Coding of Images in Long Term Memory

Annabelle Goujon, Simon Thorpe

Coding of Images in Long Term Memory: The Fate of Visual Memories across Weeks in Adults and Children

Annabelle Goujon (annabelle.goujon@univ-fcomte.fr)
University of Franche-Comté

Simon Thorpe
French National Centre for Scientific Research

Research Article

Keywords:

Posted Date: March 1st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1327360/v1

Abstract

What is the content and the format of visual memories in Long Term Memory (LTM)? Is it similar in adults and children? To address these issues, we investigated, in both adults and 9-year-old children, how visual LTM is affected over time and whether visual vs semantic features are affected differentially. In a learning phase, participants were exposed to hundreds of meaningless and meaningful images presented once or twice for either 120ms or 1920ms. Memory was assessed using a recognition task either immediately after learning or after a delay of three or six weeks. The results suggest that multiple and extended exposures are crucial for retaining an image for several weeks. Although a benefit was observed in the meaningful condition when memory was assessed immediately after learning, this benefit tended to disappear over weeks, especially when the images were presented twice for 1920ms. This pattern was observed for both adults and children. Together, the results call into question the dominant models of LTM for images: although semantic information enhances the encoding & maintaining of images in LTM when assessed immediately, this seems not critical for LTM over weeks.

Introduction

How are the landscapes of your last trip, the layout of the bedroom in which you grew up, the face of your teacher when you were eight years old seared into your memory? How are images from unique visual episodes encoded, then consolidated to emerge as memories or recycled in the construction of new percepts? Studying the formation and the consolidation of sensory memories raises the problem of the content and format of such memories in Long Term Memory (LTM). In this respect, the present study aimed at investigating how visual LTM is affected by time and whether visual features vs semantic/conceptual information in visual LMT are affected differently.

In a closely related field, the literature on mental imagery has traditionally opposed two main classes of hypotheses to account for the coding of images in LTM. The first refers to the propositional position, which assumes that symbolic codes are used for LTM (for reviews [1,2]). These codes represent something conceptual and sometimes arbitrary as opposed to perceptual. In this view, coding in memory would be a sentence-like description of the image. By contrast, the functional-equivalency hypothesis supposes that the coding of images in memory has the same structure as the information being represented [3–5]. In this view, symbolic codes are not required to account for LTM. At the interface, the dual-code theory assumes that both analogue (or perceptual codes), and arbitrary symbols or verbal codes are used when retrieving representations of pictures from memory [6,7].

Questions about the content and format of visual memories have also been addressed in the field of the perception of visual scenes through research aimed at assessing both the capacity of visual LTM and the fidelity of our representations of visual stimuli. In the 1960s and 70s, research using large scale memory procedures revealed that people have an extraordinary capacity to remember thousands of images presented for only a few seconds each [8,9]. These studies concluded that the number of visual items that can be stored in LTM is potentially unlimited, that such memories last for at least several days, and that

memory performance depends primarily on the distinctiveness between the target stimulus and the concurrent stimulus (foil stimulus) in the memory task (e.g., recognition)[10]. Nonetheless, because of the substantial visual and semantic heterogeneity between the used stimuli, those studies did not provide relevant information regarding the coding of visual memories into LTM.

Three decades later, this issue received renewed interest following research reporting the phenomena of change blindness and inattentional blindness[11]. The dramatic inability to detect even massive changes in the visual input led many authors to claim that memory representations for real-world stimuli are impoverished, sparse, volatile and lack visual details [12–16]. Influential theories in the early 2000s postulated that representations in visual LTM are gist-like and semantic in nature (e.g.[17]). This position was later examined and undermined. The ability of participants to detect changes when they are tested with either forced-choice paradigms or with longer exposures provided strong evidence that visual episodes leave a more complete memory trace that includes "visual" (or perceptual) information and not just the gist[18]. Large scale memory studies have subsequently strongly supported this conclusion, showing the massive capacity to store visual details from objects or scenes in visual LTM (for reviews, [19,20]). For instance, participants initially exposed to 2500 objects for 3 seconds performed at 92% in a two-forced choice recognition task when the target and the foil stimulus belonged to a different category, 88% when they belonged to the same basic-level category and 87% when the same object was presented in a different state or pose[21].

Recent research aimed at determining what makes an image memorable suggests, nonetheless, that high-level properties, such as distinctiveness, atypicality, emotional valence and semantic attributes strongly contribute to its memorability. In contrast, low-level image properties, such as the salience, color or other simple image features make relatively weak contributions[22–24]. While objects without semantics might not be effective at predicting memorability, the presence of semantic labels associated with objects or photographs could improve it. For example, the possibility to provide a single label for each image (i.e. a single gist) might explain most of what makes an image memorable[22]. Scene semantics would therefore be a primary substrate of memorability.

Thus far, most models and theories of VLTM give more weight to conceptual features than perceptual features in the coding used to retrieve visual representations in memory[25–29]. "*Being perceptually rich and distinctive might be not sufficient to support VLTM*. (…) *VLTM representations are hierarchically structured, with conceptual or category specific features at the top of the hierarchy and perceptual or more category-general features at lower levels of the hierarchy"* (Brady et al., 2011, p19[19]). According to Mary Potter (2012a, p1[28]), "*although some specific visual information persists, the form and content of the perceptual and memory representations of pictures over time indicate that conceptual information is extracted early and determines most of what remains in LTM".*

However, in most studies on visual LTM, the contents of memory were examined either immediately after learning or the next day. Thus, the question of how memories for images evolve over time remains

unanswered. Yet this issue is crucial to determine how visual representations are transformed and consolidated into visual memories. In this framework, the goal of this study was to examine how visual and semantic features were affected by delays and to test whether the hypothesis according to which "*conceptual information is extracted early and determines most of what remains in LTM"* extends to memories that persist beyond several weeks. This hypothesis was examined in both adults and nine-year-old children.

The literature on memory development across the life span suggests large developmental differences in many aspects of memory, especially working memory [30] and declarative memory [31,32]. Nonetheless, visual recognition memory is usually thought to be an early emerging form of memory, which can be measured from the first months of life[33]. Using an abbreviated version of the materials developed by Brady et al. (2008), Ferrara, Furlong, Park, and Landau[34] reported impressive visual memory performance by four-year-old children, both in terms of the large number of items and the level of details required for recognition. Although the number of images was substantially less than in the experiments conducted in adults, the patterns of results were similar. However, to our knowledge very few studies, if any, have examined how memory for images evolved over weeks and whether this evolution differed across the development.

In this framework, in two experiments we investigated how the recognition of images evolves over time, depending on whether they were meaningful or meaningless, in both adults (Experiment 1) and nine-year-old children (Experiment 2). The methods were similar in both experiments. In a learning phase, participants were exposed to hundreds of meaningless and meaningful images (Fig. 1). Because most models on visual memory were based on research using *Rapid Serial Visual Presentation* (RSVP) procedures or *large scale memory* procedures (for examples that combines both procedures, see [35,36]), two exposure durations were examined. Indeed, based on this literature, exposure duration seems to have different impact on memory performance and specifically on the extraction of visual vs. semantic features. Thus, change blindness might be due to a lack of encoding time or attention to each object instead of memory limitations for visual details[37]. The images were thus presented for either 120ms or 1920ms during the learning phase. We also examined the impact of another factor that potentially plays a critical role in memorization, that is, the repetition of the images. Thus, the images were presented either once or twice during the learning phase.

