W. (1991). Thinking too much: introspection can reduce the quality of preferences and decisions. *J. Pers. Soc. Psychol.* *60*, 181–192.
68. Diller, D.E., Nobel, P.A., and Shiffrin, R.M. (2001). An ARC-REM model for accuracy and response time in recognition and recall. *J. Exp. Psychol. Learn. Mem. Cogn.* *27*, 414–435.
69. Murdock, B.B., and Dufty, P.O. (1972). Strength theory and recognition memory. *J. Exp. Psychol.* *94*, 284–290.
70. Nobel, P.A., and Shiffrin, R.M. (2001). Retrieval processes in recognition and cued recall. *J. Exp. Psychol. Learn. Mem. Cogn.* *27*, 384–413.
71. Norman, D.A., and Wickelgren, W.A. (1969). Strength theory of decision rules and latency in retrieval from short-term memory. *J. Math. Psychol.* *6*, 192–208.
72. Weidemann, C.T., and Kahana, M.J. (2016). Assessing recognition memory using confidence ratings and response times. *R. Soc. Open Sci.* *3*, 150670.
73. Berlinger, M., Bottini, G., Basilico, S., Silani, G., Zanardi, G., Sberna, M., Colombo, N., Sterzi, R., Scialfa, G., and Paulesu, E. (2008). Anatomy of the episodic buffer: a voxel-based morphometry study in patients with dementia. *Behav. Neurol.* *19*, 29–34.
74. McCarthy, G., Puce, A., Constable, R.T., Krystal, J.H., Gore, J.C., and Goldman-Rakic, P. (1996). Activation of human prefrontal cortex during spatial and nonspatial working memory tasks measured by functional MRI. *Cereb. Cortex* *6*, 600–611.
75. Klingberg, T. (2006). Development of a superior frontal-intraparietal network for visuo-spatial working memory. *Neuropsychologia* *44*, 2171–2177.
76. White, M.G., and Mathur, B.N. (2018). Frontal cortical control of posterior sensory and association cortices through the claustrum. *Brain Struct. Funct.* *223*, 2999–3006.

77. Smith, J.B., Lee, A.K., and Jackson, J. (2020). The claustrum. *Curr. Biol.* **30**, R1401–R1406.
78. Bergström, F., and Eriksson, J. (2018). Neural evidence for non-conscious working memory. *Cereb. Cortex* **28**, 3217–3228.
79. King, J.-R., Pescetelli, N., and Dehaene, S. (2016). Brain mechanisms underlying the brief maintenance of seen and unseen sensory information. *Neuron* **92**, 1122–1134.
80. Soto, D., and Silvanto, J. (2016). Is conscious awareness needed for all working memory processes? *Neurosci. Conscious.* **2016**, niw009.
81. Mashour, G.A., Roelfsema, P., Changeux, J.-P., and Dehaene, S. (2020). Conscious processing and the global neuronal workspace hypothesis. *Neuron* **105**, 776–798.
82. Mudrik, L., Faivre, N., and Koch, C. (2014). Information integration without awareness. *Trends Cogn. Sci.* **18**, 488–496.
83. Scott, R.B., Samaha, J., Chrisley, R., and Dienes, Z. (2018). Prevailing theories of consciousness are challenged by novel cross-modal associations acquired between subliminal stimuli. *Cognition* **175**, 169–185.
84. Frith, C., and Dolan, R. (1996). The role of the prefrontal cortex in higher cognitive functions. *Brain Res. Cogn. Brain Res.* **5**, 175–181.
85. Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* **17**, 307–321.
86. Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* **17**, 450–461.
87. Spaniol, J., Davidson, P.S.R., Kim, A.S.N., Han, H., Moscovitch, M., and Grady, C.L. (2009). Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia* **47**, 1765–1779.
88. Cabeza, R., Prince, S.E., Daselaar, S.M., Greenberg, D.L., Budde, M., Dolcos, F., LaBar, K.S., and Rubin, D.C. (2004). Brain activity during episodic retrieval of autobiographical and laboratory events: an fMRI study using a novel photo paradigm. *J. Cogn. Neurosci.* **16**, 1583–1594.
89. Hassin, R.R. (2013). Yes it can: on the functional abilities of the human unconscious. *Perspect. Psychol. Sci.* **8**, 195–207.
90. Cabeza, R., Stanley, M.L., and Moscovitch, M. (2018). Process-specific alliances (PSAs) in cognitive neuroscience. *Trends Cogn. Sci.* **22**, 996–1010.
91. Cohen, M.X., van Gaal, S., Ridderinkhof, K.R., and Lamme, V.A.F. (2009). Unconscious errors enhance prefrontal-occipital oscillatory synchrony. *Front. Hum. Neurosci.* **3**, 54.
92. Einarsson, H., Lengler, J., and Steger, A. (2014). A high-capacity model for one shot association learning in the brain. *Front. Comput. Neurosci.* **8**, 140.
93. Leibold, C., and Kempster, R. (2008). Sparseness constrains the prolongation of memory lifetime via synaptic metaplasticity. *Cereb. Cortex* **18**, 67–77.
94. Medina, D., and Leibold, C. (2014). Re-encoding of associations by recurrent plasticity increases memory capacity. *Front. Synaptic Neurosci.* **6**, 13.
95. Nadal, J.-P. (1991). Associative memory: on the (puzzling) sparse coding limit. *J. Phys. Math. Gen.* **24**, 1093–1101.
96. Anderson, M.C. (2003). Rethinking interference theory: executive control and the mechanisms of forgetting. *J. Mem. Lang.* **49**, 415–445.
97. Kuhl, B.A., Rissman, J., Chun, M.M., and Wagner, A.D. (2011). Fidelity of neural reactivation reveals competition between memories. *Proc. Natl. Acad. Sci. USA* **108**, 5903–5908.
98. Norman, K.A., Newman, E.L., and Detre, G. (2007). A neural network model of retrieval-induced forgetting. *Psychol. Rev.* **114**, 887–953.
99. Wixted, J.T. (2004). The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.* **55**, 235–269.
100. Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M.S., Højsgaard, S., Fox, J., Lawrence, M.A., Mertens, U., Love, J., et al. (2020). *afex: Analysis of factorial experiments*.
101. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48.
102. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). *lmerTest Package: tests in linear mixed effects models*. *J. Stat. Softw.* **82**, 1–26.
103. Diedenhofen, B., and Musch, J. (2015). cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS ONE* **10**, e0121945.
104. Chapman, L.J., and Chapman, J.P. (1987). The measurement of handedness. *Brain Cogn.* **6**, 175–183.
105. Kovács, G., Vogels, R., and Orban, G.A. (1995). Cortical correlate of pattern backward masking. *Proc. Natl. Acad. Sci. USA* **92**, 5587–5591.
106. Ögmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Netw.* **6**, 245–273.
107. Betsch, C. (2004). Präferenz für Intuition und Deliberation (PID). *Z. Differ. Diagn. Psychol.* **25**, 179–197.
108. Deichmann, R., Schwarzbauer, C., and Turner, R. (2004). Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. *Neuroimage* **21**, 757–767.
109. R Core Team (2020). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing).
110. RStudio Team (2018). *RStudio: Integrated Development for R* (RStudio, PBC).
111. Lieberman, M.D., and Cunningham, W.A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Soc. Cogn. Affect. Neurosci.* **4**, 423–428.
112. Price, C.J., and Friston, K.J. (1997). Cognitive conjunction: a new approach to brain activation experiments. *Neuroimage* **5**, 261–270.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Deposited data</b>		
Raw and analyzed behavioral and fMRI data	N/A	<a href="https://doi.org/10.5061/dryad.5tb2rbp3z">https://doi.org/10.5061/dryad.5tb2rbp3z</a>
<b>Software and algorithms</b>		
MATLAB (v. R2019b)	The MathWorks, Natick, Massachusetts, United States ( <a href="https://www.mathworks.com/products/matlab.html">https://www.mathworks.com/products/matlab.html</a> )	RRID: SCR_001622
MATLAB toolbox “SPM12”	<a href="https://www.fil.ion.ucl.ac.uk/spm/software/spm12/">https://www.fil.ion.ucl.ac.uk/spm/software/spm12/</a>	RRID: SCR_007037
MATLAB toolbox “Psychophysics Toolbox” (v3.0.11)	<a href="http://psychtoolbox.org/">http://psychtoolbox.org/</a>	RRID: SCR_002881
RStudio (v1.1.447)	RStudio Team	<a href="https://www.rstudio.com">https://www.rstudio.com</a>
R (v4.0.2)	R Core Team	N/A
R-package afex (0.28-0) <sup>100</sup>	N/A	<a href="https://cran.r-project.org/web/packages/afex/index.html">https://cran.r-project.org/web/packages/afex/index.html</a>
R-package lme4 (v1.1-23) <sup>101</sup>	N/A	<a href="https://doi.org/10.18637/jss.v067.i01">https://doi.org/10.18637/jss.v067.i01</a>
R-package lmerTest (v3.1-2) <sup>102</sup>	N/A	<a href="https://doi.org/10.18637/jss.v082.i13">https://doi.org/10.18637/jss.v082.i13</a>
R-package cocor (website version 1.1-3) <sup>103</sup>	N/A	<a href="http://comparingcorrelations.org/">http://comparingcorrelations.org/</a>
Presentation	Neurobehavioral Systems, Berkely, California, United States ( <a href="https://www.neurobs.com">https://www.neurobs.com</a> )	RRID: SCR_002521
<b>Other</b>		
Benq MX764 DLP Projector (XGA conference room projector)	BenQ	<a href="https://www.projectorcentral.com/BenQ-MX764.htm">https://www.projectorcentral.com/BenQ-MX764.htm</a>
Lumina LS-PAIR response pad	Cedrus, San Pedro, California, United States ( <a href="https://cedrus.com">https://cedrus.com</a> )	<a href="https://cedrus.com/lumina/">https://cedrus.com/lumina/</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact Katharina Henke.

