Journal of Experimental Child Psychology 215 (2022) 105316

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Predicting explicit memory for meaningful cartoons from visual paired comparison in infants and toddlers

Osman Skjold Kingo *, Trine Sonne, Peter Krøjgaard

Center on Autobiographical Memory Research, Department of Psychology and Behavioral Sciences, Aarhus University, DK-8000 Aarhus C, Denmark

A R T I C L E I N F O

Article history:
Received 13 January 2021
Revised 12 October 2021

Keywords:
VPC
ΔVPC
Memory development
Dynamic stimulus material
Implicit memory
Explicit memory

A B S T R A C T

We tested the memory of 18-, 33-, and 39-month-olds ($N = 120$) for dynamic stimulus material (simple cartoons) after 6 months in a visual paired comparison (VPC) task. We also tested the explicit recognition memory (ERM) for the same material. Only the oldest age group (39-month-olds) showed a significant visual (familiarity) preference at the test. Similarly, only the oldest group reliably chose the correct cartoon in the ERM test. Data from the VPC and ERM tasks did not correlate in any age group. However, we suggested a novel score (coined ΔVPC) measuring how much visual preference changes during the test phase in the VPC task. We found that this ΔVPC score (and vocabulary) predicted children's performance in the ERM task, whereas other potential predictors such as age and conventional novelty preference did not. We discuss the impact of these findings in relation to the development of implicit and explicit memory. Furthermore, we propose that VPC measures are associated with explicit memory only when the participants processed the stimuli conceptually. In such cases, we suggest that the ΔVPC score is an approximation of how demanding it is to construct the mental representation of the familiar stimulus during the test phase.

* Corresponding author.
E-mail address: osman@psy.au.dk (O.S. Kingo).

https://doi.org/10.1016/j.jecp.2021.105316
0022-0965/© 2021 The Author(s). Published by Elsevier Inc.

Introduction

Developmental scientists have long been interested in infant looking behavior because it provides a window into the perceptual and cognitive world of the preverbal human infant (Reynolds, 2015). In the study of memory during infancy, the visual paired comparison (VPC) procedure in particular has been a popular method of choice for researchers because it does not rely on verbal instructions but also because it has proven to be viable for a very broad age range of participants from infancy to adulthood (Krøjgaard, Sonne, & Kingo, 2020; Morgan & Hayne, 2011; Richmond, Sowerby, Colombo, & Hayne, 2004).

In the VPC procedure, two visual stimuli are presented simultaneously: one stimulus for which the participant has had the opportunity to gain some kind of familiarity (the familiar stimulus) and one stimulus that is novel to the participant (the novel stimulus). Any looking time above chance level at either the novel stimulus (a novelty preference) or the familiar stimulus (a familiarity preference) at test has been interpreted as memory for the familiar stimulus (e.g., Krøjgaard et al., 2020; Richmond, Colombo, & Hayne, 2007). The interpretation of familiarity and novelty preferences has typically been based on Sokolov's (1963) comparator model, which proposes that the level of processing of the familiar stimulus prior to the test phase determines the visual preference for the familiar stimulus versus the novel stimulus. Thus, a novelty preference is assumed to reflect recognition of a fully processed familiar stimulus, whereas longer looking to the familiar stimulus is assumed to reflect further encoding of a stimulus that has not yet been fully processed (Reynolds, 2015).

Other factors also play a role, however. Hunter and Ames (1988) proposed that multiple factors such as the age of the infant, the amount of previous exposure to the stimuli, and task difficulty (stimulus complexity) would also influence visual preference. For instance, the prediction would be that more complex stimuli such as very detailed or dynamic material would be more difficult to encode and hence remember than simplified and static stimuli (e.g., Kingo & Krøjgaard, 2015; Sonne, Kingo, & Krøjgaard, 2018) but also that older infants/children would need less exposure time to a given stimulus during encoding than younger individuals for a comparable memory trace (e.g., a novelty preference) to form (e.g., Richards, 1997; Rose, 1983). Similarly, Bahrick and Pickens (1995) suggested that the retention interval (between encoding and test) for the visual recognition memory would be important. They proposed a model in which visual preference in the VPC procedure would change as a function of retention described in four phases: A short retention interval (associated with a strong memory trace) would result in a novelty preference, intermediate retention intervals would produce a null preference, long retention intervals (associated with weak memory traces) would produce a familiarity preference, and very long retention (associated with inaccessible or lost memory traces) would again result in a null preference. It is difficult to settle on a convention as to what qualifies as ‘‘short,” ‘‘intermediate,” ‘‘long,” and ‘‘very long” retention intervals because many factors affect visual preference. However, in Bahrick and Pickens' (1995) case, this four-phase pattern was confirmed with 3-month-olds and with the following intervals: short = 1 min, intermediate = 1 day, and long = 1–3 months.

One aspect is the factors affecting the visual behavior in the VPC procedure, and another is the question of which kind of memory we are tapping into with the procedure. Of particular interest is whether the kind of memory documented by the VPC task is explicit in the sense that we can become aware of it and communicate it to others (i.e., declarative memory) or whether it is implicit in nature—simply something that guides our visual behavior without any direct association with declarative memory (e.g., Hayne, 2004; Kingo, Staugaard, & Krøjgaard, 2014; Manns, Stark, & Squire, 2000; Snyder, Blank, & Marsolek, 2008; Wagner, Gabrieli, & Verfaellie, 1997). Manns et al. (2000), for instance, found that adults' performance in the VPC task was predictive of subsequent recognition memory performance, whereas perceptual priming was unrelated to subsequent recognition memory performance. They concluded that the VPC task is a measure of declarative (explicit) memory. In contrast, Snyder et al. (2008) used a classical conditioning paradigm to test whether novelty preferences reflected a stimulus-driven bias toward novelty in the VPC task or explicit memory for old material. Conditioning was found to affect adults' looking behavior in the VPC task but did not affect explicit recognition memory (ERM) judgments. Snyder and colleagues concluded, that novelty preference might reflect attentional processes and implicit memory to a greater degree than explicit memory. As is often the case, the studies from Manns et al. (2000) and Snyder et al. (2008) vary on several methodological parameters—namely, age of the adult participants, measure of implicit memory (priming vs. conditioning), and memory test (forced choice vs. yes/no recognition). This makes it difficult to conclude why their findings differ (but see Snyder et al., 2008, for a discussion of some of these differences). However, we should point to the fact, that Manns et al. (2000) used pre-experimentally familiar stimuli (i.e., photos of common objects that could readily be named by the participants), whereas Snyder et al. (2008) used novel and random polygons. As such, there was a substantial difference in the conceptual meaningfulness between the stimuli in the studies. This may very well have been a key factor for the seemingly conflicting results of these two studies. A third stance on these matters comes from Newcombe and colleagues, who argued that the VPC task may well measure explicit memory but that, especially early in development, it may index (explicit) familiarity to a higher degree than recollective (hippocampally dependent) processes (Olson & Newcombe, 2014; see also Rose, Feldman, Jankowski, & Van Rossem, 2011). We return to these issues later.