Immediately after the learning phase, or after a delay of three weeks or six weeks, the memory of the participants was assessed through a recognition task that included old and new meaningless and meaningful images. Among the new meaningful images, some belonged to a basic-level category not used during the learning phase (novel-gist images), and some belonged to a basic-level category that had already been used during the learning phase (old-gist lures). This is illustrated in Fig. 2.

Participants were first asked to judge whether the image was old or new and then to indicate how confident they were in their response using a 4-point confidence scale ("Confidence? 1 = just guessing, 2 = not sure, 3 = confident, 4 = very sure). Collecting those confidence ratings aimed at determining the most

relevant measure to compare meaningful and meaningless conditions, given potentially different response biases in the meaningful and meaningless conditions[38]. An examination of receiver operating characteristic curves (ROC), derived from signal detection theory (SDT) should help to provide the best model to apply to our data[39].

The hypothesis that conceptual information is extracted earlier and determines most of what remains in LTM[28] leads to three predictions: 1) For very brief exposures, only meaningful images should be accessible to recognition; 2) Meaningless images should be more subject to forgetting over weeks than meaningful images; 3) False recognition for the “old-gist lures” images should increase over weeks; 4) In view of the literature on children visual memory, we expected similar patterns of results on nine-year-old children and in adults, with nevertheless lower performance [33,34].

Experiment 1

Experiment 1 aimed to explore in adults how the recognition of images evolves over time, depending on whether they were meaningful or meaningless.

Results And Discussion

The hits (i.e., when the image is old and the participant's response is old) and the false alarms (FA, i.e., when the image is new and the participant's response is old) observed in the recognition task depending of the type of images, the exposure duration, the number of exposures and the delay are shown in Supplementary materials, Table 1. The ROC curves in each condition derived from the confidence ratings are shown in Supplementary materials, Fig. 1. Examination of the zROC (which corresponds to z scores of hits and FA plotted as coordinates) revealed a slope almost always different than 1, suggesting Gaussian distributions of unequal variance in responses. Therefore, recognition accuracy was calculated using the discriminability measure of d_a [38]. Each d_a was computed separately from the false-alarm and hit rates for each subject and for each type of image (meaningless vs. meaningful) and exposure condition (120 vs. 1920ms and 1 vs. 2 exposures) and was corrected by the slope of the zROC in each condition. The d_a is calculated as follows:

$$d_a = \left(\frac{2}{1+s^2}\right)^{1/2}\left(zH - szF\right)$$

where s corresponds to the zROC slope, zH to the z scores of the hits and zF to the z-scores on the FA. The d_a values are illustrated Fig. 3.

To minimize both Type I and Type II errors, we conducted several analyses instead of only one that would include the all factors. In a first analysis, we explored the impact of the number of exposures and of the exposure duration on memory. A repeated-measures ANOVA was conducted on the d_a with Delay (immediate, 3-weeks, and 6-weeks) as between-subject factors, with Exposure duration (1920ms vs. 120ms) and Number of exposures (1 vs. 2) as within-subject factors. The analysis revealed a main effect of each factor: Delay, $F(2,33) = 26.535$, $p < .001$, $\eta^2_G = 0.55$, Exposure duration, $F(1,33) = 329.608$, $p < .001$, $\eta^2_G = 0.49$; Number of exposures, $F(1,33) = 137.536$, $p < .001$, $\eta^2_G = 0.29$. The interactions [Number of exposures x Exposure duration, $F(1,33) = 64.069$, $p < .001$, $\eta^2_G = 0.084$], [Delay x Exposure duration, $F(2,33) = 8.183$, $p < .001$, $\eta^2_G = 0.046$] and [Delay x Number of exposure x Exposure duration, $F(2,33) = 4.020$, $p < .05$, $\eta^2_G = 0.01$] were all reliable. The interaction [Delay x Number of exposures, $F(2,66) < 1$, $\eta^2_G = 0.002$] was not reliable.

Those results show that 1) memory for images decayed strongly over weeks; 2) memory benefited from multiple and extended exposures; 3) multiple and extended exposures had a potentiating effect on memory; 4) the benefit of an extended exposure was even more pronounced across weeks.

In a second series of analyses, we compared how memory for meaningless vs. meaningful images evolved across the weeks separately within each exposure condition, i.e. 120ms-1exposure, 120ms-2exposures, 1920ms-1exposure, 1920ms-2exposures. We carried out four repeated-measure ANOVA on the d_a with Delay (immediate, 3-weeks, and 6-weeks) as between-subject factor, and Type of images (meaningful vs. meaningless) as within-subject factor.

The analysis of the condition “120ms − 1exposure” showed a main effect of Delay, $F(1,33) = 13.776$, $p < .001$, $\eta^2p = 0.455$, Type, $F(1,33) = 29.349$, $p < .001$, $\eta^2p = 0.471$, and a reliable interaction between both factors, $F(2,33) = 10.863$, $p < .001$, $\eta^2p = 0.397$. Consistently with our prediction, the results revealed a strong impact of semantic information when memory was tested immediately after learning. Nevertheless, they also suggest that almost nothing remains in memory after three weeks, whether the images were meaningless or meaningful. Note that the reliable interaction was probably due to the fact that memory performance for the meaningless images had dropped to chance level across weeks.

The ANOVA conducted on the d_a observed in the condition “120ms − 2 exposures” yielded a main effect of Delay, $F(1,33) = 19.307$, $p < .001$, $\eta^2p = 0.539$ and Type, $F(1,33) = 58.296$, $p < .001$, $\eta^2p = 0.639$. The interaction between both factors was marginally significant, $F(2,33) = 3.118$, $p = .058$, $\eta^2p = 0.159$. Again, the results show a strong impact of semantic information on memory and a strong impact of the delay in both conditions. For the meaningful images only, a second exposure had allowed to maintain some memories for at least 6 weeks.

For the condition “1920ms − 1exposure”, the analysis revealed a main effect of Delay, $F(1,33) = 35.365$, $p < .001$, $\eta^2p = .682$ and Type, $F(1,33) = 25.910$, $p < .001$, $\eta^2p = .440$, but no reliable interaction between both factors, $F(2,33) = 1.895$, $p = .166$, $\eta^2p = .103$. The analysis conducted in the condition “1920ms − 2exposures” yielded a main effect of Delay, $F(1,33) = 17.719$, $p < .001$, $\eta^2p = .519$ and Type, $F(1,33) =$

8.119, $p < .01$, $\eta^2p = .197$, as well as a reliable interaction between factors, $F(2,33) = 6.503$, $p < .005$, $\eta^2p = .283$. Overall, the results observed for longer exposures suggest a strong effect of semantic information when memory was accessed immediately after learning, but this benefit tended to disappear across weeks. Contrary to the 120ms-exposure conditions, the reliable interaction was not due to the fact that memory performance approached a floor effect over weeks. After 6-weeks, there was no evidence of benefit for the meaningful images as compared to the meaningless ones when there were presented twice. This result goes against our initial prediction

False alarms

The false alarms are illustrated in Table 1. The FA observed in the novel-gist meaningful and meaningless conditions were compared using a repeated-measures ANOVA. This analysis showed a marginal effect of Type, $F(1,33) = 4.105$, $p = .051$, $\eta_p^2 = 0.11$, a reliable effect of Delay, $F(1,33) = 6.659$, $p < .005$, $\eta_p^2 = 0.288$ and a reliable interaction [Type x Delay, $F(2,66) = 4.201$, $p < .05$, $\eta_p^2 = 0.20$]. Those results indicated that the false alarms increased more across weeks for meaningful images than for meaningless images.

