



Mere exposure effect(s) in the context of explicit memory search

David Grybinas¹ · Justin Kantner² · Ian G. Dobbins¹

© The Psychonomic Society, Inc. 2019

Abstract

Prior stimulus exposure often increases later ratings of positive affect (e.g., pleasantness ratings). This phenomenon – the mere exposure effect (MEE) – appears robust following subliminal and incidental exposures. However, its expression in the context of explicit memory judgment remains unclear. In four studies, memory and pleasantness ratings were combined to investigate how memory conclusions (e.g., “studied” or “unstudied”) might moderate exposure effects. Experiment 1 examined basic recognition, Experiment 2 manipulated incentives for recognition decisions, and Experiments 3 and 4 examined source memory and paired-associate recall respectively. In general, items endorsed as recognized, attributed to the queried source, or accompanied by successful recall of a paired associate (i.e., confirmations) were rated as more pleasant than baseline norms. As important, items endorsed as unstudied, rejected as originating from a queried source, or failing to yield successful recall of a paired associated were rated as less pleasant than baseline norms. This suggests that it is the outcome of memory search that alters pleasantness ratings in the context of retrieval demands, and we discuss how this *confirmation of search* (COS) hypothesis accounts for current and prior findings.

Keywords Recognition · Memory · Decision making · Judgment

Introduction

Prior exposure to stimuli increases later judgments of positive affect (Zajonc, 1968). This mere exposure effect (MEE) is robust even if the prior exposure is incidental or subliminal (e.g., Monahan, Murphy, & Zajonc, 2000). Though traditionally measured through self-report (e.g., pleasantness rating), these affective responses have been assayed through low-level autonomic responses (Winkielman & Cacioppo, 2001) and neuroimaging (Elliot & Dolan, 1998). Numerous models have been proposed to explain MEE phenomena. Here, we focus on the perceptual-fluency misattribution model of Bornstein and D’Agostino (1992) and explicit-retrieval accounts of the MEE (e.g., Newell & Shanks, 2007) because they provide the most specific predictions concerning the relationship between explicit memory decision-making and the MEE.

Perceptual fluency, recognition, and the mere exposure effect (MEE)

In a meta-analysis, Bornstein (1989) examined 134 studies reporting 208 MEE contrasts with one key question being whether previously exposed stimuli should have been explicitly recognizable during affective ratings. As explained by Bornstein, this question had been examined through attempts to show that MEEs were present even when explicit recognition was statistically controlled, or by using stimulus encoding procedures that were independently verified to produce chance, or near chance levels of subsequent explicit recognition (although, see Berry, Shanks, & Henson, 2006; Newell & Shanks, 2007). In general, both approaches suggest that explicit recognition is not necessary for MEEs to occur (see de Zilva, Vu, Newell, & Pearson, 2013, for direct contradiction). In fact, Bornstein (1989) suggested that explicit recognition might impair or inhibit the expression of MEEs because the average effect size (across studies) for subliminally presented stimuli was larger than the average effect size in the meta-analysis as a whole.

As a direct test that recognition inhibits MEEs, Bornstein and D’Agostino (1992) experimentally contrasted subliminal (5 ms) and supraliminal (500 ms) exposures for two groups viewing stimuli between one and 20 times during encoding. Afterwards, subjects rated via booklet their liking of materials and their recognition of these same materials in a trial-wise fashion (order counterbalanced). Novel materials were also

✉ David Grybinas
davidgrybinas@wustl.edu

✉ Ian G. Dobbins
idobbins@wustl.edu

¹ Department of Psychological and Brain Sciences, Washington University in St. Louis, Saint Louis, MO 63130, USA

² Department of Psychology, California State University, Northridge, CA, USA

presented in the booklet. The key finding across two studies was that the 5-ms materials showed MEEs with minimum signs of explicit recognition, and that the MEE increased with prior exposure frequency. In contrast, MEEs in the 500-ms materials were absent (Experiment 1) or dampened (Experiment 2), although explicit recognition clearly increased with increased prior exposure frequency. This pattern led Bornstein and D'Agostino (1992) to advance a perceptual-fluency misattribution account (PFM) of the phenomenon based on the memory research of Jacoby and Kelley (1987) and Jacoby and Whitehouse (1989).

Under this account, MEEs reflect the increased fluency afforded by prior processing of studied materials, which participants misattribute as positive affect. However, if participants detect that materials are being drawn from a prior study phase, then they correctly attribute fluency to prior study, weakening or eliminating the MEE. We refer to this hypothetical shift in the attribution of fluency from liking to prior study as fluency “discounting” to emphasize that under this model, subjects discount the perceived fluency during affective ratings if they believe the materials are drawn from a prior experience (see also Bornstein & D'Agostino, 1994). Similar to the PFM model, Winkielman et al. (2003) also emphasize the role of fluency in producing MEEs. Under this hedonic fluency account, however, fluency is thought of as an indicator of cognitive progress, which in turn leads to greater positive affect. Though a viable model of the MEE, this conceptualization makes no theoretical commitment to affective ratings under memory decision-making and thus will not be included in subsequent discussions.

Recognition facilitates the MEE

Notwithstanding the aforementioned work indicating recognition dampens MEEs, there are studies in which explicit recognition of previously exposed stimuli appears to increase positive affect towards those stimuli (Anand, Holbrook, & Stephens, 1988; Berry, Shanks, & Henson, 2006; Brooks & Watkins, 1989; de Zilva et al., 2013; Fang, Singh, & Ahluwalia, 2007; Fox & Burns, 1993; Lee, 1994, 2001; Newell & Shanks, 2007; Stafford & Grimes, 2012; Sawyer, 1981; Wang & Chang, 2004). For example, Newell and Shanks (2007) only observed MEEs for stimulus-encoding conditions yielding above chance recognition performance when testing subjects on a two-alternative forced choice recognition and pleasantness rating procedure. These data would suggest that MEEs were strictly dependent on the recovery of explicit recognition evidence itself.

Aims of the current study

To explore the relationships between explicit memory judgments and putative MEEs, we made the following design

choices. First, we collected a normed baseline of the pleasantness ratings of a set of verbal items outside the context of recognition judgments. This allowed us to directly test whether the context of rendering memory decisions itself globally alters affective ratings. Second, we obtained explicit recognition, source memory, and paired associate recall judgments in concert with the pleasantness ratings. This allowed us to consider whether the subjective or objective memory status of the items influenced the affective ratings, and to determine if recognition effects extended to associative retrieval outcomes.

To preview, Experiment 1 demonstrated a pattern that was inconsistent with the PFM account and instead suggested that shifts in the affective ratings of recognition probes might reflect an increased motivation to respond “old” versus “new” during recognition. Moreover, these data, for the first time, demonstrated that items judged “new” were rated as significantly less pleasant than during pleasantness norming, indicating the “new” conclusion led to kind of devaluation. Experiment 2 tested the motivation hypothesis using incentives but failed to support it, while replicating the basic pattern of Experiment 1. In Experiments 3 (source memory) and 4 (paired associate cued-recall), we tested an alternate “confirmation of search” (COS) hypothesis that predicts that confirming an initiated memory search yields a positive affective response while failing to confirm it yields a negative affective response. The results supported this new hypothesis and for the first time demonstrated that source memory and cued-recall outcomes also affect the rated pleasantness of memoranda.

Experiment 1: Single-item recognition memory and the MEE

The purpose of Experiment 1 was to establish the basic relationship between recognition memory outcomes, MEEs, and the normative pleasantness of the materials. Since the materials were normed for pleasantness outside of recognition procedures, we were able to directly test whether memory decision outcomes yielded reliable increases or decreases with respect to baseline.

Method

Participants Twenty-six Washington University in St. Louis undergraduates participated in exchange for course credit. Two students were discarded from analysis due to software failure. Informed consent was obtained in compliance with the Institutional Review Board of Washington University in St. Louis. Testing occurred in groups of one to four people using computer carousels.