In summary, there is no simple answer to the question of whether the VPC task informs us on implicit or explicit memory or maybe on elements of both. The question is important, however. If measures of VPC are (also) associated with explicit memory, it would enable us to use the VPC task (with all its merits such as no verbal requirements and very broad age applicability) to gain valuable insights into the underlying processes and development of more complex declarative kinds of memory such as episodic and autobiographical memory (cf. Dahl, Sonne, Kingo, & Krøjgaard, 2013; Kingo et al., 2014; Krøjgaard et al., 2020). Consequently, one aim of the current study was to investigate associations between performance in a VPC task and performance in an ERM task.

The current study

We had four major aims with the current study: First, we wanted to test the strength of a memory trace for dynamic material in different age groups after, by most accounts, a long or very long retention interval of 6 months. Second, we wished to compare the VPC measure with performance in an ERM task. Third, we wished to investigate the visual preference as it unfolds over time during the test phase. Finally, we wanted to investigate potential predictors of ERM performance. Each of these aims is qualified in detail below.

The first aim was to test whether we could show memory for dynamic stimulus material with the VPC task after a very long retention interval (6 months). In accordance with Bahrick and Pickens' (1995) four-phase model, we predicted that a memory trace after 6 months either would be inaccessible (thereby resulting in a null preference) or would present itself as a familiarity preference in the VPC task due to the weakness of the memory trace. We and others have previously argued that dynamic visual stimuli (involving movement, agents, and actions) allows for better generalization of data to memory for real-world events because the real world is dynamic in nature (Bahrick, Gogate, & Ruiz, 2002; Kingo & Krøjgaard, 2015). For such dynamic stimuli (i.e., short video cartoons with a simple "storyline"), we have found evidence that information from the meaningful story per se seems to facilitate memory in the VPC task to a higher degree than static information or lower-level perceptual features of the same material in 18-month-olds (Sonne et al., 2018). In addition, age is an important factor for memory in the VPC task. The expectancy is that older children will have stronger memory traces and over longer retention intervals than younger children (Bahrick & Pickens, 1995; Hunter & Ames, 1988). Performance also increases with age for other kinds of memory such as conditioning and declarative memory (Bauer, 2007; Hayne, 2004). Consequently, we tested different age groups. We tested a group of 18-month-olds, chosen because this is an age of both rapid vocabulary growth (Oates & Grayson, 2004) and important changes in self-awareness (Bard, Todd, Bernier, Love, & Leavens, 2006; Moore, Mealiea, Garon, & Povinelli, 2007). We also tested groups of older children, 33- and 39-month-olds (i.e., just below and just above 3 years), because this is the typical age for the offset of childhood amnesia (i.e., memories have a chance of being retained into adulthood; Bauer, 2007; Hayne, 2004; Kingo, Berntsen, & Krøjgaard, 2013). Therefore, changes in the explicit memory performance close to 3 years of age are of particular theoretical interest.

The second aim was to compare VPC performance with performance on an ERM task. This was to address the important question of the nature (implicit/explicit) of the memory measured by the VPC task (Hayne, 2004; Kingo et al., 2014; Manns et al., 2000; McKee & Squire, 1993; Snyder et al., 2008; Wagner et al., 1997). This need not be an either/or situation; it is possible that the VPC task holds the potential for measuring elements of both implicit and explicit memory processes. Whereas numerous studies have aimed to examine implicit memory and explicit memory separately (Reynolds, 2015; Rovee-Collier & Cuevas, 2009), research investigating the relation between different types of memory (i.e., implicit and explicit) is sparse, although arguably crucial for our understanding of memory development. Hence, we aimed to compare different measures of memory on a within-participants basis and with the same stimulus material. This, combined with the different age groups and the long retention interval, would provide us with novel and important data on the relative development of what would potentially be different kinds of memory.

The third aim relates to changes in visual preference *during* the test phase of the VPC task. Previous studies employing the VPC procedure have found that visual attention during the test phase tends to shift from familiarity to increased novelty preference (e.g., Houston-Price & Nakai, 2004; Kingo & Krøjgaard, 2015; Kingo et al., 2014; Richmond & Nelson, 2009; Roder, Bushnell, & Sasseville, 2000; Snyder et al., 2008). Although looking at in-session changes in visual preference is not a novel idea (e.g., Fantz, 1964; Sokolov, 1963), such data have typically not been analyzed as a distinct factor in VPC studies (Houston-Price & Nakai, 2004). In a previous study (Kingo & Krøjgaard, 2015), we extended the test phase to 60 s, which is beyond the typical length in comparable VPC studies, ranging from 5 to 10 s (e.g., Fagan, 1974; Roder et al., 2000; Snyder et al., 2008) to 20 s (e.g., Hayne, Jaeger, Sonne, & Gross, 2016; Richards, 1997). With this extended test phase, we found that attentional shifts from the first half to the second half of the test phase were associated with vocabulary. Low-vocabulary children would typically show a familiarity preference in the first half and a null or novelty preference in the second half, whereas high-vocabulary children would maintain a familiarity preference throughout the entire test phase. Our tentative interpretation of this was that (a) higher vocabulary was associated with better conceptual understanding of the cartoons, (b) a better conceptual understanding meant that there was more conceptual information to process (in addition to the perceptual information), and (c) this led to longer fixation at the familiar cartoon before this was "fully processed" in Sokolov's (1963) sense (Kingo & Krøjgaard, 2015). For the current study, we wished to be able to analyze these attentional shifts during an extended test phase. Based on our previous findings, we predicted that sustained attention to the familiar stimulus during an extended test phase would be associated with a better understanding of the conceptual content of the stimuli and that such an increased semantic processing in turn would be associated with an increased chance of correct explicit memory. This prediction is also in accordance with the notion that semantic memory develops early in life and may form the basis of other kinds of explicit memory such as episodic memory (e.g., Newcombe, Lloyd, & Ratliff, 2007).