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- The behavioral data and the magnetic resonance imaging data have been deposited at Dryad, Dataset, <https://doi.org/10.5061/dryad.5tb2rbp3z> and are publicly available as of the date of publication. The DOI is listed in the [Key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

#### Participants of the first part of this research

In the six behavioral experiments, we examined 224 young women and men that reported normal or corrected-to-normal visual acuity and a complete absence of previous and current neurological and psychiatric disease. We informed participants before experimentation of the applied masking protocol and the critical task to infer from the timing of five animal trajectories, which animals lingered simultaneously inside a visually impenetrable hiding place. Hence, other than in previous masking experiments, we did not keep participants naive of the fact that they will be exposed to subliminal information. All participants gave their written informed consent prior

to experimentation. Of the 224 participants, 117 (age: 25.6 (mean)  $\pm$  4.9 (SD) years; 80 women) took part in the three subliminal experiments with strong masking (LL: 33 participants; ML: 36 participants; HL: 48 participants). The remaining 107 participants (age: 24.4  $\pm$  4.5 years; 81 women) took part in the three supraliminal experiments with weak masking (LL: 33 participants; ML: 36 participants; HL: 38 participants). The study was approved by the local ethics committee.

### Participants of the second part of this research

For the second part of this research we invited 48 new participants, of which 24 were intuitive decision-makers (age 24.54  $\pm$  4.8 years; 23 women) and 24 deliberative decision-makers (age 25.54  $\pm$  4.539 years; 17 women). These participants reported normal or corrected-to-normal visual acuity and a complete absence of previous and current neurological and psychiatric disease.

### Participants of the third part of this research

Twenty-four participants were examined in the medium load experiment with strongly masked clips presented for unconscious encoding (age: 23.58  $\pm$  3.49 years; 11 women) and 24 participants were examined in the medium load experiment with weakly masked clips for conscious encoding (age: 24.00  $\pm$  3.90 years; 20 women). All participants were right-handed according to the Chapman and Chapman inventory.<sup>104</sup> They reported normal or corrected-to-normal visual acuity and a complete absence of previous and current neurological and psychiatric disease. All participants gave their written informed consent before experimentation.

## METHOD DETAILS

The current research consists of three parts that all include the same experimental task, which involves the encoding of cartoon clips and the drawing of temporal relational inferences that need to be stored for later retrieval. The first part of this research includes six behavioral experiments with manipulated information load. The cartoon clips were either presented strongly masked for unconscious encoding with low/medium/high information load or weakly masked for conscious encoding with low/medium/high load. The second part of this research comprises one behavioral replication experiment with strong masking for the unconscious encoding of cartoon clips and a medium information load. The third part of this research includes two experiments with fMRI that feature the encoding of cartoon clips with a medium information load; one fMRI experiment featured strongly masked cartoon clips for unconscious encoding and the other fMRI experiment weakly masked cartoon clips for conscious encoding. In the following, we give methodological details that concern all three parts of this research before we elaborate on each part separately.