Materials Stimuli were randomly selected from a pool of 1,216 common nouns (e.g., fox) drawn from the MRC Psycholinguistic Database (Wilson, 1988) with an average of 7.09 letters, 2.34 syllables, and Kučera-Francis frequency of 8.85. All words were presented serially via Cambria 18-pt font on an all white background administered via computers running E-Prime Software. For each participant, a subset of 150 words was randomly selected for normative rating. This subset was held out during the experiment proper, leaving 1,066 words from which to sample for the subsequent study-test cycles.

Design and procedure Subjects began by rating 150 randomly selected words for pleasantness. These were different words than those used for the subsequent recognition/MEE experiment, allowing us to establish the baseline rated pleasantness for each word outside the context of a recognition demand. For all four subsequent study-test cycles, subjects studied 60 words and were tested on 120, yielding 480 tested stimuli. During each test, old and new items were randomly intermixed and subjects classified the recognition status of the probe and then rated its pleasantness, or visa versa. The order of these ratings was counterbalanced within subjects such that the first two test blocks used one order and the final two test blocks used the other order.

For the initial norming phase, participants rated serially presented words for pleasantness via a six-point scale (1 = very unpleasant, 2 = unpleasant, 3 = mildly unpleasant, 4 = mildly pleasant, 5 = pleasant, 6 = very pleasant) in a self-paced manner. After completion, the recognition phase began. Subjects were informed that the upcoming words would be tested for memory. For the study phases, participants reported the number of syllables for each study item using a (1, 2, 3, 4+) prompt. Syllable counting was chosen to promote an intermediate level of subsequent recognition performance. Following study, subjects were informed their memory would be tested for randomly intermixed studied and non-studied words, during which they should press the ‘A’ key if they believed the item was “studied” and the ‘L’ key if they believed it “non-studied.” They were also informed that either immediately following or preceding each recognition judgment, a “pleasantness?” prompt would appear and they should rate the word on a visible six-point scale (same used above). Key assignments were chosen so that there was no natural mapping between the classification and the pleasantness keys. Recognition and pleasantness judgments were self-paced.

Results

Basic recognition performance Because neither hit rates (.63 vs. .63; $t(25) = .02$, $p = .987$) nor false-alarm rates (.13 vs. .15; $t(25) = .67$, $p = .269$) were affected by the order of judgments (recognition judgment then pleasantness rating or the reverse),

we collapsed across this factor. Overall, subjects correctly responded on 74% of the recognition trials demonstrating moderate accuracy. They were conservatively biased, responding “old” for only 39% of the trials.

Pleasantness by recognition outcomes Turning to the effects of the recognition outcomes on pleasantness ratings, we initially conducted a three-way repeated-measures ANOVA with factors of Order (recognition judgment then pleasantness rating or the reverse), Response (“old” or “new”) and Accuracy (correct or incorrect). However, because Order did not interact with the other variables (p 's > .225) and was not significant in its own right ($p = .976$), we collapsed across this factor, leaving a two-way repeated-measures ANOVA on Response and Accuracy. The model yielded a main effect of Response ($F(1,25) = 34.78$, $\eta_p^2 = .58$, $p < .001$) with “old” recognition judgments accompanied by higher pleasantness ratings than “new” judgments. However, the main effect of Accuracy was not reliable ($F(1,25) = 2.01$, $\eta_p^2 = .07$, $p = .168$), nor was the interaction between Response and Accuracy ($F < 1$). Figure 1 plots the mean pleasantness ratings associated with the four possible recognition outcomes driving the ANOVA findings.

Figure 1 not only illustrates the robust effect of recognition response on the pleasantness ratings, it also suggest that these effects are symmetric about the normative pleasantness rating outside of the context of recognition, indicated by the solid line. We tested this via one-sample t-tests, comparing each response outcome to the baseline pleasantness rating of 3.52 during the norming ratings. For “old” judgments, hits were associated with higher pleasantness ratings than baseline ($M = 3.67$, $t(25) = 2.28$, $p = .015$), whereas false alarms trended in that direction ($M = 3.63$, $t(25) = 1.67$, $p = .108$). In the case of “new” recognition judgments, correct rejections and misses accompanied reliably lower pleasantness ratings than the baseline norm ($M = 3.40$, $t(25) = -2.31$, $p = .030$ and $M = 3.35$, $t(25) = -2.63$, $p = .015$, respectively).

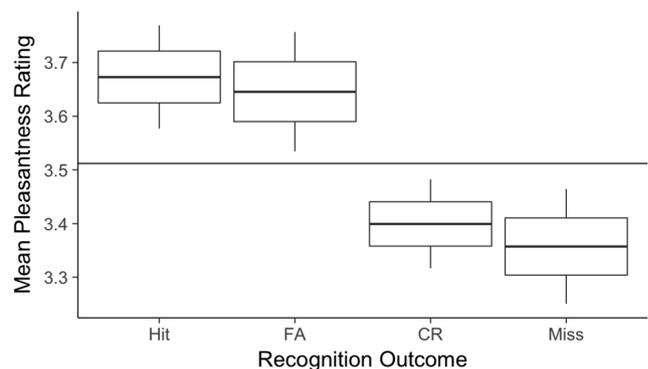


Fig. 1 Modified boxplot showing the relationship between recognition outcomes and mean pleasantness ratings of the materials. Thick lines indicate means. Box is ± 1 SEM whereas Box + Whisker is 2 SEMs. The horizontal line reflects the mean pleasantness rating observed during baseline norming of the materials (3.52)

Discussion

Participants rated items judged “old” as more pleasant than those judged “new,” regardless of the accuracy of these reports – replicating some prior findings (e.g., Lee, 2001; Wang & Chang, 2004). Moreover, “new” judgments yielded reliably lower pleasantness ratings than baseline norms. This appears to be the first documentation of this phenomenon, whereby the conclusion that an item is “new” actually lowers its perceived pleasantness relative to a baseline norming condition. Because both normed items and correctly rejected items are new to the experiment context, it must be that the act of concluding an item is “new” somehow lowers its perceived pleasantness.

Overall, these findings pose problems both for the PFM account and the explicit recognition account of MEEs. Beginning with the former, MEEs result from processing fluency, unless it is discounted during affective ratings because the subject believes the item was previously encountered. From this perspective, one would expect the fluency of hits to be heavily discounted whereas misses would presumably be processed fairly fluently but not discounted because the subject believes them to be new. Yet, as Fig. 1 shows, hits and misses fall at the extremes with respect to rated pleasantness, yielding a highly reliable difference in ratings ($t(25) = 5.51, p < .001$) with a large effect size (Cohen’s $d = 1.08$). This is not only in the reversed direction from that predicted under the PFM model; that model offers no mechanism to explain why items correctly judged as “new” yield pleasantness ratings below the normative baseline, since their processing fluency would be equivalent to baseline materials.

Aside from the PFM model, the newly observed pattern for correct rejections and misses is also not anticipated by an explicit recognition account (e.g., Newell & Shanks, 2007), although that model is consistent with the elevated pleasantness for hits and false alarms. Thus, overall, the negative affective response for items judged “new” calls for an explanation outside of the PFM and direct memory frameworks, and an ideal explanation would cover both the positive and negative affective responses documented here. The first possibility we consider is that the effects reflect the same motivational phenomenon. Under this motivated recognition account, we assume that observers view the goal of recognition tests as the identification or discovery of studied items and that goal-consistent responses yield positive affect and goal-inconsistent responses yield negative affect.

Experiment 2: Recognition memory, motivation, and the MEE

Here we manipulate motivational incentives for “studied” and “unstudied” conclusions using a design adapted from Han

et al. (2010). That study was motivated by a meta-analysis showing greater caudate activation for hits than CRs (Spaniol et al., 2009), even though caudate was historically linked with goal dependent actions, not explicit memory outcomes (Delgado, Locke, Stenger, & Fiez, 2003). Han et al. (2010), however, were able to amplify or reverse the caudate activation pattern depending upon whether “old” or “new” conclusions were incentivized and this led to the conclusion that caudate activation was signaling the goal status of recognition decisions, not the recovery of episodic information *per se*. We adapted a similar procedure to determine if the pattern in Fig. 1 could be reversed or mitigated by incentivizing either “old” or “new” recognition judgments using a point reward system.