Finally, we wanted to investigate whether we could find potential predictors of performance in ERM. Finding such predictors would provide us with a deeper understanding of elements of importance for explicit memory. We were especially interested in investigating whether measures from the VPC task would have such predictive value—both the conventional VPC measure and a measure of the attentional shifts during an extended test phase as mentioned in relation to the third aim above. However, to better establish the predictive power of these VPC measures, we also controlled for other potentially predictive factors such as age (of the participant), vocabulary, and vocabulary growth.

Method

Participants

A total of 120 infants and toddlers from three age groups participated in the study: 40 18-month-olds (M_{age} = 17.92 months, *SD* = 0.21, range = 17.6–18.4; 20 girls), 40 33-month-olds (M_{age} = 33.01 months, *SD* = 0.21, range = 32.6–33.5; 18 girls), and 40 39-month-olds (M_{age} = 39.06 months, *SD* = 0.29, range = 38.6–40.1; 20 girls). All participants were recruited via registries from the National

Board of Health and came from East Jutland in Denmark; they were predominantly Scandinavian Caucasian and lived in families with middle to higher socioeconomic status. Informed and written parental consent was obtained at the first visit. An additional 25 children were tested but excluded from the experiment: 7 due to fussiness, 4 due to technical or experimenter error, and 14 due to limited looking at the cartoons either at the encoding session (assessed by the experimenter) or during the test session (looking time > 2 standard deviations below the age group mean). The committee of science ethics in the county of Aarhus approved the study.

Equipment

A Tobii X120 eye-tracker (Tobii AB, Sweden) was used to record participants' fixations at 120 Hz (with 0.5° accuracy) on a 30-inch LCD widescreen. The total visual angle of the screen was 40° (width) × 25° (height), and the visual angle of the stimuli area was 33° (width) × 16.5° (height). The Tobii Fixation Filter (default) was used. This filter detects quick changes in the gaze point signal using sliding averaging and thus distinguishes between fixations and saccades. Interpolation of samples was used when data samples were missing (e.g., during a blink). A 5-point calibration was performed. Stimulus presentation was performed with E-Prime software (Psychology Software Tools, Pittsburgh, PA, USA). Participants were seated in a car seat mounted on top of an adjustable chair. The eyes of the participants were approximately 70 cm from the eye-tracker and level with the center of the screen.

Materials

The stimuli movies were two custom-made short cartoons (see Fig. 1), each with a duration of 30 s and of the same size. One cartoon, the S*nowman,* is set in a winter landscape. A snowman enters the scene and jumps toward a tall hat lying on the ground. The snowman looks down at the hat and starts jumping up and down. For each jump, the hat bumps higher and higher into the air until it finally lands on the head of the snowman. The snowman smiles and leaves the scene. The other cartoon, the *Crab,* is set on a beach. A crab enters the scene and moves toward a ball lying on the ground. The crab starts playing with the ball and juggling it until one of its claws punctures the ball, causing it to deflate and fall to the ground. The crab looks directly at the "camera" and then leaves the scene. There was no sound to the cartoons. In addition, all parents filled out a standardized Danish version of the MacArthur–Bates Communicative Development Inventory (CDI): Words and Gestures for the 18-month-olds and Words and Sentences for the two older age groups. The CDI was filled out at both the first and second visits.

Design and procedure

Participants visited the lab twice. At the first visit (T_1: encoding session), half of them watched the *Snowman* cartoon and half watched the *Crab* cartoon. The cartoons were presented four consecutive times (4 × 30 s) for a total of 120 s.[1] Six months later (*M* = 5.92 months, *SD* = 0.26), the participants returned to the lab (T_2: test session) and watched both cartoons simultaneously side by side (left/right positioning of the familiar and novel cartoons was balanced) two consecutive times (two iterations) for a total of 60 s while being eye-tracked. For the analysis of the eye-tracking data from T_2, the areas of interest where drawn so that they each covered one of the cartoons. The final phase of T_2 was for explicit questions (the ERM task). Immediately after the presentation of the cartoons at T_2, participants were presented with two still pictures from the *Crab* and *Snowman* cartoons respectively and side by side. The still pictures were placed in the same positions on the screen as their respective cartoons. The experimenter

[1] For this study, we did not record encoding data. As a rough encoding criterion, only participants who attended at least one of the four iterations of the cartoons during encoding were included in the remaining part of the study (assessed by the experimenter). However, in a similar study from our lab using the same stimuli and the exact same encoding procedure, we analyzed encoding data and found that 18-month-olds attended and habituated to the cartoons (Sonne et al., 2018). Therefore, we have good reason to believe that the same would be true for participants in the current study.

