

Distinguishing between attributional and mnemonic sources of familiarity: The case of positive emotion bias

MICHAEL F. VERDE, LAURA K. STONE, HANNAH S. HATCH, AND SIMONE SCHNALL
University of Plymouth, Plymouth, England

Does familiarity arise from direct access to memory representations (a mnemonic account) or from inferences and diagnostic cues (an attributional account)? These theoretically distinct explanations can be difficult to distinguish in practice, as is shown by the *positivity effect*, the increase in feelings of familiarity that accompanies positive emotion. Experiment 1 manipulated mnemonic and attributional sources of positivity via word valence and physical expressions of emotion, respectively. Both sources influenced the tendency to call items *old*, but receiver-operating characteristic analysis revealed a change in accuracy only with the mnemonic source. To further contrast the mnemonic and attributional accounts, Experiment 2 varied the ratio of positive to neutral words. A higher proportion of positive words exaggerated the pattern of increased *old* judgments and decreased accuracy for positive words, relative to neutral ones, consistent with the mnemonic account but inconsistent with the attributional account.

Knowledge that something has been previously encountered often takes the form of a simple sense of familiarity, a feeling of *oldness* that is immediate but lacking in detail or context. The nature and source of familiarity have been a focus of debate among memory theorists going back to Wundt (Feingold, 1915), and in the current literature, this takes the form of two very different explanatory approaches. Mnemonic explanations suggest that familiarity arises from direct access to memory images of the stimulus probe. This is the traditional approach taken by mechanistic models of recognition in which the characteristics of the probe are compared with those of memory images in order to calculate an index of similarity or goodness of match (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988). However, the growing appreciation for the role of metacognition in memory has given rise to attributional explanations that suggest that familiarity is not a direct product of memory retrieval, because direct access to memory representations either is limited or, at best, provides ambiguous information. Familiarity is a feeling that arises when we infer, on the basis of various clues, that a past encounter must have occurred. An example of a diagnostic clue is fluency: The ability to perceive and process something quickly and easily is evidence that the cognitive system has processed the stimulus in the past. It is the attribution of processing fluency to prior exposure that gives rise to familiarity (Jacoby & Dallas, 1981; Whittlesea, 1993, 1997).

Although the mnemonic and attributional accounts differ fundamentally at a theoretical level, they can be dif-

ficult to distinguish in practice. For example, prior exposure may prime the perceptual pathways activated by the stimulus, resulting in more fluent processing, but prior exposure also creates memory traces of this past experience. Does familiarity arise from an inference about the meaning of fluent processing or from direct contact with an image stored in memory? Signal detection analysis offers, in principle, a way to distinguish between the two accounts. In the signal detection model, the tendency when making a recognition judgment to choose one response over another (i.e., to respond *old* rather than *new*) indicates a change in response bias, not accuracy. A change in the quality of information retrieved from memory, on the other hand, may affect the ability to distinguish old from new probes, resulting in a change in accuracy. We hypothesize that attributional factors influence response tendency only, whereas mnemonic factors influence the quality of memory. Therefore, it should be possible to determine whether a change in feelings of familiarity is due to attributional or mnemonic factors by looking at corresponding changes in bias and accuracy.

In this study, we consider the case of positive emotion bias, or the *positivity effect*, the finding that positive affect inflates feelings of familiarity. The most common account of this phenomenon draws on attribution theory to suggest that affect and fluency are intertwined, so that positive affect, like fluency, provides information that can be used to infer past experience. We argue that an alternative account based on mnemonic processes is equally viable in certain circumstances and may offer a better explana-

M. F. Verde, michael.verde@plymouth.ac.uk

tion for some findings previously interpreted in terms of the attributional account. In Experiment 1, we juxtaposed the attributional and mnemonic factors within a single experiment to show that the signal detection model is able to distinguish between them. In Experiment 2, we showed that the effect of varying the proportion of positive items within a list uniquely supports the predictions of the mnemonic account. Although the focus here is on the positivity effect, this investigation is meant to serve as a case study of the confusion that exists in the wider recognition literature between attributional and mnemonic accounts of familiarity.

Positive Emotion and the Attributional Account

Researchers have long been aware of the close relationship between positive affect and previous experience. One manifestation of this relationship is the *mere exposure* effect, the finding that repeated unreinforced exposure to an item increases positive feelings toward it (Zajonc, 1968). This effect has been observed over a wide range of materials and a variety of affect-related judgments, such as liking, pleasantness, and preference (for reviews, see Bornstein, 1989; Clore & Schnall, 2005). Although some have suggested that repeated exposure simply involves affectively nonspecific activation that leads to higher ratings on any stimulus-related dimension (Mandler, Nakamura, & Van Zandt, 1987; but see Seamon, McKenna, & Binder, 1998), the larger body of evidence supports the idea that its influence is genuinely affective in nature. Mere exposure is associated with subtle contractions of the zygomatic muscle responsible for smiling (Harmon-Jones & Allen, 2001) and with positive self-reported mood (Monahan, Murphy, & Zajonc, 2000).

One consequence of experience with an item is that it facilitates subsequent processing. It has been shown that fluency by itself can be a source of positive affect. Reber, Winkielman, and Schwarz (1998) manipulated the perceptual fluency of geometric figures by varying factors such as figure-ground contrast and presentation duration. Figures that were more easily perceived due to higher contrast or longer presentation were rated as more pretty and also less ugly—the common direction of both judgments indicating that the bias was specifically positive in nature. Winkielman and Cacioppo (2001) found that flashing contour-congruent primes just before the presentation of line drawings, as well as increasing the presentation duration of the drawings, not only increased feelings of liking for the drawings, but also resulted in greater EMG-recorded activity of smile-related facial muscles. A functional view of affect is that it serves to provide fast, usually reliable information that informs all kinds of evaluative judgments—that is, considerations of whether a judgment object or situation is good or bad (Schwarz & Clore, 1983, 1988). Winkielman, Schwarz, Fazendeiro, and Reber (2003) have suggested that the hedonic marking of fluency serves such a function. Fluency indicates familiarity, and familiar things are less likely to be dangerous than unfamiliar things. Fluency is associated with prototypicality, which is indicative of biological

and genetic health. Fluency marks progress in perceptual and cognitive processing. In sum, fluency is associated with a variety of positive stimulus characteristics, and the positive affect tied to fluency provides immediate feedback about their presence.