## Experimental conditions

### Encoding

In all parts of this research, an experiment contained one experimental condition that required the cognitive capacity to draw temporal relational inferences. Experiments in all parts of this research also contained two control conditions that did not require the drawing of temporal relational inferences. The encoding part of all three conditions comprised the strongly or weakly masked presentation of cartoon clips. These clips featured either five animals (experimental condition and control condition 1) or one animal (control condition 2) that traversed a scene from left to right or from right to left. The five animals in the experimental condition and control condition 1 crossed the scene once, one-by-one, moving through a visually impenetrable hiding place. Because the hiding place was not equipped with peepholes or windows, animals were truly hidden in the hiding place such that participants could not perceive them, neither in the condition with strong masking nor the condition with weak masking. In the experimental condition, an animal could either step through the hiding place and exit it immediately or linger inside the hiding place and meet other animals that had entered before or arrived afterward. Hence, while watching the experimental cartoon clip, participants needed to maintain a mental record of which animals resided simultaneously inside the hiding place based on the temporal sequence of the five animals' entrances and exits from the hiding place. This temporal relational inference task requires working memory. The five animals' entrances and exits from the hiding place were orchestrated such that some but not all animals lingered inside the hiding place. In control condition 1, the five animals crossed the scene one-by-one moving through the hiding place without lingering inside. Hence, no animals in control condition 1 ever hid simultaneously. This circumstance alleviated the need for a temporal relational inference and left the simpler task of encoding five animals and the sequence of their appearance on the screen. Hence, the contrast of the experimental condition with control condition 1 isolates neurocognitive processes underlying the drawing of temporal relational inferences. In control condition 2, a cartoon clip featured one single animal crossing the scene five times in the same direction, moving straight through the hiding place. The control conditions were included in all experiments in all parts of this research because they were designed for the analyses performed on the fMRI data acquired in part three of this research. Please see [Video S1](#) for an impression of the visual experience in each consciousness condition.

### Retrieval

The inferences drawn in the experimental condition needed to be retained for retrieval testing. The encoding-test interval was varied in the first part of this research, along with the number of cartoon clips presented for encoding. A test trial consisted of the unmasked presentation of a clip's hiding place plus the unmasked images of two animals that had featured in the clip. Participants' task was to decide whether the two animals had hidden simultaneously or not in the respective clip (response by manual button press). There were ten retrieval trials per clip because five animals featured in each experimental condition and each control condition 1 and because each animal was presented with every other animal at test. Clips of the experimental condition were orchestrated such



that the average percentage of retrieval trials that call for a ‘simultaneous’ response was 50%. There were also ten retrieval trials in control condition 2, where only one animal had featured in the clip; this one animal was complemented with four new, not-previously-presented animals on the forced-choice test. In the two control conditions, the correct response on the forced-choice test was a ‘not simultaneous’ on each trial because animals had never hidden simultaneously inside the hiding place. When participants were oblivious of clip contents because the clips had been presented strongly masked for unconscious encoding, retrieval instructions fostered an intuitive decision. When participants had been aware of clip contents because the clips were presented weakly masked for conscious encoding, retrieval instructions fostered a conscious recall of inferences drawn when watching the clip. Participants in both consciousness conditions were informed that on average 50% of the retrieval trials call for a ‘simultaneous’ response. Hence, they were encouraged to give ‘simultaneous’ and ‘not simultaneous’ responses evenly. Please refer to [Video S1](#).

### Stimuli

We used 19 cartoon clips that consisted of colored, cartoon-style images of distinct scenes and animals drawn by two professional Swiss artists. The 19 clips pictured unique scenes (underwater scene, mountain scene, ice landscape etc.), of which 18 were used for the experiment and one (the same one for all participants) for practicing before experimentation. Eleven animals were drawn to fit each scene thematically (e.g., underwater scene: ray, orca, octopus, shark, etc.). Each scene was presented twice in the experiment, though in two different conditions (experimental condition, control condition 1, control condition 2) and with distinct sets of five animals. Of the 11 drawn animals, a set of five was used in the experimental condition and another set of five in control condition 1 (counterbalanced across participants), while the eleventh animal was always used in control condition 2. Hence, each animal appeared in only one trial and condition. A clip consisted of 35 images. An image depicted either the scene alone or one single animal crossing the scene. Seven of the 35 images presented the complete trajectory of one animal: three images displayed an animal’s arrival in the scene and three images an animal’s exit; the image of the empty scene was presented for the time when an animal lingered inside the hiding place. We counterbalanced across participants the order of the presentations of clips in a condition, the order of conditions (experimental condition, control condition 1, control condition 2) presented in an experiment, and the assignment of scenes to conditions (experimental condition, control condition 1, control condition 2).

### Clip presentation

For the masked presentation of the clips, we adopted and modified the masking protocol used by Degonda et al.<sup>59</sup> Clip frames (F) were presented for 17 ms and masks (M) for 183 ms. Each clip frame was flanked by masks and repeated three times in sequence before the next clip frame was flashed to continue an animal’s trajectory. Five adjacent masks initiated a clip, and one extra mask terminated a clip. Accordingly, the frame order in a clip was M – M – M – M – M – F1 – M – F1 – M – F1 – M (...) M – F35 – M – F35 – M – M. In the subliminal condition, we used pattern masks consisting of randomly arranged colored pixels to induce a strong masking that impedes conscious frame perception. Because pattern masks interrupt the neural responses to previously flashed images, the processing of the images does not reach consciousness.<sup>105,106</sup> In the supraliminal condition, masks were uniformly gray to induce very weak masking that still allows for the conscious perception of each clip frame. Although clips were clearly visible in the supraliminal condition, they appeared jerkily due to the intervening gray masks. Clips were consciously invisible in the subliminal condition; the conscious impression of a clip consisted in a rapid change of masks. A complete clip consisted of 105 clip frames plus 111 masks and lasted for 22.1 s. Please refer to [Video S1](#) for an impression of the visual experience in each consciousness condition.

### Technical setup

Clips were presented with a BenQ MX764 DLP video projector using a resolution of 1024 × 768 pixels and a screen refresh rate of 60 Hz. Clips were projected onto a rear projection screen (PLEXIGLAS Optical by Evonik Industries) with a viewing angle of 16° width and 9° height. Experiments were programmed using the MATLAB toolbox “Psychophysics Toolbox” Version 3.0.11 (<http://psychtoolbox.org/>, parts 1 and 2 of this research) and using the software Presentation Version 11.3 (Neurobehavioral Systems, <http://www.neurobs.com>, part 3 of this research). Responses were logged using a Cedrus Lumina Response Pad LS-PAIR (<https://www.cedrus.com/lumina>).

### Betsch personality inventory

We accommodated for the participants’ habitual decision style using the Betsch *Preference for Intuition and Deliberation* inventory.<sup>107</sup> Each participant filled out the inventory 24 hours before experimentation. The Betsch inventory consists of two independent subscales that measure the preference to use intuition or deliberation when making decisions. Based on the difference between scores obtained from the two subscales, we assigned participants to a group of intuitive or a group of deliberative decision-makers.<sup>26</sup> The number of deliberative and intuitive decision-makers per condition is provided in [Figure 1](#).