Method

Participants One hundred and fifty-four Washington University in St. Louis undergraduates participated in exchange for course credit. We increased the sample size in this study to further rule out possible effects of order of judgment. Fifty-one participants were randomly assigned to the “new”-incentive condition, 52 randomly assigned to the “old”-incentive group, and 51 randomly assigned to the control group. Seven participants were discarded because of incomplete data linked to early termination of the program. Informed consent was obtained in compliance with the Institutional Review Board of Washington University in St. Louis. Testing occurred in groups of one to four people using individual computer carousels.

Materials Stimuli were the same as Experiment 1.

Design and procedure As in Experiment 1, a norming, pleasantness-ratings phase (identical to Experiment 1) was followed by the recognition experiment proper. There were three separate incentive groups. The “new”-incentive group received the following instructions:

“We are interested in the accuracy of your ‘non-studied’ responses. For each correct ‘non-studied’ response given you will potentially earn 10 points and for each incorrect ‘non-studied’ response given you will potentially lose 10 points. Points will not be affected by any ‘studied’ responses given and will be calculated at the end for a candy prize.”

Participants in the “old”-incentive group received mirror instructions. Subjects were given a performance summary after each test (e.g., “your score for ‘non-studied’ stimuli is X ”). Although the point system accurately tracked participants’ performance throughout all test blocks, following debriefing, participants were all awarded with five pieces of candy. No

mention of incentives or points was made to control subjects, whom simply received a candy bonus at the end of participation.

Results

Basic recognition data Neither the hit ($t(152) = .97, p = .333$), nor false-alarm rate ($t(152) = .08, p = .937$) was affected by the order of recognition and pleasantness rating and so we collapse across this factor in all subsequent analyses.

To test whether incentives affected recognition judgments, an Incentive Group (“old”-incentive, “new”-incentive, control) X Response (hit or false-alarm rate) mixed ANOVA was conducted on the proportion of “old” responses (Table 1). A change in accuracy would be suggested by an interaction between Incentive Group and Response (i.e., divergence or convergence of hit and false-alarm rates across groups), whereas a change in bias by a main effect of Incentive Group (a greater or lesser tendency to respond “old”). Unsurprisingly, the results revealed a large main effect of Response ($F(1,151) = 1759.63, \eta_p^2 = .92, p < .001$) with hit rates far exceeding false-alarm rates. However, neither the main effect of Group ($F(1,151) = 2.85, \eta_p^2 = .03, p = .061$) nor the interaction of Group and Response ($F < 1$) were reliable.

Pleasantness rating by recognition outcome The potential effect of incentives on pleasantness ratings was examined using and Incentive Group by Response (“old” vs. “new”) by Accuracy (correct vs. incorrect) mixed design 3-way ANOVA. There was a main effect of Response ($F(1,151) = 67.78, \eta_p^2 = .31, p < .001$) and Accuracy ($F(1,151) = 18.37, \eta_p^2 = .11, p < .001$). The former occurred because, as in Experiment 1, items garnering “old” responses were rated as more pleasant than those garnering “new” responses. The latter occurred because accurate recognition judgments were

accompanied by higher pleasantness ratings than inaccurate recognition judgments. Had the incentive manipulation moderated the relationship between recognition responses and pleasantness ratings, we would have observed an interaction between Incentive Group and Response, but there was no evidence for such an effect ($F(2,151) = 1.50, \eta_p^2 = .02, p = .226$). However, to further confirm that the incentives did not moderate the response-linked pleasantness effects, we removed the control group, thus directly comparing the new and old incentive groups across “old” and “new” responses. This again failed to achieve significance ($F(1,101) = 1.88, \eta_p^2 = .02, p = .173$). Thus, there is no evidence that differentially incentivizing old versus new responses reliably moderates the response-linked MEEs.

Turning to the main effects, both are clear in Fig. 2, which breaks the data down into hits, false alarms, misses, and correct rejections collapsed across Incentive groups. Although we didn’t find a main effect of accuracy in Experiment 1, it nonetheless demonstrated the same numerical trend as here, with accurate reports accompanied by higher pleasantness ratings than errors (see Fig. 1), suggesting that the failure to find a reliable accuracy effect in Experiment 1 was due to the smaller sample size.

Next, we compared pleasantness ratings for each of the collapsed recognition outcomes to the baseline pleasantness rating of 3.52. Hits yielded higher pleasantness ratings than baseline ($M = 3.64, t(153) = 3.63, p < .001$), whereas correct rejections ($M = 3.43, t(153) = -4.41, p < .001$) and misses ($M = 3.40, t(153) = -5.38, p < .001$) yielded reliably lower pleasantness ratings than baseline. Although numerically higher, the rated pleasantness of items garnering false alarms did not reliably differ from baseline ($M = 3.57, t(153) = .58, p = .564$).

Discussion

Experiment 2 tested whether the recognition-linked MEE pattern of Experiment 1 could be explained solely in terms of a differential motivation to respond “old” versus “new” during recognition testing. If so, then a direct motivation manipulation should have also reversed, or at least reliably altered, the association between MEE ratings and recognition conclusions. It did not, since there was no reliable interaction between Response and Incentive Group and “old” conclusions continued to yield higher pleasantness ratings than “new” conclusions, regardless of the incentives in place. Indeed, even when we examined the “new”-incentive group in isolation, a simple comparison demonstrated that “old” reports were accompanied by higher pleasantness ratings than “new” reports ($t(50) = 3.10, p = .003$), despite the fact that the rendering of new judgments was selectively incentivized in this condition.

Additionally, a new phenomenon was identified in the overall analysis (collapsed across groups), suggesting that correct responses also yielded slightly higher pleasantness ratings

Table 1 Means and standard deviations for proportion “old” response outcome as a function of a 3(Group) × 2(Response Outcome) design in Experiment 2

Group	Response Outcome			
	FA		Hit	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
OI	0.18	0.11	0.68	0.14
NI	0.14	0.09	0.63	0.17
X	0.15	0.08	0.63	0.13

Note. *M* and *SD* represent mean and standard deviation, respectively. Group nomenclature: NI = “new” incentivized, OI = “old” incentivized, X = control (no incentive)

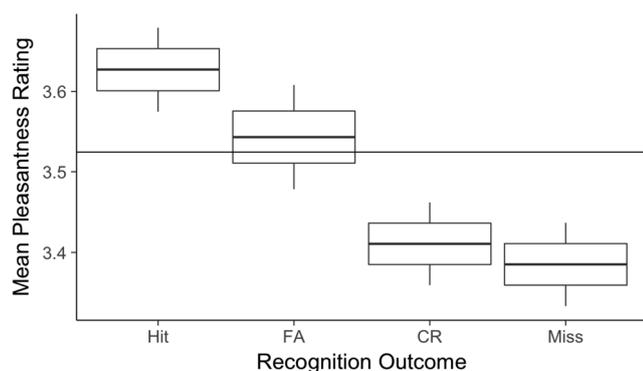


Fig. 2 Modified boxplot showing the relationship between recognition outcomes and mean pleasantness ratings of the materials (collapsed across Incentive group). Thick lines indicate means. Box is ± 1 SEM whereas Box + Whisker is 2 SEMs. The horizontal line reflects the mean pleasantness rating observed during baseline norming of the materials (3.52)

of the materials than incorrect responses (Fig. 2). Though this effect was not reliable in Experiment 1, examination of Fig. 1 shows that hits yielded ratings numerically more pleasant than false alarms and that correct rejections yielded ratings as numerically more pleasant than misses. Thus, the data tentatively suggest accurate judgments engender higher pleasantness ratings than inaccurate judgments.