The link between fluency and positive affect is the basis for an attributional account of the positivity effect in recognition. If fluency is diagnostic of a past encounter and is hedonically marked, positivity should itself suggest prior experience. Consistent with this prediction, a number of studies have shown that positive affect inflates feelings of familiarity even in the absence of true memory. Monin (2003) misled subjects to believe that they had seen a subliminal word list and then tested their memory for the list with a novel set of words. Positive words were more likely to be falsely recognized than were neutral words. Claypool, Hall, Mackie, and Garcia-Marques (2008) had subjects read a happy or a neutral article. Some were then asked whether they felt happy or sad. Finally, all were told that they had earlier seen a subliminal list and were given a sham recognition test. False recognition was higher in the happy mood condition, but only for subjects not asked to introspect about their mood. Claypool et al. proposed an attributional explanation: The subjects in a happy mood misattributed their affect-related feelings to familiarity but discounted the feelings when introspective questions made them aware of their source.

Of course, it is important to show that the positivity effect occurs during actual tests of memory. During recognition, Phaf and Rotteveel (2005) induced subjects to contract their facial muscles in ways consistent with either smiling or frowning. The subjects in the smiling condition were more likely to respond *old* to both studied and new items. In another experiment, Phaf and Rotteveel briefly displayed primes just before the presentation of each recognition probe. The probes were neutral words, and the primes were novel positive or negative words. Priming with positive words led to an increase in *old* responses. In a similar vein, Garcia-Marques, Mackie, Claypool, and Garcia-Marques (2004) preceded recognition probes with subliminal (12-msec) primes that were either empty circles or circles containing happy faces. Priming with happy faces led to an increase in false alarms but a nonsignificant decrease in hits.

Other studies have investigated the positivity effect by manipulating the affective valence of studied items and memory probes. Positive words have been observed to produce more hits and false alarms (Kapucu, Rotello, Ready, & Seidl, 2008) or only more false alarms (Cook, Hicks, & Marsh, 2007; Dougal & Rotello, 2007) than neutral words. Garcia-Marques et al. (2004) found that smiling faces produced more false alarms than did faces with neutral expressions. Monin (2003) found that more attractive faces produced more hits and false alarms than did less attractive faces, although a second study by Corneille, Monin, and Pleyers (2005) observed only the increase in false alarms. On the whole, the evidence supports the notion that positive-valence items tend to feel more familiar. This can be interpreted as meaning

that the positive affect produced by the items is misattributed to familiarity. However, an alternative interpretation is that positive items are more familiar because of their similarity or relatedness to the many other positive items that have been recently encountered and stored in memory.

The Mnemonic Account: Semantic Relatedness

In theories that assume direct access to memory representations, an item feels familiar because it matches a representation of a previous encounter. Of course, one would rarely expect such a match to be exact, because the surrounding context, the perceptual conditions, even integral parts of an item can change over time. Thus, it would make functional sense for familiarity to vary continuously with the degree of match. This allows familiarity to be informative, given the fluctuating environment, but it also means that familiarity is fundamentally ambiguous, because even novel items will bear some resemblance to things encountered in the past. Signal detection theory provides a means with which to model this ambiguity. In Figure 1 (top panel), *memory evidence* is a continuous dimension that represents some aggregate of the matches

between a recognition probe and the memory images of recently studied items (an assumption common to many formal models of recognition; e.g., Clark & Gronlund, 1996). Old items are naturally more familiar, on average, than new items, and the distributions of evidence are assumed to be normal but not necessarily of equal variance. Recognition judgments are made by choosing a criterion level of evidence, C . Any probe whose evidence exceeds this criterion is judged to be *old*. The proportions of the old item and new item distributions falling to the right of the criterion are equivalent to recognition hit and false alarm rates, respectively. Note that in this model, the terms *familiarity* and *bias* take on restricted meanings. *Familiarity* is often used (e.g., by attribution theories) to refer to a subjective feeling; here, *familiarity* is synonymous with the amount of memory evidence. *Bias* is often used in the literature to describe any change in hits and false alarms; here, *bias* refers to the location of the recognition criterion.

The signal detection model makes a distinction between accuracy and bias. Accuracy represents the ability to discriminate old from new items and is defined by the manner in which their evidence distributions overlap. Bias represents the level of evidence required to judge a probe *old* and is defined by the location of the criterion on the evidence axis.¹ Any manipulation that leads to a change in recognition performance can be characterized by how it affects accuracy and bias, which are independent of one another. Consider two conditions, A and B. The top panel of Figure 1 shows the evidence distributions of probes in Condition A with recognition criterion C . Imagine that, in Condition B, the distributions are identical but a more liberal criterion, C_L , is chosen. Shifting the criterion in this way affects bias and increases hits and false alarms, relative to Condition A. It does not affect accuracy, because shifting the location of the criterion has no effect on the underlying evidence distributions. Next, imagine that Condition B differs from Condition A in that both old and new items have increased in average familiarity, shifting their distributions to the right while maintaining the same criterion, C , as shown in the lower two panels. Like the previous example, this will increase hits and false alarms, relative to Condition A. It may or may not result in a change of accuracy. As long as the manner in which the distributions overlap remains unchanged, the ability to discriminate old from new is unaffected. This would be the case if the means of the distributions shifted rightward without affecting their variances and relative distance, as shown in the middle panel. On the other hand, accuracy would decline if the distance between the means decreased or if their variances increased, as shown in the bottom panel; both changes would lead to greater distributional overlap. In other words, a change in the evidence distributions for the items in Condition B can lead to changes in hits, false alarms, and accuracy, relative to Condition A, but the nature of these changes will depend on how a manipulation specifically affects memory evidence. The crucial point made by Figure 1 is that although both changes in response bias and changes in memory