### First part of this research: Six behavioral experiments with manipulated information load

#### Experimental procedure

In the LL, we presented a single cartoon clip for encoding with the ten respective retrieval trials following clip presentation after a few seconds because a clip awareness rating intervened clip presentation and retrieval testing ([Figure 1](#)). The conscious visual impression in the strongly masked condition was of a 22 s-clip displaying a rapid sequence of pixel-masks. In the LL condition, information load was low and long-term memory was at most minimally engaged. In the ML condition, we presented three adjacent cartoon clips

for encoding (one clip of the experimental condition, one clip of control condition 1, and one clip of control condition 2). The conscious visual impression in the strongly masked condition was of three 22 s-clips, each clip displaying a rapid sequence of pixel-masks. A rating of clip awareness was given following each clip. The adjacent presentation of three clips was meant to provide for a moderate information load and to induce a moderate interference between clips during encoding and consolidation. A two-minute rest interval separated the last clip awareness rating and the retrieval testing of the three clips. The ML condition spanned an encoding-test interval of 3.5 minutes, which requires long-term memory (Figure 1). In the HL condition, we presented nine adjacent clips for encoding (three clips of the experimental condition, three clips of control condition 1, and three clips of control condition 2). Within subjects, a fixed sequence of conditions was maintained, e.g., experimental condition (EC), control condition 1 (CC1), control condition 2 (CC2), EC, CC1, CC2, EC, CC1, CC2. Between subjects, the sequence of conditions was varied systematically. The conscious visual impression in the strongly masked condition was of nine 22 s-clips, each clip displaying a rapid sequence of pixel-masks. A rating of clip awareness was given following each clip. The adjacent presentation of nine clips was meant to provide for a high information load and to create substantial interference between clips during encoding and consolidation. A two-minute rest interval separated the last clip awareness rating and the retrieval testing of the nine clips. The HL condition provided for an encoding-test interval spanning at least 6 minutes, which requires long-term memory (Figure 1). The LL condition contained 36 encoding-test runs, the ML condition 12 encoding-test runs, and the HL condition four encoding-test runs.

The strongly (subliminal) and weakly (supraliminal) masked encoding parts were carried out with the same frame timing, same frame sequence, same psychophysical conditions, and the same encoding instructions. The only difference between the subliminal and supraliminal condition consisted in the filling of the masks (pattern mask versus uniform gray mask). However, the wording of retrieval instructions differed between the strongly and weakly masked conditions. When participants were oblivious of clip contents because the clips had been presented strongly masked for unconscious encoding, retrieval instructions fostered an intuitive retrieval response. When participants had been aware of clip contents because the clips were presented weakly masked for conscious encoding, retrieval instructions fostered a conscious recall of the inferences drawn when watching the clip.

Before experimentation, all participants practiced the task with two weakly masked (supraliminal) practice runs to become familiar with the encoding material and encoding requirements. This supraliminal familiarization was also carried out with participants, who were scheduled for an experiment featuring strongly masked (subliminal) clips in order to give them an idea of the encoding material and encoding requirements. If participants were scheduled for an experiment featuring strongly masked clips, the two supraliminal practice runs were followed by two subliminal practice runs to accustom participants to the strongly masked presentation mode.

### **Second part of this research: Replication experiment with medium information load and strong masking for unconscious encoding**

We examined whether the retrieval effect following strong clip masking and a medium information load was replicable before applying this experimental design in the fMRI experiment planned for part three of this research. We ran again the medium load version of the experiment with strong masking using the same protocol and procedure as in the initial experiment but inviting 48 new participants, half of which were intuitive decision-makers, and half were deliberative decision-makers.

### **Third part of this research: fMRI experiments with strongly and weakly masked cartoon clips and a medium information load**

Twenty-four participants were examined in the medium load experiment with strongly masked clips presented for unconscious encoding and 24 further participants were examined in the medium load experiment with weakly masked clips for conscious encoding. Because only habitually intuitive decision-makers had provided evidence of successful subliminal encoding in part one and two of this research, we restricted both fMRI experiments, i.e., the experiment on unconscious clip processing and the experiment on conscious clip processing, to habitually intuitive decision-makers. All participants fulfilled the inclusion and exclusion criteria for MRI. This MRI study was approved by the local ethics committee for human studies (Kantonale Ethikkommission Bern).

### **Experimental procedure**

The MRI session started with a T1-weighted MRI sequence. Next, participants took the practice trials to become familiar with the experimental task and procedure. Then, they performed the fMRI experiment that included 12 fMRI time-series that corresponded to the 12 encoding-test runs. In each run, three clips were presented masked for encoding: one clip of the experimental condition, one clip of control condition 1, and one clip of control condition 2. Each clip constituted one block in a block fMRI design. A two-minute rest interval separated the encoding and test part within a run. The retrieval part of a run encompassed 30 retrieval trials, ten retrieval trials per condition. Retrieval trials were presented in a rapid event-related fMRI design with a random jitter of 500 ms - 3000 ms between retrieval trials. The intervals between retrieval trials were filled with the presentation of a white fixation cross on a gray background that participants were asked to look at.

### **MRI data acquisition**

We ran the anatomical and functional image acquisition on a 3T Siemens Magnetom Trio whole-body scanner (Siemens Medical Solutions, Erlangen, Germany) with a 64-channel head coil. Anatomical T1-weighted image acquisition followed a 3D modified driven

equilibrium Fourier transform pulse sequence (MDEFT<sup>108</sup>) with a spatial resolution of  $1 \times 1 \times 1 \text{ mm}^3$  (176 sagittal slices; time of repetition (TR) = 7.93 ms; echo time (TE) = 2.49 ms; flip angle (FA) =  $16^\circ$ ; field of view (FOV) =  $256 \times 256 \text{ mm}^2$ ). Functional T2\*-weighted multi-slice single-shot images were acquired using a blood-oxygen-level-dependent (BOLD) sensitive, interleaved simultaneous multi-slice accelerated diffusion-weighted echo-planar imaging sequence with an accelerating factor of 3 and a spatial resolution of  $2 \times 2 \times 2 \text{ mm}^3$  (60 transversal slices; TR = 1220 ms; TE = 30 ms; FA =  $80^\circ$ ; FOV =  $192 \times 192 \text{ mm}^2$ ).

## Subjective and objective measures of clip awareness

### Subjective awareness measure

Participants scored the visibility of each cartoon clip immediately after its presentation on a 4-point perceptual awareness scale (PAS;<sup>52</sup>) with the levels: 1) no awareness at all; 2) a feeling that something was present, either static or moving, 3) an impression of the scene or animals, 4) a clear image of the scene and animals.