The failure to find a motivation effect on pleasantness ratings is a null finding and so the possibility exists that the point manipulation may have not been salient enough. However, research from our lab (in progress) using a similar point payout procedure elicits marked differences in the dilation of the pupil during recognition decisions as a function of whether “old” or “new” responses are incentivized. As with the current data, neither recognition accuracy nor bias were affected, which is consistent with the fact that the point payout is neutral. Given the link between pupil dilation, arousal and attentional orienting, these data suggest that simple point payouts influence motivation-linked physiological processes. Moreover, point payouts that use unbalanced payouts appear to easily shift recognition decision biases (e.g., Curran, DeBuse, & Leynes, 2007) and so we feel it is unlikely that subjects are simply ignoring the contingencies. Nonetheless, below we provide an alternative model that, if supported, would render the motivational account unnecessary because it would accommodate the current findings (including Experiment 1) and the data of two following experiments.

Before turning to this new model we briefly explain why the order of pleasantness and memory decisions does not seem to matter for the affective responses. Similar to Experiment 1, the order of which memory and affective judgments were queried did not affect the pattern of ratings found; this same order insensitivity was documented by Newell and Shanks (2007). Inspection of reaction times from both Experiment 1 and Experiment 2 suggests that the order insensitivity reflects the fact that the latter decision is initiated prior to the execution

of the prior decision. For example, in Experiment 2 when memory decisions followed pleasantness ratings, they were remarkably faster (mean median 548 ms) than when they preceded the pleasantness ratings (mean median 1470 ms) ($t(152) = 19.5$, $p < .001$). Conversely, when pleasantness ratings followed the recognition decisions, they were reliably faster (mean median = 780 ms) than when they preceded them (mean median = 1946 ms) ($t(152) = 17.40$, $p < .001$). Thus, it appears the decisions are never reached in isolation, with the second judgment initiated well before the first has completed. Experiment 1 produced the same reaction time order effects for recognition ($t(50) = 9.86$, $p < .001$) and pleasantness judgments ($t(50) = 8.79$, $p < .001$). This suggests that when subjects know they must make both decisions for the same probe, the decisions are rendered at least in part coincidentally. Given this, and the absence of order effects on recognition or pleasantness conclusions, we used a single ordering of decisions for Experiments 3 and 4, collecting the pleasantness responses first.

The confirmation of search model

In Experiment 3, we test the hypothesis that the pattern of pleasantness findings in Experiments 1 and 2 reflect a confirmatory bias in memory search operations. Under this account, subjects use a directed memory search in an attempt to confirm a candidate model of a prior experience. Indeed, most accounts of explicit memory retrieval assume that memory evidence is evaluated in light of a rough template or model of the candidate prior experience (e.g., Johnson, Hashtroudi, & Lindsay, 1993; Polyn, Natu, Cohen, & Norman, 2005). Somewhat analogously, the finding that memory retrieval appears content addressable supports the idea that observers use a coarse retrieval description when querying memory (Norman & Bobrow, 1979). This raises the possibility that confirmatory bias in the search of memory colors the affective response to materials depending upon the search outcome, given subjects preference for confirmation (Nickerson, 1998).

More specifically, recovering memory evidence consistent with a retrieval description (a search success or confirmation) may yield a positive emotional response whereas failing to retrieve evidence consistent with that description may yield a negative emotional response (a search failure or disconfirmation/negation). Moreover, the effort expended during the search may also mediate the emotional response, with greater search effort yielding more negative affect. Thus, under this confirmation of search (COS) hypothesis, recognition judgments are viewed as search or monitoring operations in which an observer seeks memory evidence capable of confirming a candidate hypothesis that the item originates from a particular context. Successfully recovering sufficient evidence yields a slight positive emotional response, but failing to do so yields a

slight negative response. Moreover, the effort expended in evaluation or monitoring may also moderate an emotional response, leading to more negative (or less positive) responses during errors than correct judgments. This would merely reflect that errors require more deliberation or effort than correct judgments. Critically, as with the PFM model, this confirmation of search account is still a misattribution model because under the model the observer is conflating emotional responses following search outcomes with an emotional response to the stimulus itself. In Experiment 3, we test the COS model by actively manipulating the memory characteristics that serve as confirmations during source memory.

Experiment 3: Source memory and the MEE

During source memory paradigms observers study materials in two differing contexts and are later asked to discriminate these materials (and perhaps novel items) by correctly indicating their origin. Under the PFM model, items from the two sources should share similar fluency levels, which means that affective ratings for the two sources should be similar as a whole.

In contrast to the PFM explanation, the COS model assumes that the association between memory judgments and the rated pleasantness of materials reflects the outcome of confirmatory memory searches, with confirmed searches yielding higher pleasantness ratings than disconfirmed searches. This can be easily tested in the context of a source memory experiment by altering the framing of the source memory question. For example, Dobbins and McCarthy (2008) demonstrated that the confirmatory framing of source memory queries ('Source A?' vs. 'Source B?') could reveal recognition based confirmatory biases when the probes originating from these two sources were differentially familiar.

In the current study, we use this source cue framing manipulation ('Source A?' vs. 'Source B?') to test the COS model. For items studied under the two sources, the prediction is that when subjects are asked if items originate from a particular source (Source A?) they will provide higher pleasantness ratings when responding positively than negatively to the source memory query for studied items. Conversely, when asked if items originate from Source B they will provide higher pleasantness ratings when responding positively than negatively. If so, then even though the items from these two sources are matched in terms of fluency, their perceived pleasantness would be altered by the outcome of the initial confirmatory memory search. This is the first time that source memory outcomes have been investigated with respect to MEEs and the first time that confirmatory status of source memory judgments has been predicted to alter affective responses to materials.

Method

Participants Thirty-four Washington University in St. Louis undergraduates participated in exchange for course credit. Informed consent was obtained in compliance with the Institutional Review Board of Washington University in St. Louis. Testing occurred in groups of one to four people in the same lab context as Experiments 1 and 2.

Materials Stimuli were the same as Experiments 1 and 2. However, words were presented via computers running PsychoPy software (Peirce, 2007).

Design and procedure Identical to Experiments 1 and 2 there was a norming phase (150 items) followed by a memory experiment. Both study-test cycles of the memory experiment consisted of 100 to-be-encoded source words divided evenly into two encoding sources (abstract/concrete judgments or pronounceability judgments). During subsequent testing these were combined with 50 novel items and source memory was tested (150 words per test). For each test, source memory was queried in one of two manners (i.e., "Pronounceability task?" or "Concrete/Abstract task?"). The order of the query across the tests was counterbalanced. During each of the two study phases half the words were classified as either abstract or concrete, or easy or difficult to pronounce, in a randomized manner. During subsequent source memory testing, the pleasantness of each memory probe was rated before the memory judgment and rendered using the same scale as the initial normative ratings. Then, depending upon which test block the participant was in, he or she was asked to confirm or disconfirm a source using the query "did you study this word with the abstract/concrete task?" or the query "did you study this word with the pronounceability task?" for the entire test. Participants used the 'A' key to indicate 'yes' for the source query and the 'L' key to indicate 'no.' Key assignments were, again, chosen so that there was no natural mapping between the memory classification and the pleasantness keys (numeric key pad). All responding was self-paced.

Results

Source recognition data Table 2 below shows the 'yes' rates as a function of the format of the source query and the origin of the probes.

To evaluate source accuracy for the two queries, we restricted the analysis to studied items and contrasted the proportion of correct responses for the two cue framing conditions. Subjects were reliably more accurate under the abstract/concrete cue (ACT?) ($M = .66$) than under the pronounceability cue (PC?) ($M = .56$, $t(32) = 4.00$, $p < .001$). However, they were similarly inclined towards 'yes' responses under the two cue conditions ($M = .50$ vs. $.51$; $t(32) = -0.32$, $p = .751$).

Table 2 Means and standard deviations for ‘yes’ rates as a function of a 2(Query) × 3(Source) design in Experiment 3

Query	Source					
	ACT		NEW		PT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ACT?	0.67	0.19	0.15	0.16	0.34	0.19
PT?	0.45	0.24	0.18	0.16	0.57	0.15

Note. *M* and *SD* represent mean and standard deviation, respectively. *ACT* abstract/concrete task, *PT* pronounceability task

Turing to novel probes, these intruded during the source judgments at a similar rate under the two cue framings (Table 2), which was not reliably different via *t* test ($t(32) = 1.21, p = .236$).