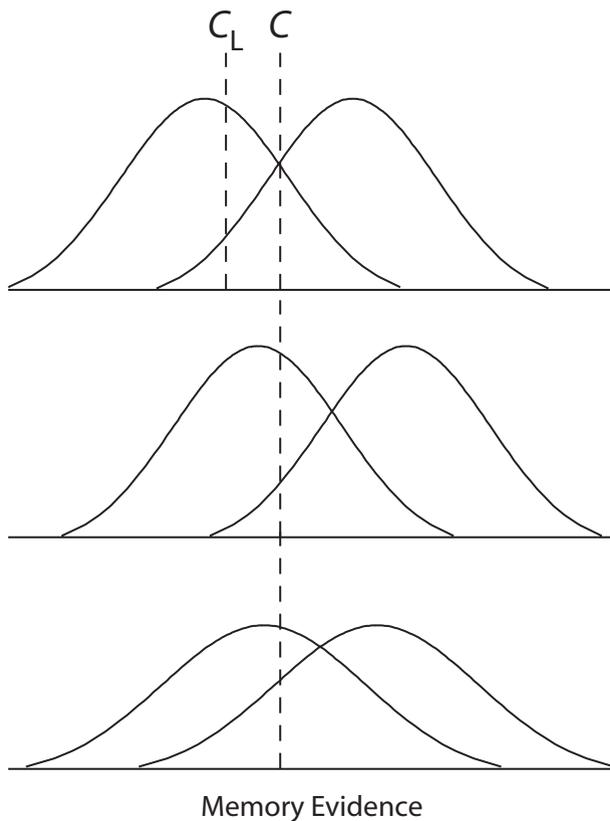


Figure 1. Hypothetical old- and new-item evidence distributions with recognition criterion C . **Top:** Criterion C_L reflects a liberal response bias with no change in memory evidence. **Middle:** Evidence distributions shift with no change in variance or relative distance. **Bottom:** Evidence distributions shift with a change in variance and relative distance, leading to decreased accuracy.

quality can lead to increases in hits and false alarms, only the latter will produce changes in accuracy.

The positivity effect is often examined by comparing recognition of positive- and neutral-valence words. If positive words have a greater degree of interitem categorical or semantic overlap, how will this affect the quality of memory? There are two ways to think about this. If memory evidence is a graded function of the match between a probe and the contents of memory, the familiarity of positive probes should be inflated, relative to neutral probes, leading to an upward shift of the evidence distributions, as in the lower two panels of Figure 1. The increase in familiarity will be a function of both the degree of categorical or semantic overlap and the number of overlapping images in memory. Increased semantic overlap may also lead to a loss of interitem distinctiveness and discriminability, a situation described by the increasing overlap of distributions in the bottom panel of Figure 1. The results of investigations of the role that categorical similarity plays in recognition have been consistent with both of these predictions: Increasing the number of studied exemplars from a category typically has led to an increase in false alarms (and to a lesser extent, hits), as well as a drop in accuracy for probes belonging to the category (Criss & Shiffrin, 2004; Dewhurst & Anderson, 1999; Shiffrin, Huber, & Marinelli, 1995; see also Verde, 2004).

The mnemonic account of the positivity effect predicts that semantic relatedness can lead to changes in accuracy, and this is one way to distinguish it from the attributional account, which arguably predicts no such change. Consider the ways that attribution has been described by various theorists. Monin (2003) suggested that liking is experienced as a "warm glow" that people use as a heuristic to infer familiarity. Garcia-Marques et al. (2004) likewise described positivity as a cue or a signal of familiarity. According to Phaf and Rotteveel (2005), positivity and familiarity "activate the same behavior system leading to less critical and more heuristic decision tendencies" (p. 310). In signal detection terms, all of these descriptions might be interpreted to mean that exposure to positive affective cues makes one more willing to indicate that something has been encountered before; in other words, it encourages a more liberal response bias.

Isolating attributional from mnemonic factors requires that emotion be manipulated independently of the characteristics of the recognition probe. Phaf and Rotteveel (2005) did this by manipulating facial expression and also the valence of word primes shown briefly before the test probes. In both cases, they observed a change in response bias but not in accuracy, which was measured using the two-high threshold index, P_r . On the other hand, Garcia-Marques et al. (2004) primed test probes with either happy faces or empty circles and found that the faces increased false alarms but not hits, suggesting a drop in accuracy. Before drawing conclusions from these apparently conflicting results, however, it is important to consider problems with the measurement of accuracy.

If accuracy is defined as the overlap of evidence distributions, a reliable measure of accuracy requires some

knowledge about the shape of these distributions. One way to do this is with receiver-operating characteristics (ROCs; Macmillan & Creelman, 2005): Recognition confidence ratings produce hit and false alarm rates at multiple points along the evidence axis, allowing one to estimate the relative height of the distributions at each point. A typical recognition test, however, asks for a binary (*old* vs. *new*) judgment, which provides hit and false alarm rates at only a single point on the evidence axis. Because the shape of the distributions on either side of this point are unknown, commonly used indexes of accuracy for single-point data (such as d' or P_r) rely on an assumption that old and new item distributions are normal with equal variance. Unfortunately, the latter assumption does not, in general, hold true for recognition memory (Mickes, Wixted, & Wais, 2007; Ratcliff, Sheu, & Gronlund, 1992). When variances are unequal, single-point indexes no longer provide a reliable measure of accuracy independently of bias (Verde, Macmillan, & Rotello, 2006). The crucial point, as Verde and Rotello (2003) have shown with regard to other illusions of recognition, is that binary recognition judgments like those used by the studies mentioned above may not provide enough information to reliably differentiate bias from accuracy effects.

A number of studies have manipulated emotion via the characteristics of the recognition probe itself. Kapucu et al. (2008) observed more hits and false alarms but no difference in accuracy between positive and neutral words. Cook et al. (2007), Corneille et al. (2005), and Garcia-Marques et al. (2004) found significant effects only in false alarms, suggesting a drop in accuracy. Grider and Malmberg (2008) observed an increase in accuracy for positive words. Both Kapucu et al. and Grider and Malmberg used ROC-based measures of accuracy, so it is hard to attribute the inconsistent results to the measurement problems noted above. For the moment, however, it is sufficient to note that a change in memory evidence may or may not affect accuracy. Attribution should not affect accuracy, and this is the key to our attempt in Experiment 1 to distinguish between the two types of explanations.