### Objective awareness measure

An objective test of clip awareness was carried out at the end of a subliminal experimental session that featured strongly masked clips. The objective awareness test was run to find out whether participants were able to consciously discern certain aspects of the strongly masked cartoon clips. We re-used the same cartoon clips and animals that we had subliminally presented in the experiment. There were 36 trials. Over the three conditions in the original experiment, we had presented each of the 18 scenes (e.g., underwater scene, mountain scene, etc.) twice, though with distinct sets of five animals. Now, we re-presented these same 18 scenes in the objective awareness test, presenting each scene twice, to account for the 36 trials. In each cartoon clip, one single animal moved across the scene, entering and leaving the central hiding place once, which amounted to a clip-duration of only 5.3 s. Clips were presented in the same strongly masked fashion and with the same psychophysical set-up as during the experiment. The forced-choice test followed the presentation of the strongly masked clip immediately (no exposure-test delay). For the forced-choice test, participants were presented with the unmasked central hiding place and with two unmasked animals: the target animal and a lure animal that had featured in the original experiment but not in the previously presented clip. Participants' task was to decide which of the two animals had featured in the preceding clip. Of the five animals presented in the original experiment along with a certain scene, we randomly selected two animals, of which one constituted the target and the other the lure. The sequence of trials was randomized anew for each individual participant. Instructions were direct in both this awareness test and in the original experiment because participants knew from the outset that they would be presented with masked clips, whose visibility they were to rate following presentation. The difference between the original experiment and this awareness test consisted in the processing demand (temporal relational inference versus perception of one single animal), the encoding-test delay (minutes versus immediate), and stimulus novelty (novel in the original experiment versus déjà-vu in the awareness test). Given the simplified clip version applied in the awareness test, a lack of behavioral effects of conscious clip awareness would strongly speak to unconscious clip processing in the original experiment.

## Results of the objective awareness test

### Part one of this research

We computed an analysis of variance (ANOVA) with the performance index accuracy on the objective awareness test (i.e., percentage of correctly chosen target animals) and the between-subjects factors Information Load (low load, medium load, high load) and Decision Style (intuitive/deliberative decision-makers). The intercept in the ANOVA was not above chance level (50%;  $F(1, 111) = 0.001$ ,  $p = 0.98$ ) and Decision Style had no significant influence on accuracy ( $F_{DS}(1, 111) = 0.87$ ,  $p = 0.35$ ). We also calculated Bayes Factors (BF) to estimate whether mean accuracy was at chance level and did not differ as a function of Decision Style.<sup>57</sup> BFs were calculated using a half-normal prior distribution with a mode of 0% (reflecting chance-level accuracy or no group-difference) and a standard deviation of 5% (expected above chance accuracy or group difference in accuracy). If the reported effects in the main experiment were due to conscious instead of unconscious processes, accuracy on the awareness tests should match the effect sizes reported for the main experiment. To estimate the expected accuracy, we multiplied the pooled effect size for the difference in reaction times following strong masking in intuitive decision-makers of all experiments ( $d_z = 0.58$ ) with the pooled standard deviation of accuracy (SD = 8.53%) in the objective awareness test. This yielded a value of  $\sim 5\%$ . Bayes Factors suggested that there is substantial evidence for chance level performance ( $M_{Acc} = 49.84\%$ , SE = 0.80%,  $BF_{10} = 0.13$ ), but only anecdotal and inconclusive evidence for the absence of an effect of Decision Style ( $M_{\Delta Acc} = 1.29\%$ , SE = 1.61,  $BF_{10} = 0.64$ ). Post hoc t tests for each decision style group and each level of Information Load against chance (50%) yielded no significant results, and most BFs favored the null assumption of chance-level performance (Table S6).

### Part two of this research (replication experiment with strong masking and a medium information load)

We computed an ANOVA with the dependent variable accuracy on the objective awareness test (i.e., percentage of correctly chosen target animals) and the between-subjects factor Decision Style. The intercept was not above chance level (50%;  $F(1, 46) = 1.133$ ,  $p = 0.29$ ); decision style had no significant influence on accuracy ( $F(1, 46) = 0.071$ ,  $p = 0.79$ ). Bayes Factors suggested that there is substantial evidence for chance level performance ( $M_{Acc} = 48.61\%$ , SE = 1.29%,  $BF_{10} = 0.13$ ), but only anecdotal and inconclusive evidence for the absence of an effect of Decision Style ( $M_{\Delta Acc} = 0.69\%$ , SE = 2.61,  $BF_{10} = 0.56$ ). Post hoc t tests for each decision style group against chance level (50%) yielded no significant results (Table S6), and BFs suggested that there is substantial evidence for the H0 (chance-level performance).

**Part three of this research (fMRI experiment with strong masking, a medium information load, and only participants with an intuitive decision style)**

Accuracy on the objective awareness test (i.e., percentage of correctly chosen target animals) was not significantly above chance (Table S6), but the BF suggested that there is anecdotal evidence for above-chance accuracy.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**Statistical analyses of behavioral data**

Behavioral data were analyzed with R (v4.0.2)<sup>109</sup> in R-Studio (v1.1.447)<sup>110</sup> running on a Windows 7 computer. ANOVAS were performed using the function `aov_ez()` of the R-package `afex` (0.28-0.)<sup>100</sup> Linear mixed models were estimated using the function `lmer()` of the R-package `lme4` (v1.1-23).<sup>101</sup> Significance testing for linear mixed models was performed using the `Anova()` function provided by the R-package `lmerTest` (v3.1-2).<sup>102</sup> The R-package `stats` (4.0.2)<sup>109</sup> provided the functions `t-stats()` for t tests and `cor.test()` for correlations.

Bayes Factors were calculated using a custom R-function provided by Baguley.<sup>55</sup>

All statistical tests were two-sided. The significance threshold was  $p < 0.05$  if not indicated otherwise.

**Control analyses with the behavioral data**

In search of alternative explanations for the results obtained with strongly masked clips, we performed additional analyses that are reported in the following sections.

Intuitive decision-makers displayed shorter reaction times (RT) for correct versus incorrect retrieval responses following their presentation with strongly masked cartoon clips. We used the median rather than the mean of individual RTs because RTs were skewed to the right at the single-trial level. We tried to determine whether biases in the stimulus material or a response bias on the participants' side may have generated shorter reactions in correctly versus incorrectly responded retrieval trials.