Source memory outcome and pleasantness rating To simplify the analyses we focus on words from the studied sources, and performed separate analyses for correct versus incorrect judgments. Beginning with correct judgments, Table 3 suggests that correct endorsements of the source queries (‘yes’) yield higher pleasantness ratings than correct rejections (‘no’) of the same materials.

Confirming this impression a 2-way repeated measures with factors of Query (ACT? or PT?) and Source (ACT or PT) demonstrated a significant interaction between the factors ($F(1,32) = 25.21, \eta_p^2 = .44, p < .001$) with no main effects of Query ($F < 1$) or Source ($F(1,32) = 1.56, \eta_p^2 = .05, p = .220$). Simple pairwise comparisons demonstrated that pleasantness ratings were higher for items from the ACT source when they were correctly endorsed under the ACT? cue ($M = 3.73$) than correctly rejected under the PT? cue ($M = 3.50$) ($t(32) = 2.76, p = .010$). Analogously, pleasantness ratings were higher for items from the PT source when they were correctly endorsed under the PT? cue ($M = 3.71$) than correctly rejected under the ACT? cue ($M = 3.41$) ($t(32) = 3.46, p = .002$). Thus, these findings indicate that during correct responding, confirmations

Table 3 Means and standard deviations for pleasantness ratings as a function of a 2(Query) × 2(Source) design in Experiment 3 (correct responses only)

Query	Source			
	ACT		PT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ACT?	3.73 (yes)	0.39	3.41 (no)	0.36
PT?	3.50 (no)	0.45	3.71 (yes)	0.42

Note. *M* and *SD* represent mean and standard deviation, respectively. *ACT* abstract/concrete task, *PT* pronounceability task

of a source query yield higher pleasantness ratings of the materials than negations.

We next considered incorrect responses for the source materials to see if they demonstrated the same pattern such that confirmations yielded greater pleasantness than negations. Table 4 suggests they did.

Confirming this impression a 2-way repeated-measures analysis with factors of Query (ACT? or PT?) and Source (ACT or PT) demonstrated a significant interaction ($F(1,32) = 13.70, \eta_p^2 = .30, p < .001$) with no main effects of Query ($F(1,32) = 1.41, \eta_p^2 = .04, p = .243$) or Source ($F(1,32) = 1.41, \eta_p^2 = .04, p = .244$). Simple pairwise comparisons demonstrated that pleasantness ratings were higher for items from the ACT source when they were *incorrectly* endorsed under the PC? cue than *incorrectly* rejected under the ACT? cue ($t(32) = 3.49, p = .001$). Analogously, pleasantness ratings were higher for items from the PT source when they were *incorrectly* endorsed under the ACT? cue versus *incorrectly* rejected under the PC? Cue; however, this difference only trended towards significance ($t(32) = 1.85, p = .074$). Thus, these findings converge with those from the correct judgments in demonstrating that observers rate the pleasantness of the probes higher when they endorse them as coming from the queried source than when they reject them as arising from that source. Moreover, this occurs even when judgments are erroneous.

The next question we consider is whether, aside from the confirmation/disconfirmation effects documented above, there is also an effect of accuracy on the rated pleasantness of the materials. To do this, we again focused on studied items, but this time we collapsed across the queries and sources and simply compared correct source judgments ($M = 3.60$) to incorrect source judgments ($M = 3.56$). Although numerically larger, pleasantness ratings accompanying accurate source judgments were not reliably different from those accompanying inaccurate source judgments ($t(32) = 1.51, p = .145$).

Finally, similar to the recognition data, we again compared the pleasantness ratings accompanying the source memory outcomes against the normative baseline. Figure 3 shows the four outcomes of interest, namely, source memory hits,

Table 4 Means and standard deviations for pleasantness ratings as a function of a 2(Query) × 2(Source) design in Experiment 3 (error responses only)

Query	Source			
	ACT		PT	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ACT?	3.28 (no)	0.59	3.69 (yes)	0.55
PT?	3.69 (yes)	0.51	3.48 (no)	0.46

Note. *M* and *SD* represent mean and standard deviation, respectively

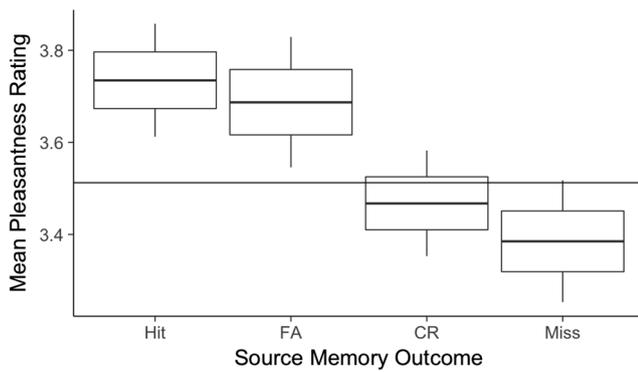


Fig. 3 Modified boxplot showing the relationship between source memory outcomes and mean pleasantness ratings of the materials. Thick lines indicate means. Box is ± 1 SEM whereas Box + Whisker is 2 SEMs. The horizontal line reflects the mean pleasantness rating observed during baseline norming of the materials (3.52)

misses, correct rejections, and false alarms to studied materials. In this context a miss is an incorrect “no” report to an item that was from the queried source. In contrast, a false alarm is an incorrect “yes” to an item that was from the inappropriate source given the query. Both source memory hits ($M = 3.74$, $t(32) = 3.43$, $p = .002$) and source memory false alarms ($M = 3.69$; $t(32) = 2.30$, $p = .028$) yielded pleasantness ratings higher than baseline. In contrast, source memory misses ($M = 3.38$, $t(32) = -2.11$, $p = .042$) but not source memory correct rejections ($M = 3.47$, $t(32) = -.99$, $p = .328$) yielded ratings reliably below baseline.

Discussion

Experiment 3 demonstrated that the rated pleasantness of probes was influenced by whether or not a source query led to a confirmation or a disconfirmation. Patterns similar to those of Experiments 1 and 2 emerged even when restricting focus entirely to studied materials that should have similar fluency. Thus, it is hard to see how the fluency-based PFM model of MEEs during recognition would also accommodate this pattern, particularly since the fluency of the studied materials presumably is unrelated to the framing of the source memory questions. The findings are also inconsistent with the explicit recognition account of Newell and Shanks (2007) in which MEEs reflect the explicit recognition of the probes. Under this account, there is no reason why correctly responding “no” to items from a particular source (source correct rejection) should yield a negative affective response. Indeed, an explicit recognition account predicts a positive affective response to these items on average, if one assumes that they are rejected because the observer realizes they originated from the wrong source. Instead, the simplest account of the pattern of Experiment 3 is that source memory confirmations yielded a positive affective response, whereas source memory

negations yielded a negative affective response, consistent with the confirmation of search model.

Across-experiment analyses (1–3)

Response and accuracy

Before turning to Experiment 4, we analyze data collapsed across all three experiments to address two questions. First, we consider whether the effect of accuracy on pleasantness ratings is generally reliable. In all experiments, inaccurate memory judgments yielded pleasantness ratings slightly lower than accurate judgments (Figs. 1, 2, and 3). When considered in isolation this was only statistically reliable in Experiment 2, but the numerical consistency across the three experiments suggests there is a likely a modest reduction in rated pleasantness accompanying erroneous memory decisions. Indeed, when the data are combined across the three experiments, an ANOVA with factors of Experiment, Response (confirmation or disconfirmation) and Accuracy yields a reliable main effect of Accuracy ($F(1,210) = 14.17$, $\eta_p^2 = .06$, $p < .001$), indicating a small but reliable reduction of rated pleasantness for inaccurate compared to accurate responding.