EXPERIMENT 1

The inflated feeling of familiarity that accompanies positive affect has most commonly been interpreted as evidence for attributional processes. However, a survey of the literature shows that such a view requires closer scrutiny, for two reasons. First, when the recognition probes themselves are the source of affective information, an explanation based on normal memory-related processes provides an adequate alternative to the attributional account. Second, although a simple and reasonable characterization of attribution is as a process that determines criterion placement, previous results provide mixed support for this. If attribution does not simply influence bias, the following question remains: How should it be characterized? One difficulty of turning to the existing literature to resolve this question is that most previous studies either do not distinguish between bias and accuracy or rely

on methods that are potentially unreliable when changes in both bias and accuracy are present in recognition data (Verde et al., 2006; Verde & Rotello, 2003). In the present study, we used ROC analysis to more reliably examine how positivity influences accuracy.

A second, equally important aspect of our methodology is that we juxtaposed attributional and mnemonic factors within a single experiment. Investigations of the positivity effect have employed a wide variety of materials and procedures, and this inevitably leads to concerns about comparing results between studies. In Experiment 1, the affective characteristics of the memory probes were manipulated by using positive- and neutral-valence words. If the increased rate of *old* judgments for positive words is due, at least in part, to the effect of interitem semantic relatedness on the quality of memory, positive words may be recognized with less accuracy than are neutral words. Positive affect was also manipulated independently of the memory probes by instructing subjects to produce expressive behaviors consistent with either happiness or sadness. The physical expressions of emotion are known to produce feelings consistent with those emotions (e.g., Schnall & Laird, 2003). Furthermore, affective expressions can influence cognitive information processing (Centerbar, Schnall, Clore, & Garvin, 2008; Schnall & Laird, 2007). If positive affective cues lead to a sense of fluency that is easily misattributed to familiarity, the smiling condition should produce higher rates of *old* judgments but no change in accuracy, relative to the frowning condition.

Method

Subjects. Fifty-four undergraduate students at the University of Plymouth participated for course credit. They were assigned randomly in equal numbers to the two physical expression conditions.

Materials and Design. Stimuli were taken from Bradley and Lang's (1999) *Affective Norms for English Words*, which provides ratings of valence and arousal.² Word frequency counts were taken from Kučera and Francis (1967). The pool of critical items consisted of 60 positive words (e.g., *friend*, *kiss*; means: valence = 7.97, arousal = 5.86, word frequency = 42 per million) and 60 neutral words (e.g., *pencil*, *ankle*; means: valence = 5.08, arousal = 3.92, word frequency = 41 per million). Additional neutral words were used as filler items. The study list consisted of 30 positive and 30 neutral words, with three neutral filler items placed at the beginning and end of the list. The test list consisted of all of the studied words, as well as 30 new positive and 30 new neutral words, with three filler items placed at the beginning of the list. The assignment of words to studied or new status and the order of words within lists were randomized uniquely for each subject.

Procedure. The subjects were seated at individual computers, with up to 4 subjects in a room per session. Following the methods of Schnall and Laird (2007) and Centerbar et al. (2008), the subjects were told at the outset that the experiment would investigate the effect of anaerobic exercise on word recognition and memory, and they read a cover story describing legitimate research on the link between anaerobic exercise and cognition. The study phase followed, in which the subjects were instructed to remember the list of words for the coming memory test. Each word was shown on the computer screen for 2,000 msec, followed by a 500-msec blank. The recognition test followed immediately. During each trial, a single word was presented on the screen, and the subjects indicated their confidence that the word was studied or new, using a 6-point confidence scale (1 = *very sure new* . . . 6 = *very sure old*). A 2,000-msec blank interval followed the keyboard response.

Following four practice test trials, the instructor paused the test and told the subjects that, for the remainder of the test, they would also be making certain muscle contractions that were thought to be helpful for word recognition. Unbeknownst to the subjects, the nature of these contractions was varied between groups and resulted in a happy facial expression and posture for half of the subjects and a sad facial expression and posture for the other half. Using anatomically detailed charts showing the musculature of the face and back (Centerbar et al., 2008; Schnall & Laird, 2007), the experimenter instructed the subjects in the *happy expression* condition to contract the zygomaticus and risorius muscles by turning up the corners of their mouth while leaving their mouth open a little, and to activate the trapezius and latissimus dorsi muscles by sitting up straight with their knees bent at a right angle. In the *sad expression* condition, the subjects were instructed to contract the triangularis and platysma muscles by drawing the corners of their mouth down and back and letting their head hang down a little, and to activate the trapezius and latissimus dorsi muscles by drooping their shoulders and letting their body go relatively limp while dropping their rib cage. Each subject demonstrated these postures to the experimenter's satisfaction and then continued the recognition test, maintaining the postures. Frequent breaks were interspersed throughout the test to allow the subjects to relax their muscles for a short period. The experimenter monitored the subjects to ensure that they maintained the posture during test trials, reminding them to do so when necessary. To be consistent with the cover story, at no point was there mention of an association between the physical expressions and emotional feelings or affect.

Results and Discussion

Recognition performance is reported in Table 1. The data were first examined by collapsing confidence ratings into *old* (confidence 1–3) and *new* (confidence 4–6) response categories to determine recognition hit and false alarm rates. These were submitted to a $2 \times 2 \times 2$ ANOVA with physical expression (happy vs. sad) as a between-subjects factor and word type (studied vs. new) and word valence (positive vs. neutral) as within-subjects factors. There were statistical main effects of expression [$F(1,52) = 7.67, p < .01, \eta_p^2 = .13$], word type [$F(1,52) = 253.71, p < .001, \eta_p^2 = .83$], and word valence [$F(1,52) = 79.84, p < .001, \eta_p^2 = .61$], as well as a statistical interaction between word type and valence [$F(1,52) = 7.90, p < .01, \eta_p^2 = .13$]. There were no other interactions. The main effects of both emotion-related factors were in the expected direction. The subjects in the happy expression condition treated all the words as more familiar, producing more hits and false alarms, as compared with those in the sad expression condition. In both expression conditions, the

Table 1
Experiment 1: Hit and False Alarm (FA) Rates,
Accuracy (d_a), and Bias (c_a)

Word Valence	Hit		FA		d_a		c_a	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Happy Expression								
Positive	.73	.02	.41	.02	0.84	0.08	-0.21	0.06
Neutral	.65	.03	.27	.03	0.97	0.10	0.11	0.06
Sad Expression								
Positive	.67	.02	.34	.02	0.86	0.09	-0.02	0.04
Neutral	.59	.03	.20	.02	1.00	0.12	0.30	0.04

Note—Physical expression was manipulated between groups.