**Biased stimulus material?**

The cartoon clips were identical for all participants across all experiments. Hence, the order of entrances and exits of the five animals featuring in a specific cartoon clip was kept constant (e.g., underwater scene with shipwreck: turtle enters, piranha enters, turtle exits, shark enters, ...; see Figure 2). Therefore, the correct retrieval responses for specific animal pairs remained the same across experiments and participants for a given cartoon clip (e.g., the correct response to the question whether the piranha and the shark had hidden simultaneously in the shipwreck was 'simultaneous'; see Figure 2). If participants strongly preferred one response over the other for a subset of animal pairs due to stimulus-inherent properties and if the preferred response happened to be correct and the response latency short, then the significant reaction time difference could be driven by a stimulus bias. To determine whether a stimulus bias contributed to the reaction time difference, we aggregated participants' responses by animal pair (e.g., turtle-piranha, turtle-shark, piranha-shark,...) for the 360 animal pairs. This was done separately for the group of intuitive and the group of deliberative decision-makers. We combined all responses obtained in the first part of this research (low load, medium load, high load) and in the second part of this research (replication experiment with medium load). Responses obtained in the third part of this research (fMRI experiment) were excluded because no deliberative decision-makers took part in the fMRI experiment. We computed the following parameters for each animal pair: 1) Response Accuracy (ratio of participants that gave a correct response); 2) 'Simultaneous'-rate (ratio of 'simultaneous' responses to the question of whether the two animals had hidden simultaneously inside the hiding place); 3) Bias (ratio of participants who gave repeatedly the same response irrespective of whether it was 'simultaneous' or 'not simultaneous'; 0.5 means no bias and 1.0 means all participants gave the same response); 4) Mean of ranked reaction times (RTs were rank-normalized within each participant prior to aggregation to make them comparable between participants). On average, there was a bias toward 'not simultaneous' responses ( $F_{Intercept}(1,359) = 124.94, p < 0.001, \eta_p^2 = 0.258$ ). This bias tended to be larger for responses sampled from deliberative (mean 'simultaneous'-ratio = 44.98%, standard error = 0.52%) than intuitive decision-makers (mean 'simultaneous'-ratio = 46.29%, standard error = 0.51%);  $F_{DS}(1,359) = 3.32, p = 0.069, \eta_p^2 = 0.009$ . Importantly, neither Bias nor 'Simultaneous'-rate were significantly associated with Response Accuracy or with the mean of ranked RT for intuitive and deliberative decision-makers (all  $|r| < 0.094$ , all  $p > 0.079$ ). Linear mixed models with random intercepts for animal pairs confirmed that Decision Style had no significant impact on the influence of Bias or 'Simultaneous'-rate on either Accuracy or RTs (all  $p > 0.285$  for the interaction between Decision Style and Bias or 'Simultaneous'-rate). Therefore, it is unlikely that the RT difference between correct and incorrect retrieval responses observed in the group of intuitive decision-makers is due to biased stimulus material. Nevertheless, we assessed whether the reported RT difference would disappear if biased items were excluded from the data analysis. To test this possibility, we first determined the cumulative binomial probability for obtaining a similar or larger Bias under the null assumption that responses would be equally distributed for each animal pair. A small probability (i.e.,  $< 5\%$ ) indicates that responses were significantly biased for a specific pair. We stepwise excluded biased pairs, first those with a cumulative binomial probability  $p < 5\%$  (i.e., those that were significantly biased at an alpha level of 5%), then all pairs with  $p < 10\%$ ,  $p < 15\%$ ,  $p < 20\%$ , and  $p < 25\%$ . We thus increased the alpha-value with each step, i.e., the level at which items were said to be significantly biased, thereby excluding more and more items with less extreme biases. For each exclusion criterion, we computed and plotted the mean RT-difference plus the 95% confidence interval (see Figure S3). The original values obtained before exclusion of any stimulus pairs are plotted at Exclusion Criterion = 0%. The reported RT difference did not systematically decay, which strongly suggests that the RT differences of intuitive decision-makers at retrieval following strong masking were not due to stimulus-inherent biases.

### Participant-inherent response bias?

If intuitive decision makers would systematically prefer one response over the other (e.g., responding mostly ‘simultaneous’) and if reaction times were systematically shorter for one type of response than the other, then these participant-inherent tendencies might lead to systematic RT differences between correct and incorrect retrieval responses and could hence explain the retrieval results following the watching of strongly masked clips. To test for a participant-inherent response bias, we analyzed participants’ response biases across all experiments with strongly masked clips. The bias was measured as the ratio of ‘simultaneous’ responses to the question whether two animals had hidden simultaneously in a clip. Results showed that the participants were biased toward ‘not simultaneous’ responses (% of ‘simultaneous’ responses was  $M_{BIAS} = 46.17\%$ ,  $SE_{BIAS} = 0.83\%$ ,  $t(116) = -4.604$ ,  $p < 0.001$ ,  $d_Z = -0.426$  for the first part of this research,  $M_{BIAS} = 43.44\%$ ,  $SE_{BIAS} = 1.58\%$ ,  $t(47) = -4.162$ ,  $p < 0.001$ ,  $d_Z = -0.601$  for the second part of this research, and  $M_{BIAS} = 48.65\%$ ,  $SE_{BIAS} = 1.48\%$ ,  $t(23) = -0.912$ ,  $p = 0.371$ ,  $d_Z = -0.186$  for the fMRI experiment). This bias toward ‘not simultaneous’ responses was not significantly influenced by Decision Style ( $F_{DS}(1,111) = 0.37$ ,  $p = 0.542$ ,  $\eta_p^2 = 0.003$ , for the first part of this research;  $F_{DS}(1,46) = 0.01$ ,  $p = 0.928$ ,  $\eta_p^2 < 0.001$  for the second part of this research). Therefore, the significant difference in reaction times between correct and incorrect retrieval responses in intuitive decision-makers is not due to a unique response bias in intuitive decision-makers.

Next, we tried to find out how reaction latencies were influenced by the type of response: ‘simultaneous’ versus ‘not simultaneous’. To this aim, we aggregated per participant the reaction times for Response Accuracy (correct response, incorrect responses) and for Response Type (‘simultaneous’ versus ‘not simultaneous’). In the first part of this research, Response Type had no significant effect on reaction time and did not interact with any other factor (Accuracy, Load, Decision Style, all  $p > 0.12$ ). Importantly, we observed a significant interaction between Accuracy and Decision Style after controlling for Response Type ( $F_{DS \times ACC}(1,111) = 5.00$ ,  $p = 0.027$ ,  $\eta_p^2 = 0.041$ ). Hence, intuitive decision-makers gave correct versus incorrect ‘simultaneous’ and ‘not simultaneous’ retrieval responses faster irrespective of Response Type ‘simultaneous’ or ‘not simultaneous’ ( $F_{ACC}(1,59) = 7.69$ ,  $p = 0.007$ ,  $\eta_p^2 = 0.115$ ). These findings are consistent with the main effect of decision style on RT differences reported in the results section. We obtained similar results for the data collected in the second part of this research (replication experiment): Response Type had no significant influence on reaction time and did not interact with any other factor (Accuracy, Decision Style, all  $p > 0.07$ ). We observed a significant interaction between Accuracy and Decision Style after controlling for Response Type ( $F_{DS \times ACC}(1,46) = 5.67$ ,  $p = 0.022$ ,  $\eta_p^2 = 0.110$ ). Hence, intuitive decision-makers gave correct versus incorrect ‘simultaneous’ and ‘not simultaneous’ retrieval responses faster irrespective of Response Type ‘simultaneous’ or ‘not simultaneous’ ( $F_{ACC}(1,23) = 4.88$ ,  $p = 0.037$ ,  $\eta_p^2 = 0.175$ ). In the third part of this research (fMRI experiment), both Response Type ( $F_{TYPE}(1,23) = 4.39$ ,  $p = 0.047$ ,  $\eta_p^2 = 0.160$ ) and Accuracy ( $F_{ACC}(1,23) = 4.31$ ,  $p = 0.049$ ,  $\eta_p^2 = 0.158$ ) influenced reaction time significantly but these two factors did not interact ( $F_{TYPE \times ACC}(1,23) = 1.02$ ,  $p = 0.324$ ,  $\eta_p^2 = 0.042$ ): Participants responded faster in correct versus incorrect retrieval trials and - independently of this effect - they also responded faster when giving a ‘not simultaneous’ versus a ‘simultaneous’ response. In conclusion, these additional analyses suggest that the intuitive decision-makers’ RT differences between correct and incorrect retrieval trials result from the intuitive decision-makers’ implicit memory of the temporal relational inferences formed when watching the strongly masked clips.