Item-selection artifacts

The second question we consider is the relationship of each of the four possible memory outcomes (hit, false alarm, correct rejection or miss) to the normative baseline. In the analyses within each experiment, we contrasted each outcome to the grand mean of the normative pleasantness ratings of 3.52. This is warranted and powerful as long as there is no interaction between normative item pleasantness and actual memorability. For example, if more pleasant items were better encoded, then subsequent hits would be rated as more pleasant than subsequent misses but this would not necessarily mean that the memory decision itself was directly affecting the pleasantness rating. Alternatively, items garnering false alarms might have had generally higher pre-experimental fluency (hence driving the false alarm behavior) and thus, under fluency accounts, would be expected to yield high pleasantness ratings. More generally then, any systematic relationship between baseline item pleasantness and outcomes, or between baseline item pleasantness and encoding, might lead to an item selection artifact such that differences in the outcome bins do not directly reflect decision processes altering affect. Instead, these differences would reflect that items with differing levels of pre-experimental pleasantness or familiarity are ending up in different memory outcome bins.

To address this, it is necessary to control for item selection when considering potential effects of memorial decisions on rated pleasantness, which cannot be done by using the overall baseline pleasantness as a comparison value. For instance, if

during recognition a subject missed “pickle,” “napkin,” and “freedom,” the question is whether his or her rated pleasantness for these particular materials differs from (i.e., is lower than) the normative ratings for “pickle,” “napkin,” and “freedom,” specifically. If so, then it is clear that the memory decision outcome is altering the ratings for these items, as opposed to their having been missed due to lower pre-experimental pleasantness or familiarity.

For each memory outcome (hits, false alarms, correct rejections, and misses) we conducted a two-way mixed ANOVA with a between groups factor of Experiment and repeated measures factor of Rating (subject’s mean rating vs. the normative mean *for those specific items*). In none of the four analyses did Experiment yield a significant main effect or interact with Rating. For hits there was a reliable main effect of rating ($F(1,210) = 18.08, \eta_p^2 = .08, p < .001$) demonstrating they increased rated pleasantness relative to baseline. For false alarms the ratings were numerically higher than baseline, but not reliably so ($F(1,210) = 2.76, \eta_p^2 = .01, p = .10$). For correct rejections the ratings were reliably below baseline ($F(1,210) = 9.40, \eta_p^2 = .04, p = .002$), which also occurred for misses ($F(1,210) = 16.52, \eta_p^2 = .07, p < .001$). These findings converge with those of the prior analyses, and rule out an item selection interpretation. Correct confirmatory memory reports increase rated pleasantness, whereas both correct and incorrect disconfirmatory reports actually decrease rated pleasantness. False alarms did not reliably increase rated pleasantness but this likely reflects the fact that they are infrequent (increasing variability) and that errors yield lower pleasantness ratings than correct responding (pushing pleasantness towards baseline).

Experiment 4: Paired associate cued-recall and the MEE

Paired-associate learning tasks involve encoding of A-B pairs and subsequent cued recall of one of the paired items when only one associate is re-presented (the A cue). We tested the confirmation of search hypothesis by having subjects rate the pleasantness of A-cues prior to each cued-recall trial. Crucially, since all of the A-cues are pre-exposed in the same manner during study, any differences in rated pleasantness across successful and unsuccessful recall are unambiguously the result of the recall outcome. The predictions of the COS model are thus straightforward in that observers should demonstrate increased pleasantness rating of the A-cue when retrieval of B-associate occurs and decreased pleasantness ratings when recall does not occur (relative to a normative baseline). To further ensure that recall outcome, as opposed to pre-experimental characteristics of the cues, was the driving factor we used meaningless cues (see below). This is the first investigation of the effects of cued-recall outcomes on the MEE and

the first to propose that retrieval of B-associates affects the rated pleasantness of A-cues.

Method

Participants Thirty-two Washington University in St. Louis undergraduates participated in exchange for course credit. Informed consent was obtained in compliance with the Institutional Review Board of Washington University in St. Louis. Testing occurred in groups of one to four people in the same lab context as all previous Experiments.

Materials Word pairs were constructed from a set of 120 Lithuanian (e.g., *batas*) and 120 English nouns (e.g., *tomato*), A-B pairs. Word pairs were randomized so as not to be direct translations. Stimuli were presented using PsychoPy software (Peirce, 2007).

Design and procedure Similar to Experiments 1–3, a norming phase preceded the memory experiment during which participants rated 60 words not used in his or her session. These ratings were for the Lithuanian cue materials. Following this, two study-test cycles of the memory experiment consisted of 35 to be encoded Lithuanian-English (A-B) word-pairs.

The study phase used the *keyword method* (Atkinson & Raugh, 1975) to facilitate A-B encoding. Each Lithuanian-English pair (*batas-tomato*) was presented serially on screen for 15 seconds, and participants were instructed to find a “keyword” in the Lithuanian word familiar in English (“bat” in “*batas*”) and use imagery to combine the English “keyword” with the English associate (*tomato*) in bizarre interactions. Explicit examples were given (e.g., someone beating a tomato with a bat). These procedures yield decent recall with moderate levels of study (McDaniel & Pressley, 1989).

For each test phase 35 previously studied A-cues (*batas*) were presented serially on screen and in random order. Participants rated the pleasantness of each A-cue, before attempting cued recall. Following the pleasantness rating, the cue stayed on screen and the B associate was typed in via keyboard (*batas - _____*). Pressing the “enter” key allowed subjects to confirm their typed response and continue to the next trial. If the subject could not recall the associate, they pressed “enter” to proceed to the next trial. All responding was self-paced.

Results

Recall performance Subjects correctly recalled approximately half of the targets ($M = .49, SD = .23$). For unsuccessful trials, omissions (“Opt Out” trials) were more numerous ($M = 27.47, SD = 12.26$) than intrusions ($M = 8.09, SD = 9.00$).

Cued-recall outcome and pleasantness rating A one-way ANOVA on mean pleasantness ratings indicated significant differences across the cued recall outcomes (Successful recall, Intrusion, and Opt-Out) ($F(2, 62) = 16.60, \eta_p^2 = .35, p < .001$) (Fig. 4).

As with recognition and source memory, we tested the pleasantness ratings accompanying each cued-recall outcome against the mean normative pleasantness rating (3.59) to determine if confirmations or disconfirmations altered rated pleasantness relative to the mean normative value. While pleasantness ratings prior to intrusions failed to differ from baseline ($M = 3.64, t(31) = .370, p = .714$), those preceding successful cued-recall were greater than baseline ($M = 3.91, t(31) = 3.68, p < .001$) and those preceding “Opt Out” trials ($M = 3.31, t(31) = -3.05, p = .005$) were lower than baseline.

Item analysis

Finally, as with the combined data sets above, we also contrasted subjects’ rated pleasantness matched at the item level. For example, if a subject rated 12 items before committing intrusions during the subsequent cued recall judgment, his or her mean rated pleasantness for those 12 items was compared against the mean of the normative values for the same 12 items. Paired t-tests replicated the item analyses above. Pleasantness ratings prior to intrusions failed to differ from the matched item norm ($t(31) = .03, p = .972$), those preceding successful recall were greater than the matched item norm ($t(31) = 2.65, p = .013$), whereas those preceding “Opt Out” trials ($t(31) = -2.72, p = .011$) were lower than the matched item norm.