To examine memory distortion regarding the new images that depicted a category used in the learning phase, a series of paired samples t-Test was conducted between the novel-gist images and the old-gist lures (see Table 1). Those analyses showed that the false alarms were higher for the old-gist lures than for the novel-gist images in the conditions [Immediate, 120ms-1exposure, t = 2.65, $p < .05$] and [Immediate, 120ms-2exposures, t = 2.46, $p < .05$] and that the false alarms were marginally higher for the novel-gist images in the conditions [6-weeks, 120ms-2exposures, t = 2.081, $p = .062$] and [6-weeks, 1920ms-2exposures, t = 2.196, $p = .051$]. We also conducted Bayes Factor analyses for any null statistical outcomes to evaluate the degree of evidence for the null versus alternative hypothesis. Specifically, we computed Bayes Factor10, which indicates the likelihood ratio of evidence given both the null and alternative hypotheses (e.g., BF_{10} = (likelihood of data given H_1 / likelihood of data given H_0)). Thus, the outcome value for BF_{10} indicates the likelihood of the data to occur in the alternative compared to the null hypothesis. Importantly, the outcomes of Bayes Factor analyses are considered on a continuous scale reflecting the degree of evidence for the null versus alternative hypothesis. Those Bayes factors analyses (see Table 1) confirmed the conclusions of paired samples t-Test, except for the conditions [6-weeks, 120ms-2exposures, t = 2.081, $p = .062$] and [6-weeks, 1920ms-2exposures, t = 2.196, $p = .051$], for which a difference between the novel-gist images and the old-gist lures was privileged compared to the null hypothesis. Therefore, in regard to the Bayesian statistics, the old-gist lures triggered more false alarms in the immediate conditions (when the images depicting a related gist were presented briefly in the learning phase). However, this effect disappeared across weeks. Surprisingly, the opposite patterns tended to emerge at 6-weeks: the old-gist lures for images that were presented twice generated less false alarms than the new images that depicted a novel gist. Though the last statistics remain weak to firmly accept this hypothesis, they nevertheless allow us to reject the hypothesis that the false alarms increase more for the old-gist lures than for the novel-gist images. Thus, contrary to our prediction, the false

alarms for the old-gist lures did not increase more across weeks than the false alarms on new images depicting an original gist.

Experiment 2

Experiment 2 aimed to explore in nine-year-old children how the recognition of images evolves over time, depending on whether they were meaningful or meaningless. We did not have strong predictions regarding this issue, but because perceptual recognition memory is usually supposed to be mature early in development, we expected similar patterns of results that those observed in adults. The method and the analyses conducted were both similar to those used in Experiment 1.

Results & Discussion

The hits and the false alarms observed in the recognition task depending of the type of images, the exposure duration, the number of exposures and the delay are shown in Supplementary materials, Table 2. The ROC curves in each condition derived from the confidence ratings are shown in Supplementary materials, Fig. 2. As in adults, because examination of the zROC suggests Gaussian distributions of unequal variance in response, recognition accuracy was calculated using the discriminability measure of d_a. The d_a values are illustrated Fig. 4. The analyses we conducted were similar to those conducted in Experiment 1.

First, a repeated-measures ANOVA was conducted on the d_a with Delay (immediate, 3-weeks, and 6-weeks) as between-subject factors, and Exposure duration (1920ms vs. 120ms) and Number of exposures (1 vs. 2) as within-subject factors. The analysis revealed a main effect of each factor: Delay, $F(2,33) = 19.898$, $p < .001$, $\eta^2_G = 0.38$, Exposure duration, $F(1,33) = 144.330$, $p < .001$, $\eta^2_G = 0.32$; Number of exposures, $F(1,33) = 93.851$, $p < .001$, $\eta^2_G = 0.118$. The interactions [Number of exposures x Exposure duration, $F(1,33) = 26.928$, $p < .001$, $\eta^2_G = 0.045$], [Delay x Exposure duration, $F(2,33) = 17.081$, $p < .001$, $\eta^2_G = 0.1$] and [Delay x Number of exposures, $F(2,66) = 3.788$, $p < .05$, $\eta^2_G = .01$] were all reliable. The interaction [Delay x Number of exposure x Exposure duration, $F(2,33) < 1$, $\eta^2_G = 0.003$] was not reliable.

The pattern of results was close to those observed in adults. In sum, multiple and extended exposures had a potentiating effect on memory and delayed the forgetting in memory. We can nevertheless mention that the children had poorer memory performance than the adults.

In a second series of analyses, four repeated-measure ANOVA on the d_a , with Delay (immediate, 3-weeks, and 6-weeks) as between-subject factor, and Type of images (meaningful vs. meaningless) as within-subject factor.

The analysis of the condition “120ms − 1exposure” revealed a main effect of delay, $F(1,33) = 5.971$, $p < .05$, $\eta^2p = 0.266$, Type, $F(1,33) = 17.639$, $p < .001$, $\eta^2p = 0.348$ and a reliable interaction between both factors, $F(2,33) = 6.304$, $p < .005$, $\eta^2p = 0.276$. Likewise, the ANOVA conducted on the d_a observed in the

condition “120ms − 2 exposures” showed a main effect of delay, $F(1,33) = 15.031$, $p < .001$, $\eta^2 p = 0.48$ and Type, $F(1,33) = 38.742$, $p < .001$, $\eta^2 p = 0.540$, as well as a reliable interaction, $F(2,33) = 8.025$, $p < .001$, $\eta^2 p = 0.327$. This pattern of results is closed to this observed in adults. There was a strong benefit of semantic information when memory was accessed immediately after learning, but nothing remains after six weeks when the images were presented once, whether the images were meaningless or meaningful. Although semantic information clearly helped encoding in memory, memories strongly worsened across weeks.

The ANOVA conducted in the condition “1920ms − 1exposure”, yielded a main effect of delay, $F(1,33) = 30.229$, $p < .001$, $\eta^2 p = .647$ and Type, $F(1,33) = 98.881$, $p < .001$, $\eta^2 p = .750$, as well as a reliable interaction between both factors, $F(2,33) = 30.229$, $p < .001$, $\eta^2 p = .43$. For the condition “1920ms-2exposures”, the analysis revealed a main effect of delay, $F(1,33) = 17.623$, $p < .001$, $\eta^2 p = .516$ and Type, $F(1,33) = 5.653$, $p = .007$, $\eta^2 p = .382$ and a marginally reliable interaction between factors, $F(2,33) = 3.186$, $p = .054$, $\eta^2 p = .162$. Like in adults, the main conclusion of those results goes against our initial prediction. The strong benefit of semantic information when memory is accessed immediately after learning tended to disappear across weeks, especially when the images were presented twice.

False alarms

The false alarms are illustrated in Table 2. The false alarms observed with the novel-gist meaningful images and those observed with the novel meaningless images were compared using a repeated-measures ANOVA. This analysis showed an effect of Type, $F(1,33) = 5.428$, $p < .05$, $\eta_p^2 = 0.141$, no reliable effect of Delay, $F(1,33) = 1.198$, $p = .314$, $\eta_p^2 = 0.068$, but a reliable interaction [Type x Delay, $F(2,66) = 4.337$, $p < .05$, $\eta_p^2 = 0.208$]. Like in adults, those results indicated that the false alarms increased more across weeks for the meaningful images than for the meaningless images.