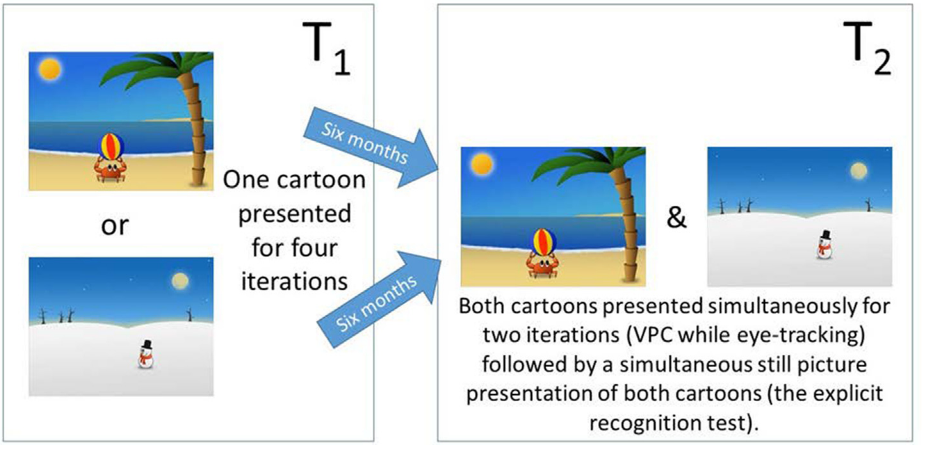

Fig. 1. Schematic depiction of the design and procedure, including static pictures from the two cartoons. VPC, visual paired comparison; T_1, encoding session at lab; T_2, test session at lab.

then said, "You were here a long time ago, where you saw one of these cartoons. Do you remember which one you have seen before? Was it this one [pointing to the left picture]? Or this one [pointing to the right picture]?" Participants could then either verbally or by pointing give an explicit reply. If no reply was given, the questions were repeated once. Finally, the experimenter scored the participants for giving either a correct reply, an incorrect reply, or no relevant reply (including "I don't know/remember").

Results

The presentation of results falls in three sections: eye-tracking data, ERM task, and combined data/ prediction models.

Eye-tracking data

First, we ran a preliminary analysis with total looking time at the cartoons during the 60-s T_2 presentation as the dependent measure and with age group (18, 33, or 39 months), sex (female or male), and specific cartoon (*Snowman* or *Crab*) as the independent measures. There was an effect of specific cartoon; participants looked longer at *Snowman* ($M = 30.45$) than at *Crab* ($M = 21.24$) ($p < .001$). This, however, did not interact with either age group or sex (all *ps* > .20), and as mentioned the presentation of the two cartoons was balanced during encoding at T_1; data from our own lab show that 18-month-olds remember this material after 2 weeks (familiarity preference) using the exact same procedure (Kingo & Krøjgaard, 2015), meaning that any difference in stimuli salience for this material can be overcome given the right retention interval. There was no main effect of or interactions with sex, but there was a main effect of age group ($p < .001$). A Bonferroni post hoc test revealed that the 18-month-olds looked less ($M = 48.31$) than the 33-month-olds ($M = 54.14$) and 39-month-olds ($M = 52.60$) (all *ps* < .02), whereas the two oldest age groups did not differ ($p = .87$). This also means that the three age groups fixated the movies from 81% to 90% of the time during the test phase. There was no difference in the total looking time between the first and second iterations of the movies in the two youngest age groups (18- and 33-month-olds, *ps* > .77), whereas the 39-month-olds looked longer in total at the first iteration ($M = 26.72$) than the second iteration ($M = 25.61$) ($p = .013$).

The primary dependent measure was the proportion of time spent looking at the novel stimulus during the test session compared with a .50 chance level (the novelty preference score; see, e.g., Kingo & Krøjgaard, 2015; Richmond et al., 2007; Roder et al., 2000). Only data from the first iteration (30 s) were considered for this analysis because we knew from previous studies that the preference could change from the first iteration to the second iteration (with individual variation; see also Roder et al., 2000). Furthermore, because the inclusion of the second iteration during the test was our own invention, we found data from the first iteration alone to be the “cleanest” conventional VPC measure. Data from the second iteration (the extended test phase) were included in subsequent specific analyses (see “Combined data/Prediction models” section below). Results of the main analyses are shown in Table 1. Only the 39-month-olds showed a preference (familiarity) different from chance and thereby were exhibiting memory for the familiar cartoon after 6 months (see Fig. 2 for a graphic display). In addition, we analyzed the first 5 s of the test, looking for early or transient familiarity or novelty preferences. None of the age groups showed any such early preference (all $ps > .68$).

ERM task

The results for the explicit questions are displayed in Table 2. Only the 39-month-olds were able to pick the correct cartoon above chance level. It is also notable that very few of the 18-month-olds (28%) gave any relevant reply, whereas 93% and 100% of the two oldest age groups, respectively, gave a relevant reply. This indicates that replying to these verbal and explicit questions was difficult for the 18-month-olds (see Hayne et al., 2016, for similar numbers for 1-year-olds). There was no significant correlation between giving a correct reply and the novelty preference score either overall ($r = .057$, $p = .597$) or within each age group (all $rs < .15$, all $ps > .392$).

Combined data/Prediction models

We were interested in potential predictors for giving a correct reply on the explicit test after 6 months of retention. We had the following candidates for such predictors: age, vocabulary (CDI), vocabulary growth (from T_1 to T_2), novelty-preference, and what we have called VPC preference change or ΔVPC (explained below).

Age

Age was a natural predictor candidate, and we have already seen that the age groups differed in memory performance. To avoid ordinal data in the regression analysis, the exact age at T_1 was used as a potential predictor here.

Vocabulary and vocabulary growth

Vocabulary was another predictor candidate because the explicit question at T_2 was delivered verbally and because a verbal reply to this question was an option. In addition, productive vocabulary has been found to be related to the ability to process and remember event sequences (e.g., Bauer & Wewerka, 1995; McGuigan & Salmon, 2004). We also included a measure of vocabulary growth (i.e., CDI T_2 minus CDI T_1) as a predictor. The vocabulary of all age groups increased significantly over the 6 months (all $ps < .001$ in paired-samples t test) but with considerable individual variation (standard deviations ranging from 58 to 120 words), which might explain some of the variance in the explicit memory test.

Table 1
Descriptives and results from the one-sample t tests for the three different age groups (against null preference = .50).