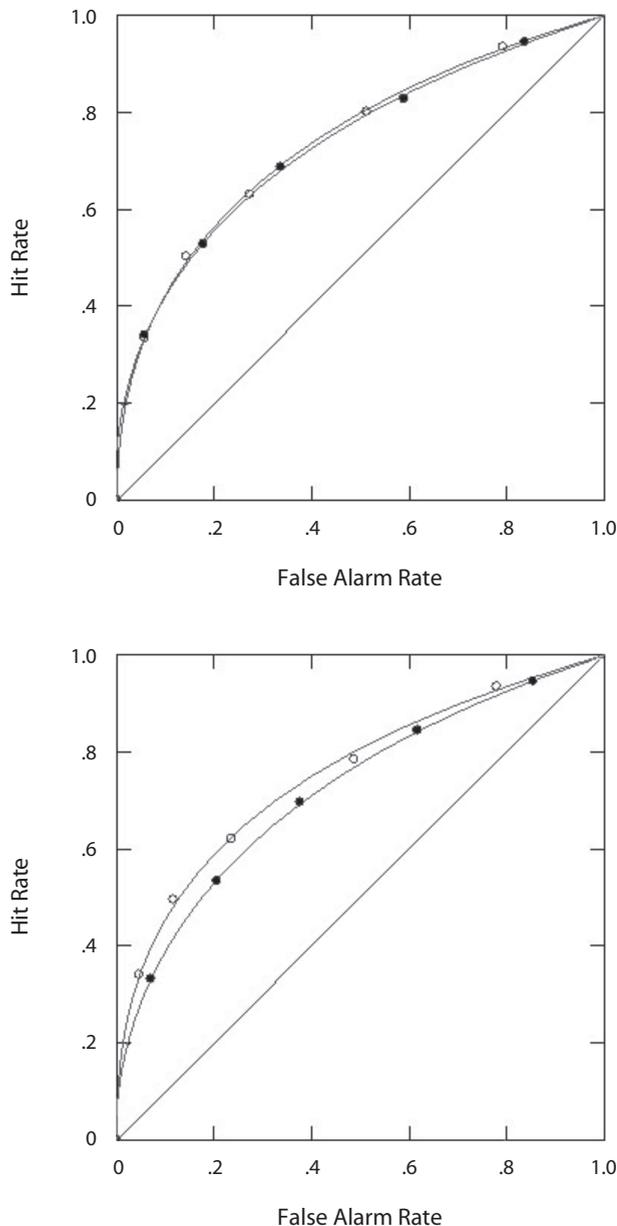


Figure 2. Experiment 1: Aggregate receiver-operating characteristic curves with best-fitting normal model superimposed. Top: Comparison of happy (dark circles) and sad (light circles) expression conditions. Bottom: Comparison of positive-valence (dark circles) and neutral-valence (light circles) word conditions.

subjects treated positive-valence words as more familiar, producing more hits and false alarms, as compared with neutral words. The interaction between word type and valence was due to a larger increase in false alarms than in hits for positive words.

From the confidence ratings, individual ROC curves were constructed for each subject and condition, and maximum-likelihood estimation was used to derive measures of accuracy (d_a) and bias (c_a ; analogous to d' and c , respectively, but not tied to assumptions about distribution

variance; Macmillan & Creelman, 2005). The values of d_a were submitted to a 2×2 ANOVA with physical expression as a between-subjects factor and word valence as a within-subjects factor. There was a statistical main effect of word valence, with lower accuracy for positive words ($d_a = 0.85$) than for neutral words ($d_a = 0.98$) [$F(1,52) = 4.99, p < .05, \eta_p^2 = .09$]. Neither the main effect of expression nor the interaction between expression and valence was statistically significant. Similar analysis of c_a values revealed statistical main effects of word valence [$F(1,52) = 75.08, p < .001, \eta_p^2 = .59$] and physical expression [$F(1,52) = 7.79, p < .01, \eta_p^2 = .13$], consistent with a more liberal bias for positive words and for the happy expression condition. Word valence and expression did not interact. It is important to keep in mind, however, that c_a indexes the location of the criterion relative to the evidence distribution means.³ Thus, the main effect of valence should be interpreted with caution: Although it is consistent with a liberal bias for positive words, it might also be tied to changes in the quality of memory associated with lower positive word accuracy.

The two emotion-related manipulations, physical expression and word valence, produced positivity effects, consistent with previous findings. Both factors had similar effects on hit rates, but word valence had a much larger effect on false alarm rates. The similarity with respect to hit rates makes it hard to argue that a difference in magnitude between the two manipulations could be responsible for different effects on accuracy. That the two factors did have qualitatively different influences on recognition was borne out in the analysis of the ROC data, which showed that physical expression only influenced bias, whereas word valence influenced recognition accuracy. This can be seen clearly in the ROC curves in Figure 2, which are based on aggregate data collapsed so as to compare only facial expression (top panel) or only word valence (bottom panel). Pairs of hits and false alarms that fall along a common curve represent different criterion locations but equivalent accuracy. Happy and sad expression conditions differ in bias but clearly fall on the same ROC curve. Positive- and neutral-valence word conditions fall on different curves. The middle diagonal, along which hit and false alarm rates are equal, represents chance performance. The closer proximity of the positive-valence curve to the middle diagonal indicates lower accuracy for that condition.

EXPERIMENT 2

The larger message from Experiment 1 is that the influence of positive affect on familiarity is not comfortably accommodated by a single theoretical account. Using the label *positivity effect* to describe all such findings implies that they derive from a single source that influences recognition performance in a consistent manner. On the contrary, attributional and mnemonic processes seem to contribute separately and in qualitatively different ways to the positivity effect. Facial expression and body posture involve affective cues that are unrelated to the memory task. However, people have presumably learned through

prior experience the contingency between positive affect and memory. They treat positive feelings as a cue for familiarity, and this makes them more willing to call something *old*. Of course, it seems reasonable to assume that this response bias should arise whether the source of the affect is incidental to the memory probe or the probe itself. The decline in accuracy observed with positive items does not rule out the presence of a bias, but it does indicate a change in underlying memory evidence, which rules out attribution as the sole cause of the positivity effect.