A series of paired samples t-Test conducted between the novel-gist meaningful and the old-gist lures conditions showed that the false alarms were higher for the old-gist images in the conditions [Immediate, 120ms-2exposures, $t = 2.39$, $p < .05$] only (see Table 2). However, the outcome value for BF_{10} privileged a condition effect to the null hypothesis in the conditions: [Immediate, 120ms-1exposure, $BF_{10} = 2.04$], [Immediate, 120ms-2exposures, $BF_{10} = 2.02$], as well as in the condition [3-weeks, 1920ms-1exposure], although this last ratio was very weak ($BF_{10} = 1.09$). In conclusion, as in adults, although there were more false alarms for the old-gist lures than for the novel-gist images when memory was tested immediately after learning, the results suggest that they did not increase more across weeks.

General Discussion

The purpose of the present study was to provide insight regarding the format and the content of the representations of pictures in visual LTM. More specifically, by examining how memory for meaningless and meaningful images evolved across weeks, we tested the hypothesis that conceptual information is extracted earlier and determines most of what remains in LTM[25–27]. Because the literature on visual

memory used to report memory performance on very brief and longer exposures to the stimuli, the images were presented for either 120ms or 1920ms. Moreover, because we expected that one exposure may be not enough to maintain an image across weeks in memory, the images were presented once or twice. The hypothesis we examined lead to three predictions: 1) For very brief exposures, only meaningful images should be accessible to recognition; 2) Meaningless images should be more forgotten over weeks than meaningful images; 3) False recognition for the “old-gist lure” images should increase more over weeks than the “novel-gist” images; 4) Similar patterns were expected for both children and adults.

In concordance with the first prediction, for 120ms exposures, the recognition indexes (d_a) were much better for the meaningful images than for the meaningless images. This suggests that indeed, for brief exposures, conceptual information considerably enhances recognition memory[27,28,40]. Nonetheless, whether the images were meaningless or meaningful, they tended to be dramatically forgotten over weeks. Although a second exposure enhanced memory and then reduced the decay for meaningful images, it seems that two brief exposures are not sufficient to maintain a memory for a very long term. This decay was even more pronounced for the children than for the adults, reaching quickly chance level. Our results contrast nevertheless with the RSVP literature suggesting that with presentations shorter than around 250ms, only the gist is retained in LTM[29]. Indeed, the performance in the meaningless condition was above chance level when the testing phase was presented immediately after learning or three weeks later, showing that 120ms of exposure is sufficient to maintain much more than the gist in LTM, at least in adults. It is also noteworthy that in preliminary experiments using a similar procedure, we even observed a learning effect for meaningless images presented for only 30ms (see also[35,36]).

The second aspect of the results concerns memory for longer exposures. Again, the results show a strong benefit for the meaningful images when memory was assessed immediately after learning. They also show how a second exposure considerably enhances recognition memory and delays the decay in memory. There was indeed a reliable interaction between the factor Delay (Immediate, 3-weeks and 6 weeks) and Type (meaningful vs. meaningless), but of interest, this interaction was in the opposite direction to what we predicted [29]. As a result, at six weeks, there was no longer any benefit for the meaningful images presented twice as compared to the meaningless images, suggesting that the semantic facilitation disappeared over weeks. This pattern of results was observed in both adults and children. This thus fails to validate the prediction that meaningless images should be forgotten more easily over weeks than meaningful images. An unpublished experiment conducted in our laboratory revealed a similar pattern of results with four-year-old children exposed to an abbreviated version of the materials.

The third prediction was related to the false alarms for the meaningful “old-gist” lures with respect to the meaningful “novel-gist” images. When the recognition took place immediately after learning, the false recognition for the novel images that belonged to a basic-level category already used in the learning phase was above the false recognition observed with novel categories (Novel-gist condition), but only for very brief presentations. Similar patterns were observed in both adults and children. This suggests that,

indeed, the gist were used in the retrieval of memory when it was assessed immediately after learning. However, essentially nothing remained of those false memories after three weeks, regardless of the duration or the number of exposures. Again, this result goes against to our initial prediction.

The last prediction was related to the age effects on images memory. In congruence with our initial prediction, although memory performance was weaker in nine-year-old children than in adults, the global pattern was very similar. The children's capacity to form and retrieve visual representations demonstrates the existence of a visual memory system very similar to the visual memory system of adults. We can nevertheless note that children were even more inclined to forget images presented very briefly than adults. Several reasons are likely to explain this result. This might be the signature of a kind of immaturity of the attentional, working memory, or declarative memory systems. It can be noted that they also had much more difficulty using all the panel of the confidence scale. A simpler scale would be more appropriate for a young population.

Together, the results obtained in the present study call into question the models of VLTM for images that assume that conceptual information determines most of what remains in LTM, e.g.[25,26,28,28,41]. Though conceptual/semantic information and even linguistic labels enhance the encoding & maintaining of representations in LTM considerably, through a dual-coding for example[6,7], semantic codes or even the gist do not seem to be what remains primarily in LTM over weeks. By contrast, VLTM has a strong capacity to store visual features of images, even independently of pre-existing conceptual features, provided that the exposure is long enough and repeated. In addition, memory for visual information contained in images seems to be more robust over time than memory for conceptual information that would be independent of visual features, as suggested by the result that the false alarms did not increase more in the old- gist lure condition than in the novel-gist condition.

However, this study shows also that interfering effects and false memories constitute a problem when investigating recognition memory[42]. In congruence with the literature on *memory distortion*, false alarms were higher in the meaningless condition when memory was tested immediately after learning, but this effect tended reversed over weeks. As a result, the stronger impact of the delay on the d_a in the meaningful condition as compared to the meaningless condition (for extended and multiple exposures) was not due to a stronger impact on the hits (i.e. impact on decay) but to a stronger impact on the false alarms (i.e. impact on interference).

The present results raise several issues. First, what makes an image memorable over time[22,43]? This study shows a potentiating effect of repetition and exposure duration on memory over weeks, and suggests that multiple and extended exposures are probably required to maintain an image in LTM over time. Second, the present results highlight important changes in memory effects across the weeks, with a reduction of the facilitating effect of the meaningful cues in the repeated and prolonged exposure condition. Thus, we hypothesize that multiple and prolonged exposures, the uniqueness of an image, as well as its distinctiveness relative to what is already in memory are good predictors of which images will be sensitive to long term recognition [44].

Second, the present study raises the question of how the different kinds of consolidation mechanisms (synaptic vs systemic), as well as how the different memory/processing systems interact during the encoding and the consolidation of visual memories [45,46]. Classical theories of memory (e.g.[47–49]) used to propose a distinction between explicit/declarative and implicit/nondeclarative memory systems. In this respect, a hypothetical sketch is that memory of images results from interactions between different memory “systems”. An “integrating system”, usually associated with explicit/declarative memory, might play a critical role in the integration and association of distributed sensory and conceptual information. The hippocampus might be a good candidate for such integration and memory formation. This system would underlie VLTM that is strongly enhanced by the retrieving of semantic cues. However, the associations maintained in this system would rapidly decay over weeks because of important neuronal recycling. In parallel, learning mechanisms relying on the mere extraction of visual information would develop at a lower level of visual processing. Information coded by this system is visual by nature. This memory requires both longer and multiple exposures to a specific stimulus to support familiarity, but it would be more robust over time and less subject to interference effects. Pervasive cortical plasticity phenomena (e.g. Spike Timing Dependent Plasticity) are good candidates to account for the formation of such sensory memories[50,51].