Discussion

Experiment 4 showed that pleasantness ratings of cues were higher than normative when preceding successful cued recall

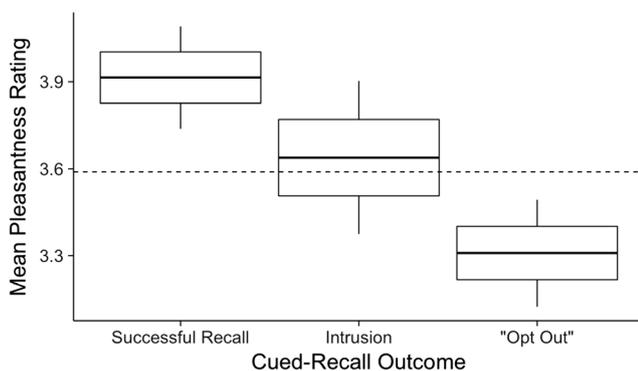


Fig. 4 Modified boxplot showing the relationship between cued-recall outcomes and mean pleasantness ratings of the materials. Thick lines indicate means. Box is ± 1 SEM whereas Box + Whisker is 2 SEMs. The horizontal line reflects the mean pleasantness rating observed during baseline norming of the materials (3.59)

(i.e., search confirmations), and lower when preceding “opt out” trials (i.e., search failures). They were intermediate and not reliably different from normative when intrusions followed. This likely reflects the fact that while responses are produced during intrusions, potentially elevating pleasantness, the trials are also presumably difficult given that the responses are incorrect. As shown earlier, errors are associated with declines in pleasantness ratings and so this may have offset any increase in ratings. Regardless, neither the PFM nor explicit recognition accounts make clear cued-recall predictions for the correct and opt-out conditions. In the case of the explicit recognition account, one might suggest that the successful retrieval of the associate should elevate rated pleasantness. However, to date, that account has assumed that it is item recognition of the evaluated probe that causes the increase (Newell & Shanks, 2007, p. 117). If so, then there is no reason to necessarily expect an increase during associative retrieval since the retrieved information is of another item that is not the focus of the pleasantness rating. Nonetheless, if a framework were developed to broaden the direct retrieval account to associative retrieval, it still would not explain the devaluation occurring on “Opt-Out” trials. Thus, the simplest account is that cued-recall confirmations yielded a positive affective response, whereas cued-recall failures (“Opt-Out” trials) yielded a negative affective response.

General discussion

We sought to better understand why the rated pleasantness of memory probes covaries with memory status and judgments. Under the PFM model, this is assumed to reflect subject attributions about probe fluency. This model encountered several difficulties. For example, it failed to predict that misses during recognition would be rated as the least pleasant of all outcomes; indeed, the model would seem to suggest that misses should yield positive affective responses because they are processed fluently yet believed to be new. Additionally, hits yielded increases in pleasantness ratings despite observers correctly realizing that they had been encountered, which under the PFM model should have led to fluency to being actively discounted when rating pleasantness. Thus, hits should be rated near baseline levels for pleasantness. Finally, the PFM model does not predict differences for studied probes during source memory paradigms as a function of source memory conclusions and cue frames, and it also fails to anticipate the cued recall findings.

Though the explicit recognition account, whereby MEEs result from the recovery of explicit memory information about the probe, fares better, it only accounts for approximately half of the findings; namely, increases in rated pleasantness for memory probes that are recognized (Experiments 1 and 2) or yield accurate source memories (Experiment 3). In the case of

accurate cued-recall in Experiment 4, it is unclear whether the explicit recognition account applies because it focuses on recognition of the probe as the basis of increased perceived pleasantness. If this increase results from item memory strength, then it is not clear that successful retrieval of a paired associate should increase rated pleasantness of the cue because associative retrieval of the target doesn't necessarily increase the recognition strength of the cue itself. Indeed, in Experiment 4, the subjects knew all of the Lithuanian probes were studied materials and so item recognition of the cues is potentially moot. Nonetheless, all four experiments also showed clear devaluation of rated probes that yielded negative memory decisions, including correct rejections and misses during recognition, source memory misses, and failed cued-recall attempts. These findings are not anticipated by an explicit recognition account. While one could employ two theoretical mechanisms to explain the positive versus negative judgment effects of affective responding, this would not be parsimonious, if a single account suffices.

In contrast, the confirmation of search (COS) model correctly predicted how positive and negative memory judgments in Experiments 3 and 4 would affect pleasantness ratings and is also consistent with the recognition findings of Experiments 1 and 2 (Figs. 1 through 4). It also helps to explain findings in which memory confirmations and the accuracy of those outcomes facilitated positive affect (e.g., see Wang & Chang, 2004, for investigation of affective ratings in concert with remember/know judgments). Moreover, it is considerably simpler than the PFM model because it does not assume that subjects strategically regulate the attribution of fluency during each separate probe encounter; a strategy that would require considerable cognitive theorizing on the part the participants who are otherwise engaged in a fairly demanding memory task. Instead, the COS model simply assumes that confirmations of memory search evoke increased positive affect whereas disconfirmations yield negative affect and it is this affective response that is unwittingly ascribed to the probe during concomitant pleasantness ratings. There is no assumption that subjects actively try to regulate or direct this process during the affective judgments and, indeed, they are presumably unaware of the link between their memory conclusions and rated pleasantness.

Interestingly, if the COS model is correct, then appeals to 'mere exposure' in explaining the patterns of pleasantness ratings in the context of explicit memory demands should be reconsidered, since it is clear the mere exposure of the items does not explain the current patterns. More specifically, neither the dissociation of rated pleasantness by source confirmation/disconfirmation, cued-recall outcomes, nor the boost in rated pleasantness observed for false alarms, can be better understood by appeals to mere exposure. Although the COS model explains the effects of memory search outcomes on affective responses, it is not meant to replace or challenge prior findings that clearly show increased fluency can alter

memorial decisions (Jacoby & Whitehouse, 1989), more positive ratings follow subliminal presentations than supraliminal (e.g., Bornstein & D'Agostino, 1992) or that exposed stimuli accompany more positive ratings than novel (e.g., Zajonc, 1968). Instead, the current findings and model specifically address how the outcomes of memory decisions affect rated pleasantness of memory probes, relying on the widespread tendency of observers to favor confirmations over disconfirmations both generally, and more specifically, in the context of explicit memory search.

There are several reasons why one might assume that the outcome of memory searches would have affective consequences, particularly if one takes the idea of 'search' somewhat literally. In an actual physical search for an object (e.g., searching for one's keys), confirmation signals termination of search (and acquisition of the desired object) whereas disconfirmation signals the need to formulate another candidate location and explore that location. Thus, disconfirmations in actual search are usually linked to the need for additional effort and the delay of a desired outcome. Similarly, in memory search outside of the laboratory, the disconfirmation of a candidate memory source (Did I meet person X before at event Y?) means that an additional candidate sources must be formulated and assessed and that a firm attribution of the person's identity remains elusive. From this perspective, it seems reasonable to anticipate a positive emotional response to confirmations versus a negative response to disconfirmations of memory search attempts. However, the fact that this might occur during simple laboratory recognition and source memory tasks is somewhat surprising since these often entail mutually exclusive origins for the materials. That is, if an item is judged as not old during recognition then the necessary conclusion must be that it is new. In the case of the current source memory paradigm, if an item is not from the queried source then the necessary response is "no." In both cases, a disconfirmation of putative search does not necessarily entail the need to initiate a new search operation and expend additional effort. Thus, the negative affective response that results from disconfirmations must be assumed to operate fairly reflexively, perhaps reflecting the fact that outside of the lab, memory search disconfirmations are almost always accompanied by continued uncertainty determining the status of an encountered person or object (see also Lee, 2001).

Although the COS framework well explains the current patterns of pleasantness ratings, it does not explain why an affective response to a search outcome "bleeds into" the affective rating of the materials. However, prior work in social psychology has well documented the tendency of affective reactions to bleed into stimulus judgments; particularly, when those judgments in isolation would be somewhat ambiguous. The affect misattribution procedure (AMP) (for review, see Payne & Lundberg, 2014) is one useful example. In this paradigm, ambiguous stimuli (e.g., Chinese characters) are interleaved with photos having clear valence (e.g., puppies, snakes, etc.)

and observers are asked to rate the pleasantness of the ambiguous stimuli. Even when forewarned that the interleaved photographs may influence their responding, the valence of the photographs nonetheless affects the ratings of the characters and considerable research suggests that this influence is unconscious (Winkielman, Zajonc, & Schwarz, 1997).