We are not the first to note that semantic relatedness may be an issue with emotional materials. Other recognition studies, interested primarily in isolating the effect of emotion, have attempted to equate the interitem relatedness of positive and neutral words in various ways (Dougal & Rotello, 2007; Grider & Malmberg, 2008). The goal of the present study was not to isolate the pure contribution of emotion but, rather, to argue that semantic relatedness effects can be misconstrued as attributional effects. According to the mnemonic account, categorical or semantic overlap increases the match between a probe and an item in memory. However, just as important as the degree of overlap is the number of overlapping items. Because familiarity is based on an aggregate of matches, the effect of relatedness on familiarity is a function of the number or proportion of related items in the study list (Arndt & Hirshman, 1998; Criss & Shiffrin, 2004; Shiffrin et al., 1995). One way to support the mnemonic account would be to demonstrate that, consistent with semantic relatedness effects observed with other materials, the size of the positive word effect increases as the proportion of positive words in the study list increases. On the other hand, such a finding would be inconsistent with an attributional account.

In Experiment 2, the size of the study and test lists was held constant, and the proportion of positive words was varied between groups. The proportion of positive words was 30% in the *low-positive-ratio* condition and 57% in the *high-positive-ratio* condition. The mnemonic account predicts that as the proportion of positive words increases, the degree of aggregate match increases, leading to greater feelings of familiarity, as well as to a loss of distinctiveness for positive words. In operational terms, this should result in a greater tendency to call positive words *old*, relative to neutral words, and also diminished accuracy for positive than for neutral words. The mnemonic account predicts that both trends will be more pronounced in the *high-positive-ratio* condition. According to the attributional account, processing a positive word feels fluent; this property of fluency is intrinsic to the word and would not be expected to depend on the properties of other words. However, a recent study by Westerman (2008) suggested an alternative possibility. The Jacoby–Whitehouse illusion (Jacoby & Whitehouse, 1989), in which priming the recognition probe with an identical word elevates the rate of *old* responses, is widely interpreted as an example of misattributed fluency. Westerman observed that the size of the fluency effect was largest when identical primes made up only a small proportion of trials. This suggests that the feeling of enhanced fluency is greatest when it is relatively

uncommon and, therefore, surprising. Thus, two predictions seem to be compatible with an attributional account: Either positive ratio should be unrelated to enhanced positive word familiarity, or the positive word effect should be largest in the *low-positive-ratio* condition.

Several previous studies have provided some support for the predictions of the mnemonic account. Dougal and Rotello (2007), using a similar design, compared a *high-emotion-ratio* condition (33% positive, 33% negative, and 33% neutral words) with a *low-emotion-ratio* condition (25% positive, 25% negative, and 33% neutral words). They found that recognition was less accurate for positive than for neutral words only in the *high-emotion-ratio* condition, consistent with the mnemonic account. However, they observed no difference between positive and neutral words in number of *old* responses. Although, at first glance, this seems inconsistent with the basic positivity effect, in fact the *high-emotion-ratio* condition produced higher false alarm rates for positive than for neutral words, but the trend may have been masked by lower hit rates for positive words (hits and false alarms were not separately analyzed). Two other studies did not vary the proportion of positive words but did attempt to control semantic overlap. Kapucu et al. (2008) used only 25% positive words and found no difference in accuracy between positive and neutral words. Grider and Malmberg (2008) controlled for semantic relatedness, using latent semantic analysis, and found recognition to be more, not less, accurate for positive than for neutral words. All of these studies offer suggestive evidence in support of the prediction that increased feelings of familiarity and decreased accuracy for positive words are driven by a high degree of semantic relatedness; when relatedness is controlled, these trends weaken or disappear. However, none of these studies combined the manipulation of positive word ratio with a full analysis of hit and false alarm rates necessary for our test of the attributional and mnemonic accounts.

Method

Subjects. Fifty undergraduate students at the University of Plymouth participated for course credit or payment. They were assigned randomly in equal numbers to the two positive ratio conditions.

Materials and Design. Words were taken from the Bradley and Lang (1999) norms to create a pool of 96 positive words (e.g., *friend*, *kiss*; means: valence = 7.94, arousal = 5.88, word frequency = 57 per million) and 120 neutral words (e.g., *pencil*, *ankle*; means: valence = 5.08, arousal = 4.04, word frequency = 42 per million). Although a larger pool was used, the word characteristics were nearly identical to those of the positive and neutral words used in Experiment 1. Additional neutral words were used as filler items. The study list consisted of 84 words, with three neutral filler items placed at the beginning and end of the list. The test list consisted of all studied words as well as an equal number of new words, with three filler items placed at the beginning of the list. The ratio of positive to neutral words in these lists was manipulated between subjects. In the *low-positive* condition, 30% of the study and test words were positive, whereas in the *high-positive* condition, 57% of the study and test words were positive. The assignment of words to studied or new status and the order of words within lists were randomized uniquely for each subject.

Procedure. The subjects were seated at individual computers, with up to 5 subjects in a room per session. They were told at the

Table 2
Experiment 2: Hit and False Alarm (FA) Rates, Accuracy (d_a), and Bias (c_a)

Word Valence	Hit		FA		d_a		c_a	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
High Positive Ratio (57%)								
Positive	.71	.02	.39	.03	0.93	0.09	-0.13	0.06
Neutral	.67	.03	.22	.03	1.24	0.10	0.17	0.07
Low Positive Ratio (30%)								
Positive	.71	.03	.33	.03	1.01	0.09	-0.05	0.06
Neutral	.69	.03	.29	.03	1.14	0.11	0.04	0.05

Note—Positive word ratio was manipulated between groups.

outset that they would see a list of words that they should remember for an upcoming memory test. Presentation of the study and test lists was identical to that in Experiment 1 (absent the manipulation of physical expression).