	T_2 first iteration					
	Descriptives		*t* tests			
Age group	*M*	*SD*	*t*	*df*	*p*	*d*
18-month-olds	.47	.19	−1.11	39	.275	−.16
33-month-olds	.45	.22	−1.48	39	.147	−.23
39-month-olds	.44	.19	−2.19	39	.035	−.32

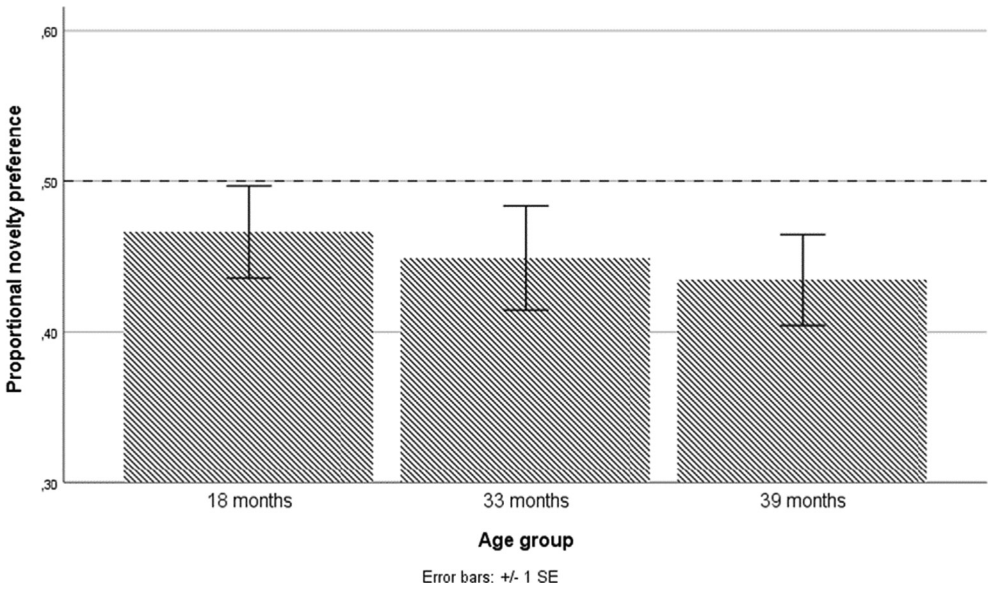

Fig. 2. Proportional looking times to the novel stimulus by age group. The dashed line is the .50 chance level.

Table 2
Descriptives and binomial test of the explicit memory recognition task

	Descriptives		Binomial tests		
Age group	Category	Observed *n*	Observed proportion	Test proportion	Exact significance
18-month-olds: Correct answer?	Yes	6	.55	.50	1.00
	No	5	.45		
	Total	11 (28%)	1		
33-month-olds: Correct answer?	Yes	24	.65	.50	.099
	No	13	.35		
	Total	37 (93%)	1		
39-month-olds: Correct answer?	Yes	30	.75	.50	.002
	No	10	.25		
	Total	40 (100%)	1		

Note. The observed *N* is always out of the 40 participants in each age group.

Novelty preference

The novelty preference score was an obvious predictor to test. This would test the possibility of a direct association between our implicit and explicit memory measures.

VPC preference change or ΔVPC

In a previous study, we showed that a change in visual preference from a first iteration to a second iteration of the cartoons during the presentation at T_2 was associated with vocabulary (Kingo & Krøjgaard, 2015). We suggested then that a sustained familiarity preference from the first iteration to the second iteration could be an expression of a more advanced understanding of the stimulus material for some of the infants (i.e., understanding the meaningful storyline of the cartoons (see also Sonne et al., 2018), meaning that more information was available for processing, resulting in increased processing time and visual fixation (Kingo & Krøjgaard, 2015). Here, we wanted to test the potential importance of the ΔVPC score by including it as a predictor of performance in the ERM test.

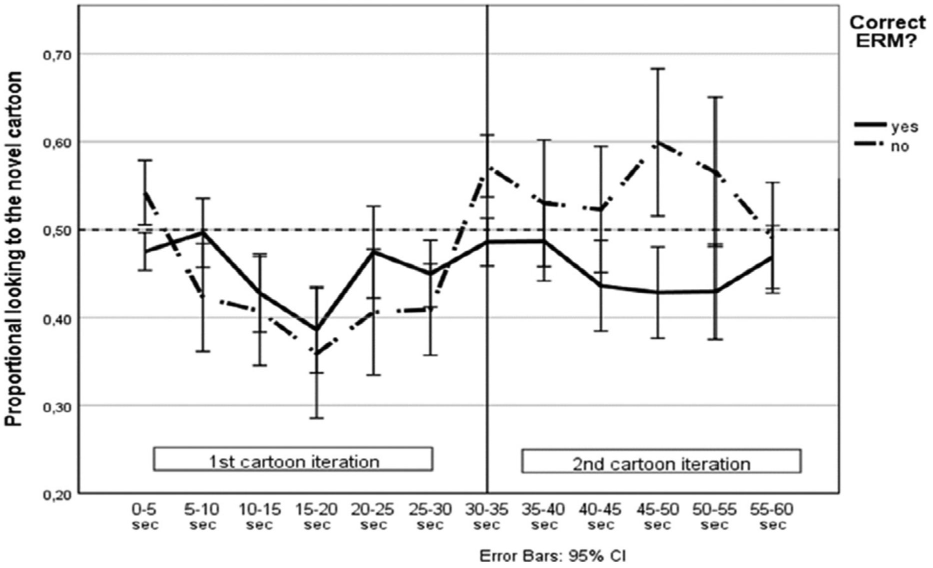

Fig. 3. Proportional looking to the novel cartoon over time in 5-s segments and split by individuals who did or did not give a correct reply in the explicit recognition memory (ERM) task. The dashed line is the .50 chance level. sec, seconds; 95% CI, 95% confidence interval.

ΔVPC was calculated by deducting the novelty preference score for the first iteration from that for the second iteration. Thereby, a positive score would mean a change toward novelty, a negative score would mean a change toward familiarity, and a score close to 0 would mean no change from the first iteration to the second iteration. A one-way analysis of variation (ANOVA) revealed no differences among the age groups in the ΔVPC score, $F(2, 117] = 2.215$, $p = .114$. When grouping all participants, we found that the ΔVPC score was negatively correlated with giving a correct reply in the explicit test ($r = -.231$, $p = .03$), indicating that an increased novelty preference from the first iteration to the second iteration was associated with giving a wrong reply in the ERM test (see also Fig. 3). For the novelty preference score for the second iteration alone, there was no correlation with ERM test performance ($r = .193$, $p = .072$), and as such it was meaningful to include the ΔVPC score as a predictor in the regression analysis.