Given the effect of confirmation emerged across three distinct episodic memory tasks, it may affect a host of judgment domains linking subtle affective responses to the act of confirmatory classification attempts (e.g., basic semantic classification tests). Thus, investigating the present breadth of this effect in other decision contexts may prove fruitful. Moreover, an exploratory analysis on the current data suggests that the magnitude of the emotional response to positive versus negative memory search outcomes may be an important individual difference. To investigate this, we z-scored each subject's pleasantness ratings in each of the four experiments and then correlated correct positive search outcomes with negative search outcomes. In Experiments 1 and 2, there were reliable negative correlations between pleasantness ratings in the context of hits versus correct rejections ($r = -.79$, $p < .001$; $r = -.86$, $p < .001$, respectively). Thus, the subjects reporting more positive pleasantness ratings during confirmations also report more negative pleasantness ratings during disconfirmations. Likewise, in Experiment 3, there was a negative correlation between rated pleasantness during source memory hits versus source memory correct rejections of studied materials ($r = -.41$, $p = .019$). Finally, in Experiment 4 there was a trend towards a reliable negative correlation for rated pleasantness in the context of successful cued-recall versus opt-out trials ($r = -.32$, $p = .072$). Overall, this suggests that emotional or motivational dispositions may mediate these decision-linked affective responses and supports the COS approach of interpreting both positive and negative effects within a common framework.

Future investigation of the COS model considering questions in consumer psychology may also be informative. Unlike the predictions of the PFM model, which prohibit exposure effects from occurring when accompanied by explicit memory, the COS hypothesis fits well with conceptions of *brand knowledge* and *brand awareness* in consumer research and brand management (see Esch, Langner, Schmitt, & Geus, 2006; Percy & Rossiter, 1992). For example, in the domain of consumer behavior one might test if confirming products as originating from certain advertising sources (e.g., Brand X's advertisement) bears greater influence on consumers' affective preferences than exposure to the product *per se*.

Conclusion

The current findings indicate that classic notions of the MEE are likely inappropriate in the context of explicit memory decision-making and demonstrate multiple new affective phenomena

during various types of memory judgment. In these situations, it is likely not the fluency of the materials that is altering affective responses but whether the memory conclusions confirm or disconfirm the initial search and the degree to which the memory judgment is arrived at easily. Critically, this confirmation of search model does not apply to situations in which explicit memory decisions are not being rendered; in such cases the fluency of the materials may play a dominant role, as suggested by models such as the PFM. Given this, we suggest maintaining a clear methodological distinction between decisional versus non-decisional influences on affective ratings. During the latter, such as in the classic MEE studies (originally envisaged by Zajonc, 1968), stimuli are exposed and then subsequently affectively rated without any additional explicit decisions accompanying the ratings. In this case, fluency misattribution may be a dominant moderator of affective ratings. In contrast, the current data demonstrate that when observers are rendering memory judgments, concurrent affective ratings are heavily moderated by the outcome of these judgments. Moreover, because the data suggest a confirmatory/disconfirmatory mechanism is at play, there are potentially a host of judgments about the stimuli, outside of explicit memory, that may flavor the affective response. Under these situations, "mere exposure" and fluency constructs are not informative.

References

- Anand, P., Holbrook, M. B., & Stephens, D. (1988). The formation of affective judgments: The cognitive-affective model versus the independence hypothesis. *Journal of Consumer Research*, 386-391.
- Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of experimental psychology: Human learning and memory*, 1(2), 126-159.
- Berry, C. J., Shanks, D. R., & Henson, R. N. (2006). On the status of unconscious memory: Merikle and Reingold (1991) revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 925-934.
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological bulletin*, 106(2), 265-289.
- Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of personality and social psychology*, 63(4), 545-552.
- Bornstein, R. F., & D'Agostino, P. R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect. *Social Cognition*, 12(2), 103-128.
- Brooks, J. O., & Watkins, M. J. (1989). Recognition memory and the mere exposure effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 968-976.
- Curran, T., DeBuse, C., & Leynes, P. A. (2007). Conflict and criterion setting in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 2-17.
- de Zilva, D., Vu, L., Newell, B. R., & Pearson, J. (2013). Exposure is not enough: Suppressing stimuli from awareness can abolish the mere exposure effect. *PLoS one*, 8(10), e77726.

- Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cognitive, Affective, & Behavioral Neuroscience*, 3(1), 27-38.
- Dobbins, I. G., & McCarthy, D. (2008). Cue-framing effects in source remembering: A memory misattribution model. *Memory & cognition*, 36(1), 104-118.
- Elliott, R., & Dolan, R. J. (1998). Neural response during preference and memory judgments for subliminally presented stimuli: a functional neuroimaging study. *Journal of Neuroscience*, 18(12), 4697-4704.
- Esch, F. R., Langner, T., Schmitt, B. H., & Geus, P. (2006). Are brands forever? How brand knowledge and relationships affect current and future purchases. *Journal of Product & Brand Management*, 15(2), 98-105.
- Fang, X., Singh, S., & Ahluwalia, R. (2007). An examination of different explanations for the mere exposure effect. *Journal of consumer research*, 34(1), 97-103.
- Fox, S. E., & Burns, D. J. (1993). The mere exposure effect for stimuli presented below recognition threshold: A failure to replicate. *Perceptual and motor skills*, 76(2), 391-396.
- Han, S., Huettel, S. A., Raposo, A., Adcock, R. A., & Dobbins, I. G. (2010). Functional significance of striatal responses during episodic decisions: recovery or goal attainment? *The Journal of Neuroscience*, 30(13), 4767-4775.
- Jacoby, L. L., & Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin*, 13(3), 314-336.
- Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118(2), 126-135.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin*, 114(1), 3-28.
- Lee, A. Y. (1994). The mere exposure effect: Is it a mere case of misattribution. *Advances in consumer research*, 21, 270-275.
- Lee, A. Y. (2001). The mere exposure effect: An uncertainty reduction explanation revisited. *Personality and Social Psychology Bulletin*, 27(10), 1255-1266.
- McDaniel, M. A., & Pressley, M. (1989). Keyword and context instruction of new vocabulary meanings: Effects on text comprehension and memory. *Journal of Educational Psychology*, 81(2), 204-213.
- Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, 11(6), 462-466.
- Newell, B. R., & Shanks, D. R. (2007). Recognising what you like: Examining the relation between the mere-exposure effect and recognition. *European Journal of Cognitive Psychology*, 19(1), 103-118.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), 175-220.
- Norman, D. A., & Bobrow, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 11(1), 107-123.
- Payne, B. K., & Lundberg, K. (2014). The Affect Misattribution Procedure. *Social and Personality Psychology Compass*, 8(12), 672-686.
- Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, 162(1), 8-13.
- Percy, L., & Rossiter, J. R. (1992). A model of brand awareness and brand attitude advertising strategies. *Psychology & Marketing*, 9(4), 263-274.
- Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, 310(5756), 1963-1966.
- Sawyer, A. G. (1981). Repetition, cognitive responses, and persuasion. *Cognitive responses in persuasion*, 237-261.
- Spaniol, J., Davidson, P. S., Kim, A. S., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia*, 47(8), 1765-1779.
- Stafford, T., & Grimes, A. (2012). Memory enhances the mere exposure effect. *Psychology & Marketing*, 29(12), 995-1003.
- Wang, M. Y., & Chang, H. C. (2004). The mere exposure effect and recognition memory. *Cognition & Emotion*, 18(8), 1055-1078.
- Wilson, M. (1988). MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, 20(1), 6-10.
- Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: psychophysiological evidence that processing facilitation elicits positive affect. *Journal of personality and social psychology*, 81(6), 989-1000.
- Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. *The psychology of evaluation: Affective processes in cognition and emotion*, 189-217.
- Winkielman, P., Zajonc, R. B., & Schwarz, N. (1997). Subliminal affective priming resists attributional interventions. *Cognition & Emotion*, 11(4), 433-465.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of personality and social psychology*, 9(2p2), 1-27.

Open Practices Statement Data and materials are available upon request. No experimenters are preregistered.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.