Results and Discussion

Recognition performance is reported in Table 2. The data were first examined by collapsing confidence ratings into *old* and *new* response categories to determine recognition hit and false alarm rates, as shown in Table 2. These were submitted to a 2 × 2 × 2 ANOVA with positive ratio (low vs. high) as a between-subjects factor and word type (studied vs. new) and word valence (positive vs. neutral) as within-subjects factors. There was no main effect of positive ratio, but there were statistical main effects of word type [$F(1,48) = 269.87, p < .001, \eta_p^2 = .85$] and word valence [$F(1,48) = 22.33, p < .001, \eta_p^2 = .32$]. The latter was the result of a positivity effect, with positive words called *old* more often than neutral words. Word type did not interact with positive ratio, but there were statistical interactions between word valence and positive ratio [$F(1,48) = 7.40, p < .01, \eta_p^2 = .13$] and word valence and word type [$F(1,48) = 13.36, p < .01, \eta_p^2 = .22$] and a three-way interaction between word valence, word type, and positive ratio [$F(1,48) = 6.62, p < .05, \eta_p^2 = .12$]. This complex set of interactions was clarified by planned comparisons that looked for the presence of a positivity effect by examining valence-induced changes in hit and false alarm rates within each ratio condition. With a low positive ratio, there were slightly more hits and false alarms to positive words, but neither of these trends was significant. With a high positive ratio, there were statistically more false alarms to positive words [$t(24) = 6.45, p < .001, d = 1.06$] and slightly more hits, although this was not statistically significant. To summarize, only when the proportion of positive words was relatively high was there a greater tendency, located mainly in false alarms, to call positive words *old*. When the proportion of positive words was relatively low, there was only a small, nonsignificant trend in this direction. This pattern is consistent with the mnemonic account, which predicts that positive words will feel increasingly familiar as semantic overlap increases.

Accuracy and bias measures were derived from confidence rating ROCs constructed for each subject and condition. Accuracy (d_a) values were submitted to a 2 × 2 ANOVA with positive word ratio as a between-subjects

factor and word valence as a within-subjects factor. Although there was no main effect of positive ratio, there was a statistical main effect of word valence [$F(1,48) = 22.89, p < .001, \eta_p^2 = .32$] and a statistical interaction between positive ratio and word valence [$F(1,48) = 4.22, p < .05, \eta_p^2 = .08$]. This interaction, shown in Figure 3, reveals that the disparity in accuracy between positive and neutral words was larger in the high-positive condition. Planned comparisons confirmed that accuracy was statistically lower for positive words in the high-ratio condition [$t(24) = 7.76, p < .001, d = 0.66$], but this was not true for the difference in the low-ratio condition. The interaction between positive ratio and word valence is consistent with the mnemonic account’s prediction that the distinctiveness of positive words will decline, relative to neutral words, as the proportion of positive words increases.

Bias values (c_a) were analyzed in a similar fashion. There was a statistical main effect of word valence [$F(1,48) = 23.83, p < .001, \eta_p^2 = .33$]. Although there was no main effect of positive ratio, there was an interaction between ratio and word valence [$F(1,48) = 7.11, p = .01, \eta_p^2 = .13$]. Planned comparisons revealed that c_a for positive words was statistically lower in the high-positive condition [$t(24) = 4.72, p < .001, d = 4.60$], but only marginally so in the low-positive condition [$t(24) = 1.85, p < .10, d = 1.63$]. The results are consistent with a liberal bias for positive words that was significant only when the proportion of positive words was high. As was discussed earlier, this interpretation must be qualified by the fact that a high ratio of positive words was associated with lower accuracy.

The claim that positive words possess a greater degree of semantic relatedness to one another than do neutral words was supported consistently in both experiments by the fact that positive words were associated with higher hit and false alarm rates and lower accuracy. The interaction shown in Figure 3, however, raises the issue of how inter-item relatedness might change between ratio conditions. The trend of declining accuracy for positive words as the

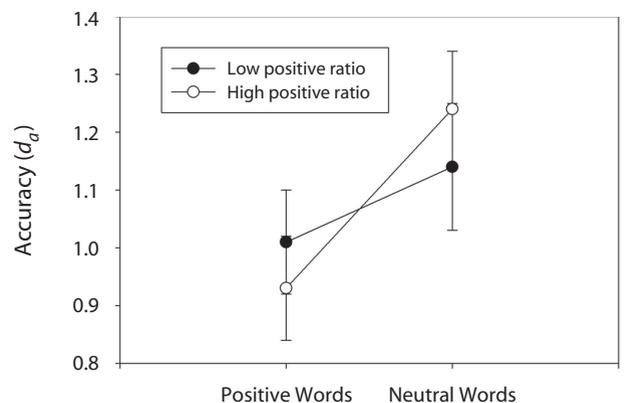


Figure 3. Experiment 2: Mean accuracy (with SE) by word valence within each ratio group. The interaction shows the increasing disparity in accuracy between positive and neutral words as the proportion of positive words increases.

positive ratio moved from low to high was expected, but the simultaneous trend of increasing accuracy for neutral words was not. A mundane explanation for this result is that the high-ratio group was simply more accurate in general, elevating accuracy for both valence conditions, relative to the low-ratio group. This could have occurred fortuitously or for unknown reasons related to altered list composition. However, it may also suggest that our discussion of semantic relatedness effects has been too simplistic.

Work in categorization suggests that the similarity or semantic overlap between items is not immutable but depends on context, modified by our active concepts and selective focus on various feature dimensions (Murphy & Medin, 1985; Nosofsky, 1986). It might be that increasing the ratio of positive words also makes subjects more conscious of the *positive* category, leading them to focus on or preferentially weight category-relevant features. This would increase the semantic overlap between positive items but would decrease their overlap with neutral items. For neutral items, the effect is to make them more distinctive. For positive items, however, there are two opposing effects. Decreased overlap with neutral items increases distinctiveness, but increased overlap with other positive items decreases distinctiveness. This hypothesis can explain two aspects of the between-groups differences in Figure 3. First, it explains the trend toward higher neutral word accuracy in the high- than in the low-positive-ratio condition. Second, it suggests that the expected trend toward lower positive word accuracy in the high- than in the low-positive-ratio condition is masked somewhat by the fact that decreased relatedness to neutral words enhances the accuracy of positive words. This explanation is not incompatible with the mnemonic account described so far. Rather, it calls for a more complete model of relatedness and distinctiveness beyond the scope of the present study. What is important to focus on is that the effect of positive ratio on the within-group differences between positive and neutral words was consistent with the basic predictions of the mnemonic account.