Prediction model/Regression analysis

To find the best predictive model, we entered all potential predictors in a logistic regression analysis with backward stepwise (Wald) exclusion of variables (see Table 3). Only vocabulary and ΔVPC "survived" in the final model (Step 4).

The contribution of vocabulary may seem very modest (the odds ratio was .997), but it is important to remember that this means that for every unit increment in vocabulary (which is one word), the likelihood of giving an incorrect explicit answer decreases by 0.3%. This, in turn, means that for a vocabulary increase of 1 standard deviation (=239.47 words), the odds of giving a *correct* explicit answer increase by 71.8%. In addition, note that the estimated standardized $\beta = -.280$.

For ΔVPC, the odds ratio was 8.868. This means that for each unit increment in ΔVPC (i.e., attentional shift toward the novel stimulus), the odds for giving a wrong answer in the ERM test increase by 786.8%. So, for a ΔVPC increase of 1 standard deviation (=.269), the likelihood of giving an incorrect answer increases by 211.6%; in other words, the odds for giving a wrong explicit answer become about three times larger. The estimated standardized $\beta = .228$ and hence is in the same range as the β for vocabulary.

Table 3
Logistic regression with backward stepwise (Wald) exclusion of variables in four steps: Coefficients predicting the (incorrect) explicit answer in the explicit recognition memory task

	Model								
	Step 1		Step 2		Step 3		Step 4		
Variable	*b*	OR [CI]	*b*	OR [CI]	*b*	OR [CI]	*b*	OR [CI]	ϐ
Novelty preference	−.406	0.667 [0.044–10.011)	–	–	–	–	–	–	–
Vocabulary growth	.002	1.002 [0.995–1.010]	.002	1.002 [0.995–1.010]	–	–	–	–	–
Age	.044	1.045 [0.916–1.192]	.040	1.041 [0.916–1.184]	.032	1.032 [0.911–1.170]	–	–	–
Vocabulary (T_1)	−.003	0.997 [0.991–1.002]	−.003	0.997 [0.992–1.002]	−.004	0.996 [0.991–1.001]	−.003*	0.997 [0.995–1.000]	−.280*
ΔVPC	2.213*	9.146 [1.182–70.776]	2.304*	10.012 [1.430–70.119]	2.217*	9.180 [1.365–61.736]	2.182*	8.868 [1.332–59.063]	.228*
Constant	−.828	0.437	−.970	0.379	−.178	0.837	.396	1.487	
R^2 (Nagelkerke)	.146		.145		.140		.136		

Note. b, unstandardized beta; ϐ, estimated standardized beta (Menard, 2011); OR [CI], odds ratio [95% confidence interval]. All variance inflation factors (VIFs) are from 1.00 to 4.11. Final model (Step 4): $\chi^2(2) = 9.004$, $p = .011$. All assumptions were met.

* $p < .05$.

Discussion

In the current study, we sought to investigate different but related aspects of memory tested with the VPC procedure in addition to an ERM test.

The first aim of this study was to test the strength of the memory trace in the VPC paradigm after 6 months of retention. We tested this in three age groups. Only the oldest age group (39 months at T_1) convincingly showed memory, and as expected the memory manifested as a familiarity preference at T_2. Following Bahrick and Pickens' (1995) four-phase model, the 6 months seems to be a "long" retention interval for the 39-month-olds, resulting in a "weak" memory trace and hence familiarity preference, whereas the same interval was "very long" for the 18- and 33-month-olds, resulting in an "inaccessible or lost" memory trace and hence a null preference. Interestingly, Morgan and Hayne (2011) found a novelty preference with dynamic stimuli (faces with blinking eyes and moving mouths) in 4-year-olds after 6 months of retention in the VPC procedure. However, the children in that study were older than our oldest age group, and the stimuli were less complex. According to theory, both older age and reduced stimulus complexity should be more likely to produce a novelty preference while holding the retention interval constant (Bahrick & Pickens, 1995; Hunter & Ames, 1988; Richards, 1997). This does indeed seem to be the case when comparing the current study with Morgan and Hayne's (2011) study.

The second aim was to compare VPC performance with performance in an explicit memory recognition task. Even though only the 39-month-olds showed memory in both the VPC task and the explicit memory recognition task, we found no straightforward association between these two tasks for any of the age groups. Similarly, the novelty preference score in the VPC task did not predict explicit memory in the regression model. Our partial findings here are in line with the findings and conclusions from Snyder et al. (2008) that novelty preference may reflect attentional processes and implicit memory to a higher degree than explicit memory. However, we return to this matter after discussing the results from analyzing the ΔVPC score.

The third and fourth aims were interrelated. For the third aim, we wished to analyze attentional shifts during an extended test phase. Consequently, we coined and calculated the ΔVPC score and found it to be associated with performance in the ERM test. Fourth and finally, we attempted to find potential predictors for ERM performance. Both vocabulary and ΔVPC yielded such unique predictions. It might not be surprising that vocabulary predicts ERM performance given that the ERM task is verbally mediated. It may be more surprising that age was *not* a good predictor given that age ranged from 18 to 39 months. Apparently, among the many abilities that develop with age for children, vocabulary is of privileged importance for verbally mediated tasks such as the ERM task. The newly coined ΔVPC score also turned out to be a good predictor of ERM performance. In our view, this is an important finding, and therefore we discuss it in some detail.