### Statistical analyses of fMRI data

We ran the preprocessing and the statistical analyses of the functional MRI volumes using the software SPM12 (Wellcome Department of Cognitive Neurology, London, UK). Functional volumes were slice-time corrected, realigned to the mean volume by rigid body transformation, coregistered to the participants’ anatomical volume, normalized to the MNI305 T1 template, and finally smoothed with a 6 mm (FWHM) isotropic Gaussian kernel. We performed first-level analyses using a general linear model that modeled the fMRI time-series as a sequence of events convolved with the canonical hemodynamic response function (HRF) provided by SPM12. The encoding time-series was modeled using three regressors of interest (replicated across the 12 encoding sessions) that corresponded 1) to the 12 clips of the experimental condition, 2) to the 12 clips of control condition 1, and 3) to the 12 clips of control condition 2. We modeled the signal changes during the encoding of a clip using a block design with block duration of 22.1 s.

The retrieval time-series was modeled using four regressors of interest (replicated across the 12 retrieval sessions) that corresponded to the responses given at test, namely 1) correct responses given in the experimental condition, 2) incorrect responses given in the experimental condition, 3) all responses given in control condition 1, and 4) all responses given in control condition 2. In addition, the retrieval time-series included 3 regressors of no interest that corresponded to the retrieval cues (cues are images of the hiding places). These cues were provided before retrieval testing to remind participants of the respective clip in order to facilitate their retrieval of inferences formed when watching that clip. We modeled the signal changes during retrieval using a rapid event-related design (duration: 0 s) for regressors of interest. The regressors of no interest referred to blocks of 5 s.

The general linear models computed for both the encoding and retrieval time-series included six nuisance regressors for the six movement parameters plus a constant covariate representing the session-specific mean over scans.

Due to a rapid TR of 1220 ms, combined with a 64-channel head coil, functional images exhibited a drop in signal intensity around the center of the brain. This is a general physical drawback of multichannel acquisitions. The signal collapses toward the center of the brain. Importantly, affected regions of interest (hippocampus and basal temporal lobe) retained a reasonable signal-to-noise ratio of  $> 10$  (hippocampus) and  $> 4$  (basal temporal lobe) in the participant with the largest drop in signal intensity. However, the implicit statistical masking of SPM under default parameters led to the exclusion of large parts of these regions of

interest from the analysis. Therefore, we constructed a custom explicit inclusive statistical mask, based on the implicit mask that replaced missing areas with normalized automated anatomical labeling (AAL) areas. Additionally, white matter was excluded from this custom explicit mask. To this aim, we segmented the average normalized structural (T1) image per group and experiment (experiment with strong masking; experiment with weak masking) using SPM's built-in segmentation function. The resulting white matter tissue map was then inverted and multiplied with the custom explicit inclusive statistical mask to create the final custom explicit mask.

We computed second-level analyses for group statistics using one-sample *t* tests on first-level contrasts. We included one group-level regressor to obtain brain-behavior correlations, which are particularly informative because they isolate brain activity that is directly related to the task-relevant encoding- and retrieval-operations rather than other, non-instructed concurrent mental activity such as rumination. We modeled the encoding and retrieval of strongly and weakly masked clips focusing on four brain-behavior correlations. To model encoding, we used first-level contrasts between the experimental condition and control condition 1. To model retrieval, we used first-level contrasts between correct and incorrect retrieval responses in the experimental condition. As a group-level behavioral regressor, we used retrieval accuracy (% correct responses) in the experimental condition when clips were weakly masked and the RT difference between incorrect and correct retrieval responses in the experimental condition when clips were strongly masked. We set the intensity threshold to a *p* value of 0.001, uncorrected, and the minimal cluster size threshold *k* to 10 voxels.<sup>111</sup> For the region of interest, the hippocampus, we masked voxels inclusively using the SPM Neuromorphometrics Mask for the bilateral hippocampus and set the intensity threshold to a *p* value of 0.005, uncorrected, with a cluster-size threshold *k* of 5 voxels. The reasons for the liberal hippocampal threshold were 1) the *a priori* hypothesis of hippocampal involvement in the memory task and 2) the wish to reduce Type II errors associated with the notoriously weak fMRI signal changes in the hippocampus, which are even weaker in protocols with subliminal presentations.<sup>48,49</sup>

To uncover commonalities in brain activity underlying the conscious and unconscious encoding and retrieval of temporal relational inferences, we computed conjunction analyses against the global null hypothesis.<sup>112</sup> Briefly, we took *T*-maps of the contrasts of interest (positive brain-behavior correlations) and set the threshold at the square root of the *p*-value initially applied to the contrast, i.e.,  $p < \sqrt{.001}$  for whole-brain and  $p < \sqrt{.005}$  for region of interest (hippocampus) analyses. Then, we computed the minimum-*T* image between the newly thresholded *T*-maps for weakly and strongly masked clips to obtain conjunction maps for encoding as well as for retrieval. The resulting conjunction clusters were considered significant at the extent thresholds of *k* = 10 (whole brain) and *k* = 5 (region of interest: hippocampus).

We examined whether the correlations used in the conjunction analyses were of statistically different magnitude when cartoon clips were consciously versus unconsciously processed. Hence, we compared the strength of correlations between consciousness conditions. We focused on hippocampal results obtained when correlating the encoding contrasts with retrieval performance and when correlating the retrieval contrasts with retrieval performance. We used the R package 'cocor'<sup>103</sup> and set the confidence level to 0.95. The null hypothesis predicated that the two correlations are equal (null value = 0).

Outlier control: We assessed whether the computed conjunction analyses of correlations included correlations with outliers (see Figure 4). For outlier control, we ran leverage- and Cook's distance tests using the MATLAB function plotDiagnostics. We chose the default recommended critical maximum values  $L_{crit}$  (leverage test) =  $2 \frac{p}{N}$ , where *p* is the number of coefficients and *N* is the number of

observations.  $D_{crit}$  (Cook's distance test) =  $3 \left( \frac{1}{N} \sum_{i=1}^N D_i \right)$ , where *N* is the number of observations and *D<sub>i</sub>* is Cook's distance of observation *i*. For the correlation of the retrieval contrast with retrieval performance in the unconscious condition, these two tests indicated that the top right data point in Figure 4 was indeed an outlier ( $L = 0.321 > L_{crit} = 0.167$ ;  $D = 2.928 > D_{crit} = 0.455$ ). When we ran this conjunction analysis again excluding this outlier, two of the three hippocampal clusters remained significant. These two clusters were located in the left anterior hippocampus (correlation plotted in Figure 4: peak at MNI = -28, -22, -20;  $t = 1.97$ ,  $p = 0.006$ ,  $k = 26$ ) and in the right posterior hippocampus (shown on sagittal brain slice in Figure 4: peak at MNI = 32, -34, -6;  $t = 1.88$ ,  $p = 0.037$ ,  $k = 5$ ). However, the right anterior hippocampal cluster failed to reach significance (shown on the coronal and sagittal brain slice in Figure 4: peak at MNI = 34, -22, -18).