However, research conducted in the fields of implicit learning and statistical learning reveal the limits of such a clear functional dichotomy between explicit and implicit memory systems that would be governed by different learning principles and that would operate in isolation from each other. Memory phenomena result in large part from both external, slow, pervasive, and cortex-based mechanisms of learning, and on transitory associative representations formed and maintained within the medial temporal lobe memory system[45,52–54]. In addition, important changes in the functional connectivity between the hippocampus and cortical areas operate during memory consolidation, especially with a progressive disengagement of the Medial Temporal Lobe and both synaptic and systemic consolidation in the neocortex[46,55,56]. How those different memory systems interact and how a redescription of knowledge operates over time and consolidation remain a challenge for further research. In this view, the present research highlights important changes in memory across weeks, which show its relevance for assessing memory after weeks and months. The weakness of most research in the field of visual LTM is that it examines memory immediately after learning only. Studying how memories evolved over time remains fundamental to understand the format and the content of memories in LTM.

To conclude, the present study shows that while semantic and conceptual information enhance learning of images in LTM systems for transitory periods, they might not be able to account for longer term memorization of images. In contrast, information stored at a lower level might be more robust over time and might be more resistant to interfering effects. This hypothesis could be examined by assessing memory over months or even years for images that are presented several times. The problem of how images are stored and manipulated within the human brain remains a fertile area for further research and to address the issue of the coding of information into memory.

Methods

Participants:

Thirty-six adult individuals (mean age = 26 years; SD=6 years, range = 17-42 years) participated in Experiment 1 and thirty-six 9 year-old children participated in Experiment 2. All were naïve to the purpose to the study and reported normal or corrected-to-normal acuity with no color vision deficiencies. The adult participants received course credits and gave written informed consent before starting the experiment. The parents of the children signed a similar informed consent form. The children were free to accept or to refuse participation in the experiment both for the learning phase and for the testing phase. The procedures were in accordance with the Declaration of Helsinki and approved by the local ethics committee “Comité d’Evaluation Ethique de l’Inserm”.

The materials and procedure were the same in Experiment 1 and Experiment 2.

Material:

The material included 360 different full-colored images, with 200 “meaningful” images and 160 “meaningless” images (for several examples, see Figure 3). An additional 8 images (4 meaningless and 4 meaningful) were used for a practice block of learning.

Procedure:

The experiment included two phases: a learning phase followed by a testing phase.

Learning phase:

In the learning phase, observers were presented with 200 different full-colored images, of which 80 were meaningless and 120 were meaningful. The 120 meaningful images were photographs of either an animal, a vegetal, an object or a landscape. The signified/gist represented in each picture belonged to a unique basic-level category and was chosen because it could be quickly labeled using a simple name (e.g. a dog, a cherry, a beach).

Among the 80 meaningless images, 20 images were presented once during 120ms, 20 were presented twice during 120ms, 20 were presented once during 1920ms and 20 were presented twice during 1920ms. Among the 120 meaningful images, 30 were presented once during 120ms, 30 were presented twice during 120ms, 30 were presented once during 1920ms and 30 were presented twice during 1920ms. This gave a total of 300 trials. Note that the additional 40 meaningful images (10 in each of the four exposure conditions) were used to create the “New old-gist” condition in the recognition task (description in the

paragraph “testing phase”). Each trial started by a 500-ms cross fixation, followed by an image, then by a 1000-ms complex mask (for an example, see Figure 5).

The participants were instructed to remember each image as well as possible for a further memory task. They additionally performed a repetition detection task to maintain focus. They were told to press a button to indicate if the current item had been presented previously. The learning phase began after 12 familiarization trials that included four repeated images. After this familiarization, an instruction indicated the beginning of the experiment. The participants were exposed to the 300 trials (100 images presented once and 100 images presented twice). The order of presentation of the images, and consequently, the exposure duration and the number of repetitions of the images were all randomized across the experiment. Every 30 trials, the participants were shown a screen allowing them to take a break. They were free to continue the experiment when they were ready by pressing the space bar. The exposure duration, as well as the number of exposures for each image were counterbalanced between the participants.

Testing phase:

Participants were split into three different “delay groups” (12 per condition), in such a way that the testing phase was either administrated immediately after the learning phase, three weeks later or six weeks later. The memory of the participants for the images was assessed in a recognition task. Observers were presented with 360 images, that is, the 80 meaningless images that were presented in the learning phase (Meaningless-Old condition), 80 new meaningless images that were never seen before (Meaningless-New condition), 80 meaningful images from the 120 that were presented during the learning phase (Meaningful-Old condition), and 80 new meaningful images that were never seen before. Among the 80 new meaningful images, 40 belonged to 40 basic-level categories that were not used during the learning phase (Meaningful New condition), and 40 belonged to 40 basic-level categories that were already used during the learning phase (Meaningful New/Old-gist condition, for an example, *see Figure 2*). Each image was displayed for 3s. The participants were asked to decide whether or not they had seen the image in the study phase. Then, they rated the confidence in their response on a scale from 1 to 4. The scale was presented as follows: “Confidence? 1= just guessing, 2 = not sure, 3 = confident, 4 = very sure. The images that were used in the new conditions vs. the images that were used in the old conditions were counterbalanced between the participants.

The procedure of the experiment was programmed on Python and the stimuli were generated with Open Sesame.

Declarations

Acknowledgments:

We thank Sophie Muratot and Nicolas Tribout for collecting data, as well as Gabriel Besson, Pierre Yves Jonin, André Didierjean and Lionel Brunel for their comments, and suggestions on this work. This research was supported by grants from l'Agence Nationale de la Recherche (ANR) for the ELMA project and from the European Research Council (ERC) for the M4 project.

Availability of Data and Materials:

The datasets used and/or analysed during the current study are available from the corresponding author on request.

References

1. Pylyshyn, Z. W. Mental imagery: In search of a theory. *Behav. Brain Sci.* **25**, 157–182 (2002).
2. Pylyshyn, Z. Return of the mental image: Are there really pictures in the brain? *Trends Cogn. Sci.* **7**, 113–118 (2003).
3. Borst, G., Kosslyn, S. M. & Denis, M. Different cognitive processes in two image-scanning paradigms. *Mem. Cogn.* **34**, 475–490 (2006).
4. Kosslyn, S. M. Information representation in visual images. *Cogn. Psychol.* **7**, 341–370 (1975).
5. Shepard, R. N. & Metzler, J. Mental rotation of three-dimensional objects. *Science (80-.).* **171**, 701–703 (1971).
6. Clark, J. M. & Paivio, A. Dual coding theory and education. *Educ. Psychol. Rev.* **3**, 149–210 (1991).
7. Paivio, A. Imagery and verbal processes. New York, NY: Holt, Rinheart & Winston. Paivio, A. 1986. Mental representation: A dual-coding approach. (1971).
8. Shepard, R. N. Recognition memory for words, sentences, and pictures. *J. Verbal Learning Verbal Behav.* **6**, 156–163 (1967).
9. Standing, L. Learning 10000 pictures. *Q. J. Exp. Psychol.* **25**, 207–222 (1973).
10. Nelson, D. L. 'Remembering Pictures and Words: Appearance, Significance and Name. *Inf. Process. Res. Advert.* 45–76 (1979).
11. Simons, D. J. & Levin, D. T. Change blindness. *Trends Cogn. Sci.* **1**, 261–267 (1997).
12. Intraub, H. Rapid conceptual identification of sequentially presented pictures. *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 604 (1981).
13. Irwin, D. E. Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **18**, 307–317 (1992).
14. O'Regan, J. K. & Noë, A. A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* **24**, 939–973 (2001).
15. Rensink, R. A. The dynamic representation of scenes. *Vis. cogn.* **7**, 17–42 (2000).
16. Simons, D. J. In sight, out of mind: When object representations fail. *Psychol. Sci.* **7**, 301–305 (1996).