The calculation of the ΔVPC score was facilitated by the decision to extend the test phase beyond the typical length in the VPC paradigm (our full/extended test phase was 60 s). The extension consisted of a full repetition of the 30-s cartoon (second iteration) with a simple storyline (see also Kingo & Krøjgaard, 2015, and Sonne, Kingo & Krøjgaard, 2016, 2017; Sonne et al., 2018). The repetitive nature of the extension may very well have been important for the pattern of results we obtained. For instance, when Fantz (1964) reported decreased attention to familiar complex patterns relative to novel patterns, repetition of the (familiar) stimulus was an integral part of the design, and the increasing novelty preference with successive exposure periods was explained by habituation to the familiar stimulus. Therefore, in the current study, when we at the test provided the participants with two full iterations of the cartoons, we gave them the opportunity to reprocess the meaningful storyline in addition to any lower-level perceptual information they may process. This was important because we know from a previous study with the same stimuli that the storyline is crucial for later memory by 18 months of age (Sonne et al., 2018). However, following the logic from Fantz (1964) as well as Hunter and Ames (1988), older (and hence more competent) children should direct their attention faster to novel stimuli than younger children. The explanation for this, again according to Sokolov's (1963) comparator model, would be that the older/more competent children process (and build a model for) the familiar stimulus faster than younger children. One would expect that the more competent children in the current study would also be the children more likely to answer correctly in the ERM task, and this was indeed what we found.

Following the existing literature cited above, one would also expect that the more competent children would direct their attention faster to the novel stimulus and thus have a higher ΔVPC score than those giving a wrong answer in the ERM task. In contrast, we found the opposite pattern, with a high ΔVPC score predicting giving a *wrong* explicit answer. This was what we had predicted from our previous studies using the same stimuli (Kingo & Krøjgaard, 2015; Sonne et al., 2018). Sustained attention to (i.e., continued processing of) the familiar stimulus predicted ERM. Our interpretation of this is that successful ERM is associated with good conceptual understanding of the material (i.e., understanding of event structure, agency, actions, and outcomes) because such conceptual understanding increases the likelihood of explicit access to the memory trace. Several previous studies have confirmed that these aspects are important for memory in infants and toddlers (Sonne et al., 2016, 2017, 2018), and studies with adults have found that familiarity-based explicit recognition may be more sensitive to conceptual processing than to perceptual processing (e.g., Wagner et al., 1997). In addition, an elaborated conceptual understanding of the stimuli is associated with sustained attention to the familiar stimulus in the extended test phase. This is because understanding the conceptual "layer" of the stimuli in addition to the lower-level perceptual features means that there is more information to process than if one is able to process only the perceptual features. Consequently (and following Sokolov, 1963), it takes longer to complete the model of the familiar stimulus before shifting attention to the novel stimulus. This interpretation is in accordance with Roder et al.'s (2000) finding that infants shifted from familiarity preference to novelty preference across trials for meaningful stimuli only, whereas nonmeaningful stimuli did not produce a familiarity preference prior to the novelty preference. In summary, in our interpretation conceptual understanding of the stimuli links the ΔVPC score and ERM performance. Furthermore, the ΔVPC score is an approximation of how demanding it is to construct the mental representation of the familiar stimulus during the test phase.

We do not think that our findings and interpretations contradict the studies from Fantz (1964) and Hunter and Ames (1988). However, we do hope to contribute to a better understanding of stimuli complexity. Fantz (1964) used relatively simple stimuli without any conceptual content, and older infants shifted their attention to the novel stimulus faster than younger infants. This is still in accordance with our interpretation given that there was no conceptual information to prolong the processing time of the familiar stimulus for the older infants. Hunter and Ames (1988) suggested that stimuli that are more complex would be more difficult to encode and hence remember. In the interpretation of the current study, this is still the case but with the addition that increased *conceptual* understanding can lead to increased perceived stimulus complexity for the individual just as an intermediate-level chess player may be presented with more difficult decisions than a beginner when looking at a complex middle game position on the chessboard.

So, is the memory measured by the VPC task more implicit or explicit in nature? Our conventional novelty preference score was unrelated to the ERM, but the fact that the ΔVPC score predicted ERM performance makes it difficult to maintain that there is no association between VPC performance and ERM as suggested by Snyder et al. (2008). Although it is clearly difficult to find an association between the conventional novelty preference score and ERM (also in the current study), the extended test phase and the ΔVPC score may present a novel opportunity to show exactly that. Given our interpretation of the current data, it also follows that an association between VPC measures and explicit memory may be depending on elements in the stimulus material that can be understood conceptually by the participants (i.e., semantic information; cf. Newcombe et al., 2007). This stance suggests that the VPC task may measure the result of *both* implicit and explicit memory processes and that the ΔVPC score may be better suited at tapping into the explicit memory contribution to VPC performance than the conventional novelty preference score. Accordingly, the ΔVPC score might not be useful *only* for conceptual information or in relation to explicit memory. In principle, ΔVPC may be seen as an application of Sokolov's comparator model to the test phase of VPC tasks. As such, it could provide useful information to any VPC study where the participants do not enter the test phase with a fully completed model of the familiar stimulus by informing us on the speed and difficulty of the current model's completion. This information could be a valuable addition to what we can learn from looking at encoding data and conventional VPC scores.

The finding of predictors of explicit memory from a VPC study may improve our ability to study and theorize about the memory system as it develops from infancy into more explicit forms of memory later in life such as episodic and autobiographical memory. While doing this, however, we must be mindful that the link between the VPC task and recollection per se may be weak initially and develop only gradually (e.g., Newcombe et al., 2007; Olson & Newcombe, 2014; Rose et al., 2011). With the current study, the suggestion of the applicability of the ΔVPC score, and the results specifically related to the predictive power of the ΔVPC score, we have attempted to suggest a fruitful path for further studies of the developing memory system.

Acknowledgments

We thank Jonna J. Dahl, Line R. Madsen, and Tirill F. Hjuler for their help as assistants. Furthermore, we thank all the participating infants as well as their parents. This study was supported by the Danish National Research Foundation (DNRF89) and the Velux Foundation (VELUX10386).