For the correlation of the encoding contrast with retrieval performance in the unconscious condition, the leverage- and Cook's distance tests disagreed on which data points were potential outliers. The leverage-test classified the same top right data point in Figure 4 as potential outlier ( $L = 0.321 > L_{crit} = 0.167$ ;  $D = 0.114 < D_{crit} = 0.133$ ), while the Cook's distance test classified the top left data point as potential outlier ( $L = 0.09 < L_{crit} = 0.167$ ;  $D = 0.199 > D_{crit} = 0.133$ ). We ran this conjunction analysis again excluding both potential outliers. The right posterior hippocampal cluster depicted in Figure 4 remained statistically significant (peak at MNI = 30, -38, -4;  $t = 2.28$ ,  $p = 0.017$ ,  $k = 10$ ).

Besides the commonalities in brain activity across consciousness conditions, we were also interested in differences in brain activity between consciousness conditions. To identify differences in brain activity underlying the conscious versus unconscious encoding and retrieval of relational inferences, we compared first-level encoding contrasts and also retrieval contrasts between consciousness conditions, i.e., between the encoding of strongly versus weakly masked clips and between the retrieval of strongly versus weakly masked clips. To this aim, we used an independent *t* test at the second level. For encoding, we used first-level contrasts between the experimental condition and control condition 1. For retrieval, we used first-level contrasts between correct and incorrect retrieval responses in the experimental condition.

### Additional fMRI results obtained in the individual experiments

#### **Experiment with strongly masked clips for unconscious encoding and retrieval**

**Encoding with strong masking:** In the encoding fMRI time-series, we used a block design for the presentation of the clips and contrasted the experimental condition with control condition 1. We then correlated this contrast with the retrieval performance (i.e., RT difference between incorrect and correct retrieval responses) between subjects to reveal BOLD signal that is associated with unconscious temporal relational inference and associated with retrieval success. This correlation yielded only two significant results and both results were located in the right inferior frontal gyrus (BA47; 51 voxels, peak at MNI 38, 46, -2;  $T = 4.77$ ,  $p_{\text{uncor}} < 0.001$ ; BA47; 10 voxels, peak at MNI 50, 40, -16;  $T = 3.85$ ;  $p_{\text{uncor}} < 0.001$ ) (Table S2). There was no significant result for the hippocampus. We would like to mention that two hippocampal activity increases appeared below threshold – one in the left anterior hippocampus (5 voxels, peak at MNI -26, -14, -24;  $T = 2.29$ ;  $p_{\text{uncor}} = 0.016$ ) and the other in the right posterior hippocampus (3 voxels, peak at MNI 30, -38, -4;  $T = 2.10$ ;  $p_{\text{uncor}} = 0.024$ ). This is interesting because these same hippocampal areas increased their activity significantly during the unconscious retrieval of inferences and/or appeared in the conjunction analyses.

The inverse correlation yielded no significant results (Table S2).

**Retrieval following strong masking:** In the retrieval fMRI time-series, we used a rapid event-related design that allowed contrasting correct versus incorrect retrieval responses given in the experimental condition. We computed the contrast of these correct versus incorrect retrieval responses and correlated this contrast with retrieval performance (i.e., RT difference between incorrect and correct retrieval responses) between subjects. Positive correlations were situated in the right inferior frontopolar cortex (BA10; 29 voxels, peak at MNI 26, 60, -8;  $T = 5.38$ ,  $p_{\text{uncor}} < 0.001$ ), the left posterior cingulate gyrus (BA23; 26 voxels, peak at MNI -6, -32, 34;  $T = 5.87$ ,  $p_{\text{uncor}} < 0.001$ ), and in the left anterior hippocampus (5 voxels, peak at MNI -30, -16, -22;  $T = 3.39$ ,  $p_{\text{uncor}} < 0.005$ ) (Table S2). The inverse correlation yielded no significant results (Table S2).

#### **Experiment with weakly masked clips for conscious encoding and retrieval**

**Encoding with weak masking:** In this encoding fMRI time-series, we used a block design for the presentation of the clips and contrasted the experimental condition with control condition 1. We then correlated this contrast with retrieval performance (i.e., retrieval accuracy) between subjects to reveal BOLD signal that is associated with conscious temporal relational inference and associated with retrieval success. This correlation yielded several significant results: in the right middle temporal gyrus (BA21; 38 voxels, peak at MNI 50, 14, -32;  $T = 4.32$ ,  $p_{\text{uncor}} < 0.001$ ); in the left anterior thalamus (11 voxels, peak at MNI -8, -16, 18,  $T = 4.73$ ,  $p_{\text{uncor}} < 0.001$ ); in the right pulvinar nuclei of the thalamus (10 voxels, peak at MNI 24, -32, 4,  $T = 3.97$ ,  $p_{\text{uncor}} < 0.001$ ); in the right intraparietal sulcus (BA40; 23 voxels, peak at MNI 32, -32, 46,  $T = 4.01$ ,  $p_{\text{uncor}} < 0.001$ ); and in the right posterior hippocampus (17 voxels, peak at MNI 24, -36, 4;  $T = 3.31$ ,  $p_{\text{uncor}} < 0.005$ ) (Table S3). The inverse correlation yielded no significant results (Table S3).

**Retrieval following weak masking:** In this retrieval fMRI time-series, we used a rapid event-related design that allowed contrasting correct versus incorrect retrieval responses given in the experimental condition. We computed the contrast of these correct versus incorrect retrieval responses and correlated this contrast with retrieval performance (i.e., retrieval accuracy) between subjects. Significant positive correlations were situated in many left and right lateral temporal semantic storage sites (e.g., left inferior temporal gyrus, BA20; 132 voxels, peak at MNI -60, -20, -34;  $T = 6.52$ ; right middle temporal gyrus, BA20/21; 211 voxels, peak at MNI 56, -35, -24;  $T = 5.83$ ;  $p_{\text{uncor}} < 0.001$ ). Additional positive correlations were located in the left middle frontal gyrus (BA8; 37 voxels, peak at MNI -18, 34, 38;  $T = 4.24$ ,  $p_{\text{uncor}} < 0.001$ ), in bilateral areas of the insula, in bilateral areas of the superior occipital gyrus (BA19) and in the right posterior hippocampus (6 voxels, peak at MNI 32, -38, -6;  $T = 3.23$ ,  $p_{\text{uncor}} < 0.005$ ) (for complete results see Table S3). The inverse correlation yielded no significant results.