17. Rensink, R. A., Kevin O'Regan, J. & Clark, J. J. On the Failure to Detect Changes in Scenes Across Brief Interruptions. *Vis. cogn.* **7**, 127–145 (2000).
18. Hollingworth, A. & Henderson, J. M. Accurate visual memory for previously attended objects in natural scenes. *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 113 (2002).
19. Brady, T. F., Konkle, T. & Alvarez, G. A. A review of visual memory capacity: Beyond individual items and toward structured representations. *J. Vis.* **11**, 4–4 (2011).
20. Schurgin, M. W. Visual memory, the long and the short of it: A review of visual working memory and long-term memory. *Attention, Perception, Psychophys.* **80**, 1035–1056 (2018).
21. Brady, T. F., Konkle, T., Alvarez, G. A. & Oliva, A. Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci.* **105**, 14325–14329 (2008).
22. Isola, P., Xiao, J., Parikh, D., Torralba, A. & Oliva, A. What makes a photograph memorable? *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 1469–1482 (2014).
23. Isola, P., Xiao, J., Torralba, A. & Oliva, A. What makes an image memorable? *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 145–152 (2011). doi:10.1109/CVPR.2011.5995721
24. Rust, N. C. & Mehrpour, V. Understanding Image Memorability. *Trends Cogn. Sci.* (2020).
25. Brady, T. F., Alvarez, G. A. & Störmer, V. S. The role of meaning in visual memory: Face-selective brain activity predicts memory for ambiguous face stimuli. *J. Neurosci.* **39**, 1100–1108 (2019).
26. Cunningham, C. A., Yassa, M. A. & Egeth, H. E. Massive memory revisited: Limitations on storage capacity for object details in visual long-term memory. *Learn. Mem.* **22**, 563–566 (2015).
27. Konkle, T., Brady, T. F., Alvarez, G. A. & Oliva, A. Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychol. Sci.* **21**, 1551–1556 (2010).
28. Potter, M. C. Conceptual Short Term Memory in Perception and Thought. *Front. Psychol.* **3**, 1–11 (2012).
29. Potter, M. C. Recognition and memory for briefly presented scenes. *Front. Psychol.* **3**, 1–9 (2012).
30. Gathercole, S. E. Cognitive approaches to the development of short-term memory. *Trends Cogn. Sci.* **3**, 410–419 (1999).
31. Brainerd, C. J., Holliday, R. E. & Reyna, V. F. Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Dev.* **75**, 505–522 (2004).
32. Richmond, J. & Nelson, C. A. Accounting for change in declarative memory: A cognitive neuroscience perspective. *Dev. Rev.* **27**, 349–373 (2007).
33. Rose, S. A., Feldman, J. F. & Jankowski, J. J. Infant visual recognition memory. *Dev. Rev.* **24**, 74–100 (2004).
34. Ferrara, K., Furlong, S., Park, S. & Landau, B. Detailed Visual Memory Capacity Is Present Early in Childhood. *Open Mind* **1**, 136–147 (2017).
35. Thunell, E. & Thorpe, S. J. Memory for repeated images in rapid-serial-visual-presentation streams of thousands of images. *Psychol. Sci.* **30**, 989–1000 (2019).

36. Thunell, E. & Thorpe, S. J. Regularity is not a key factor for encoding repetition in rapid image streams. *Sci. Rep.* **9**, 1–10 (2019).
37. Brady, T. F., Konkle, T., Oliva, A. & Alvarez, G. A. Detecting changes in real-world objects: The relationship between visual long-term memory and change blindness. *Commun. Integr. Biol.* **2**, 1–3 (2009).
38. Rotello, C. M., Masson, M. E. J. & Verde, M. F. Type I error rates and power analyses for single-point sensitivity measures. *Percept. Psychophys.* **70**, 389–401 (2008).
39. Yonelinas, A. P. Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *J. Exp. Psychol. Learn. Mem. Cogn.* **20**, 1341 (1994).
40. Kouststaal, W. *et al.* False recognition of abstract versus common objects in older and younger adults: Testing the semantic categorization account. *J. Exp. Psychol. Learn. Mem. Cogn.* **29**, 499–510 (2003).
41. Konkle, T., Brady, T. F., Alvarez, G. a & Oliva, A. Categories in Visual Long-Term Memory. **21**, 1551–1556 (2010).
42. Schacter, D. L., Guerin, S. A. & Jacques, P. L. S. Memory distortion: An adaptive perspective. *Trends Cogn. Sci.* **15**, 467–474 (2011).
43. Khosla, A., Raju, A. S., Torralba, A. & Oliva, A. Understanding and predicting image memorability at a large scale. *Proc. IEEE Int. Conf. Comput. Vis.* **2015 Inter**, 2390–2398 (2015).
44. Brown, G. D. A., Neath, I. & Chater, N. A temporal ratio model of memory. *Psychol. Rev.* **114**, 539 (2007).
45. Goujon, A., Didierjean, A. & Thorpe, S. Investigating implicit statistical learning mechanisms through contextual cueing. *Trends Cogn. Sci.* **19**, 524–533 (2015).
46. Larzabal, C., Bacon-Macé, N., Muratot, S. & Thorpe, S. J. Tracking Your Mind's Eye during Recollection: Decoding the Long-Term Recall of Short Audiovisual Clips. *J. Cogn. Neurosci.* **32**, 50–64 (2020).
47. Reber, P. J. & Squire, L. R. Parallel brain systems for learning with and without awareness. *Learn. Mem.* **1**, 217–229 (1994).
48. Squire, L. R. Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *J. Cogn. Neurosci.* **4**, 232–243 (1992).
49. Tulving, E. Episodic and semantic memory. *Organization of memory* **1**, 381–403 (1972).
50. Markram, H., Gerstner, W. & Sjöström, P. J. Spike-timing-dependent plasticity: a comprehensive overview. *Front. Synaptic Neurosci.* **4**, (2012).
51. Masquelier, T., Guyonneau, R. & Thorpe, S. J. Competitive STDP-based spike pattern learning. *Neural Comput.* **21**, 1259–1276 (2009).
52. Goujon, A., Didierjean, A. & Poulet, S. The emergence of explicit knowledge from implicit learning. *Mem Cogn.* **42**, 225–236 (2014).

53. Henke, K. A model for memory systems based on processing modes rather than consciousness. *Nat Rev Neurosci* **11**, 523–532 (2010).
54. Henke, K., Reber, T. P. & Duss, S. B. Integrating events across levels of consciousness. *Front. Behav. Neurosci.* **7**, 68 (2013).
55. Frankland, P. W. & Bontempi, B. The organization of recent and remote memories. *Nat. Rev. Neurosci.* **6**, 119–130 (2005).
56. Takeda, M. Brain mechanisms of visual long-term memory retrieval in primates. *Neurosci. Res.* **142**, 7–15 (2019).

Tables

Table 1

False alarms (FA) rates in the Meaningless Lures and Novel-gist Meaningful Lures conditions (depending of the delay) and in the Old-gist meaningful Lures conditions (depending of the delay and of the exposure duration and the number of exposures of the gist-related images presented during the learning phase). SE indicated the standard errors. *t* Tests and BF_{10} were determined by a comparison between the old-gist lures and novel gist lures for each respective delay.