References

Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development, 73*(6), 1629–1643.
Bahrick, L. E., & Pickens, J. N. (1995). Infant memory for object motion across a period of three months: Implications for a four-phase attention function. *Journal of Experimental Child Psychology, 59*(3), 343–371.
Bard, K., Todd, B., Bernier, C., Love, J., & Leavens, D. (2006). Self-awareness in human and chimpanzee infants: What is measured and what is meant by the mark and mirror test? *Infancy, 9*, 191–219.
Bauer, P. J. (2007). *Remembering the times of our lives: Memory in infancy and beyond*. Mahwah, NJ: Lawrence Erlbaum.
Bauer, P. J., & Wewerka, S. S. (1995). One- to two-year-olds' recall of events: The more expressed, the more impressed. *Journal of Experimental Child Psychology, 59*(3), 475–496.
Dahl, J. J., Sonne, T., Kingo, O. S., & Krøjgaard, P. (2013). On the development of episodic memory: Two basic questions. *Nordic Psychology, 65*(2), 189–207.
Fagan, J. F. (1974). Infant recognition memory: The effects of length of familiarization and type of discrimination task. *Child Development, 45*(2), 351–356.
Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 146* (3644), 668–670.
Hayne, H. (2004). Infant memory development: Implications for childhood amnesia. *Developmental Review, 24*(1), 33–73.
Hayne, H., Jaeger, K., Sonne, T., & Gross, J. (2016). Visual attention to meaningful stimuli by 1- to 3-year-olds: Implications for the measurement of memory. *Developmental Psychobiology, 58*, 808–816.
Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development, 13*(4), 341–348.
Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In L. P. Lipsitt (Ed.), *Advances in child development and behavior* (pp. 69–95). San Diego: Academic Press.
Kingo, O. S., Berntsen, D., & Krøjgaard, P. (2013). Adults' earliest memories as a function of age, gender, and education in a large stratified sample. *Psychology and Aging, 28*(3), 646–653.
Kingo, O. S., & Krøjgaard, P. (2015). Eighteen-month-olds' memory for short movies of simple stories. *Scandinavian Journal of Psychology, 56*(2), 151–156.
Kingo, O. S., Staugaard, S. R., & Krøjgaard, P. (2014). Three-year-olds' memory for a person met only once at the age of 12 months: Very long-term memory revealed by a late-manifesting novelty preference. *Consciousness and Cognition, 24*, 49–56.
Krøjgaard, P., Sonne, T., & Kingo, O. S. (2020). Is the eye the mirror of the soul? Exploring autobiographical memory development by means of looking-time measures. In S. Gülgöz & B. Sahin-Acar (Eds.), *Autobiographical memory development: Theoretical and methodological approaches*. New York: Routledge.
Manns, J. R., Stark, C. E. L., & Squire, L. R. (2000). The visual paired-comparison task as a measure of declarative memory. *Proceedings of the National Academy of Sciences of the United States of America, 97*(22), 12375–12379.
McGuigan, F., & Salmon, K. (2004). Time to talk: The influence of the timing of adult–child talk on children's event memory. *Child Development, 75*, 669–686.
McKee, R., & Squire, L. (1993). On the development of declarative memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 397–404.
Menard, S. (2011). Standards for standardized logistic regression coefficients. *Social Forces, 89*(4), 1409–1428.
Moore, C., Mealiea, J., Garon, N., & Povinelli, D. (2007). The development of body self-awareness. *Infancy, 11*, 157–174.
Morgan, K., & Hayne, H. (2011). Age-related changes in visual recognition memory during infancy and early childhood. *Developmental Psychobiology, 53*(2), 157–165.
Newcombe, N. S., Lloyd, M. E., & Ratliff, K. R. (2007). Development of episodic and autobiographical memory: A cognitive neuroscience perspective. In R. V. Kail (Ed.). *Advances in child development and behavior* (Vol. 35, pp. 37–85). San Diego: Elsevier.
Oates, J., & Grayson, A. (2004). *Cognitive and language development in children*. Malden, MA: Blackwell.
Olson, I., & Newcombe, N. (2014). Binding together the elements of episodes: Relational memory and the developmental trajectory of the hippocampus. In P. J. Bauer & R. Fivush (Eds.), *The Wiley handbook on the development of children's memory* (pp. 285–308). Hoboken, NJ: John Wiley.

Reynolds, G. D. (2015). Infant visual attention and object recognition. *Behavioural Brain Research, 285*, 34–43.
Richards, J. E. (1997). Effects of attention on infants' preference for briefly exposed visual stimuli in the paired-comparison recognition-memory paradigm. *Developmental Psychology, 33*(1), 22–31.
Richmond, J., Colombo, M., & Hayne, H. (2007). Interpreting visual preferences in the visual paired-comparison task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 823–831.
Richmond, J., & Nelson, C. A. (2009). Relational memory during infancy: Evidence from eye tracking. *Developmental Science, 12*(4), 549–556.
Richmond, J., Sowerby, P., Colombo, M., & Hayne, H. (2004). The effect of familiarization time, retention interval, and context change on adults' performance in the visual paired comparison task. *Developmental Psychobiology, 44*(2), 146–155.
Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy, 1*, 491–507.
Rose, S. A. (1983). Differential rates of visual information processing in full-term and preterm infants. *Child Development, 54*(5), 1189. https://doi.org/10.2307/1129674.
Rose, S. A., Feldman, J. F., Jankowski, J. J., & Van Rossem, R. (2011). The structure of memory in infants and toddlers: An SEM study with full-terms and preterms. *Developmental Science, 14*, 83–91.
Rovee-Collier, C., & Cuevas, K. (2009). The development of infant memory. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (2nd ed., pp. 11–41). New York: Psychology Press.
Snyder, K. A., Blank, M. P., & Marsolek, C. J. (2008). What form of memory underlies novelty preferences? *Psychonomic Bulletin & Review, 15*(2), 315–321.
Sokolov, E. N. (1963). *Perception and the conditioned reflex*. New York: Macmillan.
Sonne, T., Kingo, O. S., & Krøjgaard, P. (2016). Occlusions at event boundaries during encoding have a negative effect on infant memory. *Consciousness and Cognition, 41*, 72–82.
Sonne, T., Kingo, O. S., & Krøjgaard, P. (2017). Bound to remember: Infants show superior memory for objects presented at event boundaries. *Scandinavian Journal of Psychology, 58*(2), 107–113.
Sonne, T., Kingo, O. S., & Krøjgaard, P. (2018). Meaningful memory? Eighteen-month-olds only remember cartoons with a meaningful storyline. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02388.
Wagner, A., Gabrieli, J., & Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 305–323.