	Condition									
	Meaningless		**Novel-gist meaningful**		**Old-gist meaningful**					
Delay	**FA**	SE	**FA**	SE		**FA**	SE	*t*(12)	*p*	BF_{10}
Immediate	**0,181**	.034	**0,135**	.032	120ms - 1 expo.	**0,267**	.036	2.65	.023	3.09
					120ms - 2 expo.	**0,225**	.048	2.47	.031	2.38
					1920ms - 1 expo.	**0,142**	.029	0.14	.888	0.29
					1920ms - 2 expo.	**0,175**	.039	1.22	.249	0.53
3-weeks	**0,234**	.026	**0,306**	.035	120ms - 1 expo.	**0,308**	.063	0.04	.970	0.29
					120ms - 2 expo.	**0,333**	.035	0.93	.372	0.41
					1920ms - 1 expo.	**0,308**	.043	0.05	.962	0.29
					1920ms - 2 expo.	**0,342**	.061	0.63	.541	0.34
6-weeks	**0.232**	.045	**0,358**	.039	120ms - 1 expo.	**0,292**	.060	1.50	.162	0.70
					120ms - 2 expo.	**0,317**	.037	2.08	.062	1.42
					1920ms - 1 expo.	**0,417**	.055	1.41	.183	0.64
					1920ms - 2 expo.	**0,258**	.061	2.20	.050	1.64

Table 2

False alarms (FA) rates in the Meaningless Lures and Novel-gist Meaningful Lures conditions (depending of the delay) and in the Old-gist meaningful Lures conditions (depending of the delay and of the exposure duration and the number of exposures of the gist-related images presented during the learning phase). SE

indicated the standard errors. t Tests and BF_{10} were determined by a comparison between the old-gist lures and novel gist lures for each delay.

	Condition									
	Meaningless		Novel-gist meaningful		Old-Gist meaningful					
Delay	FA	SE	FA	SE		FA	SE	$t(12)$	p	BF_{10}
Immediate	**0,187**	.041	**0,154**	.024	120ms - 1 expo.	**0,192**	.034	1.14	.279	2.04
					120ms - 2 expo.	**0,267**	.062	2.35	.039	2.02
					1920ms - 1 expo.	**0,183**	.037	0.88	.397	0.40
					1920ms - 2 expo.	**0,233**	.058	1.69	.224	0.87
3-weeks	**0,157**	.055	**0,225**	.057	120ms - 1 expo.	**0,272**	.077	1.25	.239	0.54
					120ms - 2 expo.	**0,279**	.066	1.14	.277	0.49
					1920ms - 1 expo.	**0,272**	.075	1.88	.087	1.09
					1920ms - 2 expo.	**0,289**	.085	1.27	.231	0.55
6-weeks	**0.214**	.042	**0,306**	.046	120ms - 1 expo.	**0,300**	.049	0.14	.895	0.29
					120ms - 2 expo.	**0,354**	.037	1.18	.262	0.51
					1920ms - 1 expo.	**0,292**	.060	0.37	.716	0.30
					1920ms - 2 expo.	**0,337**	.052	0.84	.505	0.39

Figures

Figure 1

Examples of images used in the experiment

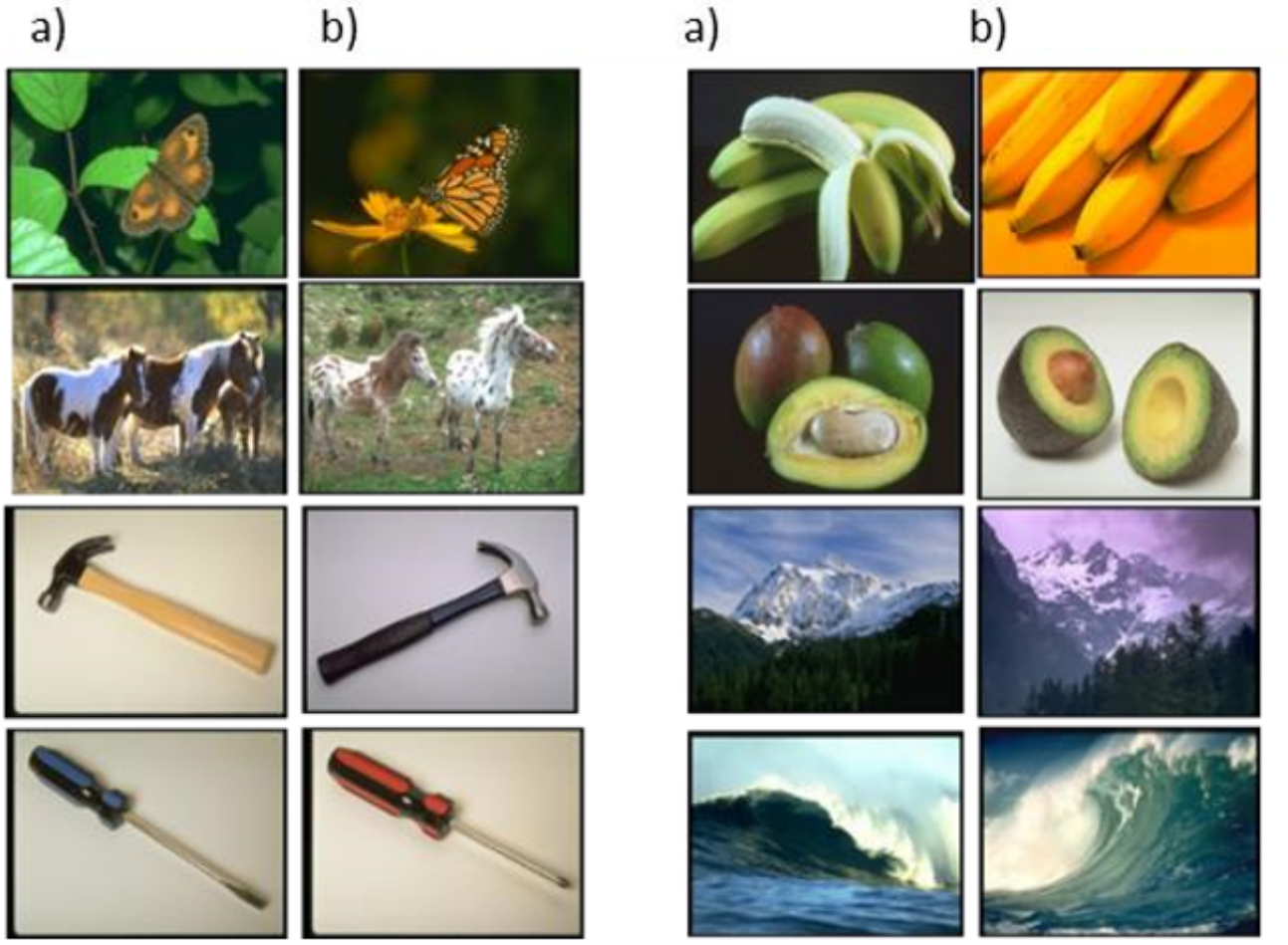

Figure 2

Examples of images used for the old-gist lures. For example, the images a) were presented during the learning phase and the images b) during the testing phase.

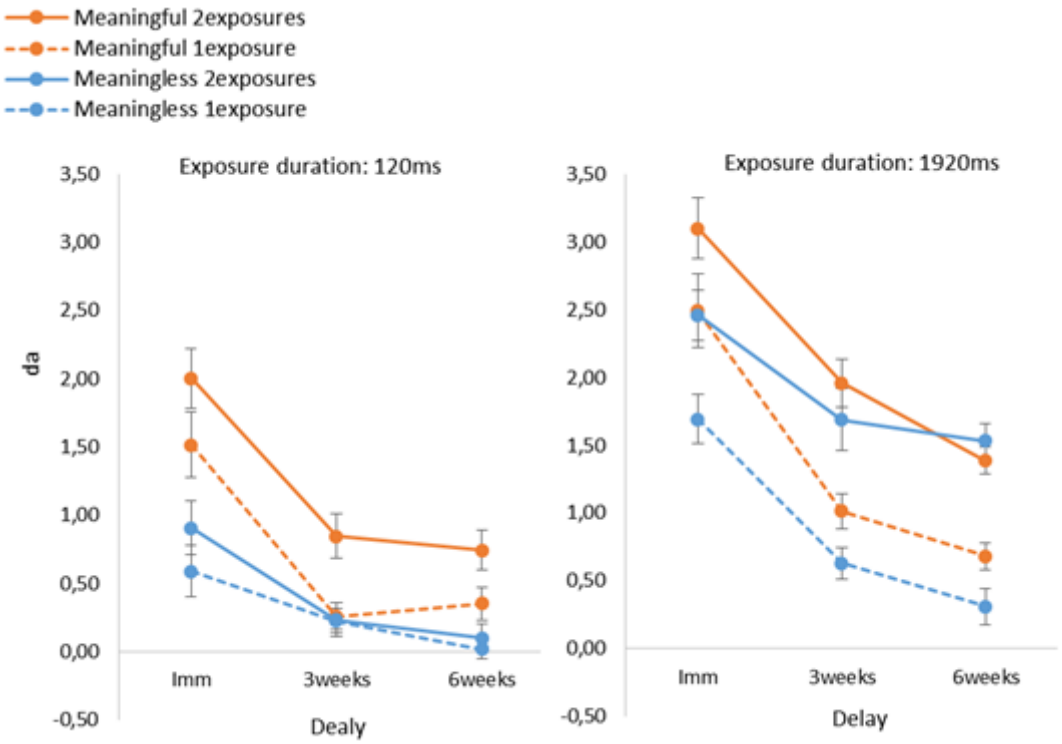

Figure 3

Mean discrimination indexes (d_a) depending of the delay (immediate, 3-weeks, and 6-weeks), the type of images (meaningful vs. meaningless), the exposure duration (120ms vs. 1920ms) and the number of exposures (1 vs. 2). The error bars show the standard error of the mean (n = 12).

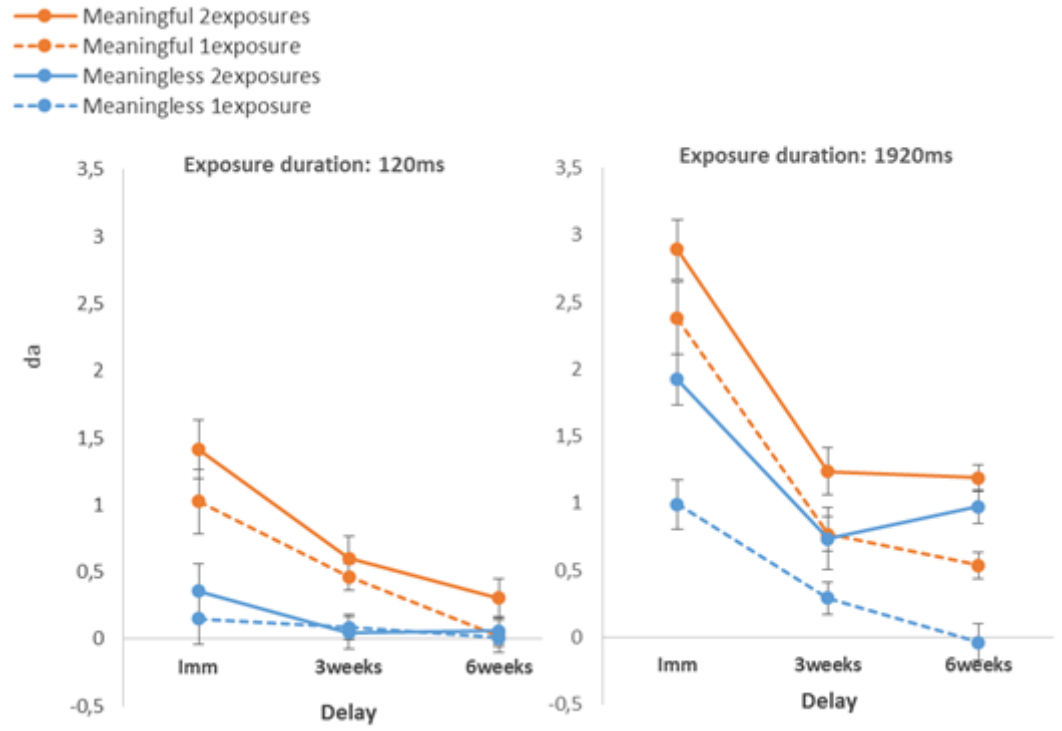

Figure 4

Mean discrimination indexes (d_a) depending of the delay (immediate, 3-weeks, and 6-weeks), the type of images (meaningful vs. meaningless), the exposure duration (120ms vs. 1920ms) and the number of exposures (1 vs. 2). The error bars show the standard error of the mean (n = 12).

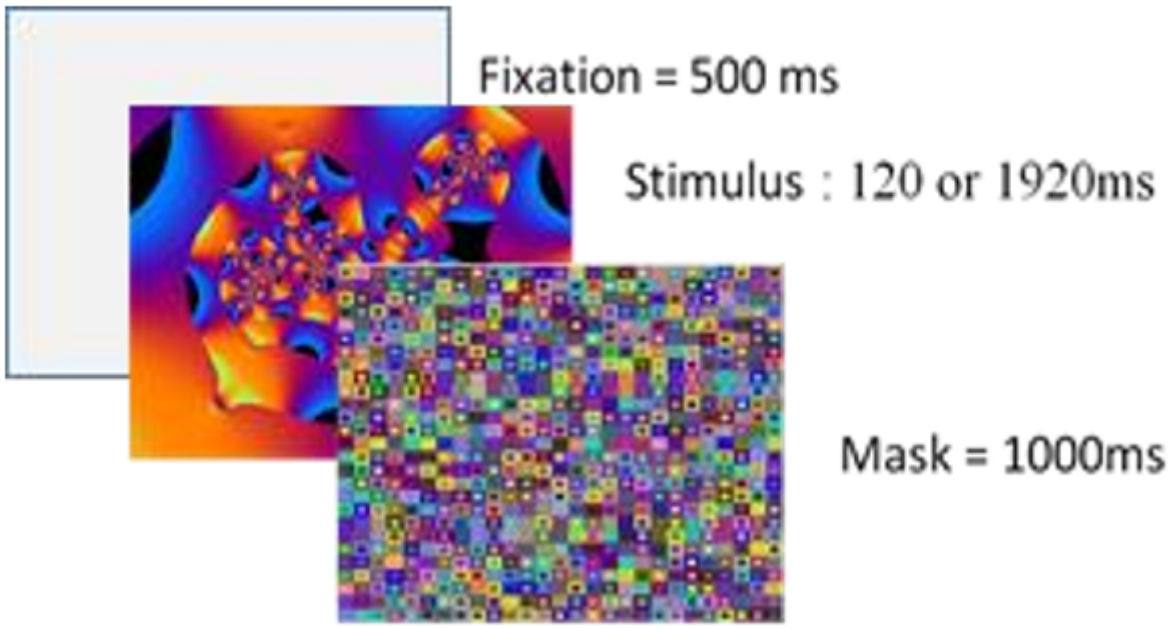

Figure 5

Sequence of a trial during the learning phase. Each trial started by a 500-ms cross fixation, followed by an image, then by a 1000-ms complex mask.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Rowdata.xlsx
- Supplementarymaterials.docx