GENERAL DISCUSSION

How much of our experience of the past depends on direct retrieval of stored images; how much depends on inferences and intelligent guesswork? This is a recurring theme in memory research; in current work on recognition, it takes the form of two parallel explanations for feelings of familiarity, one based on mnemonic processes and the other on attributional processes. Although theoretically distinct, the two types of explanations describe effects on recognition performance that are in many ways similar. To contrast the two accounts, we chose as a case study the positivity effect, the finding that positive affective cues lead to greater feelings of familiarity. The positivity effect can be produced by happy moods and exposure to positive objects and affective cues, and it is characterized by an increased tendency to call memory probes *old*. In Experiment 1, we produced a positivity effect with two

different manipulations within the same experiment. The subjects who adopted facial expressions and postures consistent with happiness (smiling, sitting up straight) produced more recognition hits and false alarms than did those who adopted expressions and postures consistent with sadness (frowning, slumping shoulders). Similarly, the subjects produced more hits and false alarms to positive words (*beauty, gift*) than to neutral words (*clock, fabric*). These findings are similar to those in other reports of the positivity effect in the literature, which typically have been interpreted as examples of how positive feelings, or the sense of fluency that accompanies positivity, can be misattributed as familiarity. The attributional account is most persuasive when the source of the positive affect is unrelated to memory. However, when the recognition probes are the source of positivity, a mnemonic account that describes familiarity as a function of the semantic relatedness among images in memory provides a viable alternative explanation.

The attributional and mnemonic accounts of familiarity are not mutually exclusive. Whittlesea (2002), for example, explicitly described recognition as a two-stage process that begins with the retrieval of information from memory but ends with an attributional process that determines how the information is subjectively experienced. However, this point is often lost, because the differential properties of these processes have not been well explored in the literature. The signal detection model provides a useful framework for doing so. Mnemonic processes lead to changes in the quality of memory evidence, which can lead to changes in accuracy. Attributional processes are essentially mechanisms that determine the criterion for recognition, influencing response bias but not accuracy. Experiment 1 provided empirical support for this characterization of the processes. A happy expression led to a more liberal bias than did a sad expression, but recognition performance for both conditions fell along a common ROC, indicating equivalent accuracy. Recognition performance for positive words, on the other hand, fell along a different, lower accuracy ROC than did neutral words.

Experiment 2 varied the ratio of positive to neutral words in a list. This provided another way to contrast the attributional and mnemonic accounts, because they make very different predictions with regard to positive word ratio. If the positivity effect arises due to the inherent fluency of a positive test probe, as is suggested by the attributional account, the characteristics of other list items should not matter. Thus, attribution theory might predict that the positive word ratio should not influence the size of the positivity effect. However, some theorists have argued that the tendency to misattribute fluency to familiarity is greater when fluency is surprising or unexpected. Westerman (2008) demonstrated this recently in a study in which priming and context manipulations were used to increase the fluency of a proportion of the recognition probes. The effect in terms of inflated rates of *old* responses was greater when fluent trials were relatively rare. For the present study, this suggests an alternative prediction of attribution theory: The positivity effect

should be greater when the proportion of positive words is low. In contrast to these two predictions, the mnemonic account predicts that, because familiarity increases with the number of semantically related items in a list, the positivity effect should increase as the proportion of positive items increases. The results of Experiment 2 supported the prediction of the mnemonic account. When the majority (57%) of the items were positive, positive items were more likely to be called *old* and also were less accurately recognized than neutral words. When positive items were in the minority (30%), these trends were present but were much smaller and not statistically significant. Although the attribution account does not seem to predict the mediating effect of the positive word ratio, should one conclude that attribution plays no role in the effect produced by word valence? On the contrary, it seems reasonable to assume that positive memory probes trigger inferential processes in the same way as other sources of positive affect. What the present study argues is that when the memory probes themselves are the source of positivity, attribution is neither the sole nor necessarily the most significant factor underlying the positivity effect.

Describing any increase in hits or false alarms as a *rise in familiarity* glosses over the distinction between bias and accuracy, which is central to understanding the factors underlying the positivity effect. Although the positivity effect is usually ascribed to the misattribution of positive feelings, a review of the literature reveals that the use of positive-valence recognition probes most often results in a rise in false alarms and a much smaller rise in hits, a pattern also observed in the present study. A large asymmetry in the effect on hits and false alarms is often indicative of a shift in evidence distributions, consistent with the mnemonic account. The implication is that the heightened feelings of familiarity for positive items may be due not to their emotional nature per se but, rather, to a factor orthogonal to emotion. An alternative explanation often put forward for the relatively larger effect on false alarms is that attribution is used primarily when there is an absence of true memory information, as would be the case with new items (Corneille et al., 2005; Garcia-Marques et al., 2004; Monin, 2003). Such an explanation has little to say about the finding in Experiment 1 that happy physical expression and positive word valence were associated with roughly equivalent increases in hits but that the latter produced a much greater increase in false alarms.

From the literature reviewed earlier, two deviations from our framework are worth mentioning. First, Garcia-Marques et al. (2004) manipulated affect independently of memory probe characteristics by subliminally priming probes with either empty circles or happy faces, finding that the latter increased false alarms only, a pattern likely consistent with a drop in accuracy. That the two prime types differed not only in valence but also in visual complexity deserves further scrutiny, but if replicable, this result might also hint at aspects of affective manipulation beyond those discussed. Second, not all studies have shown lower accuracy for positive than for neutral words. Notably, in three previous studies, ROC analysis of ac-

curacy was used while semantic overlap was controlled. Both Kapucu et al. (2008) and Dougal and Rotello (2007, Experiment 2)⁴ found no decline in accuracy for positive words, whereas Grider and Malmberg (2008) found that positive words were more accurately recognized. In fact, these findings do not conflict with our argument. We do not argue that semantic relatedness is the only way that positive valence can affect recognition. Rather, the fact that semantic relatedness plays an important role in the positivity effect shows that multiple explanations are needed. This point might be made about the influence of emotion on memory more generally. Although we have not included negative emotion in our discussion, the inconsistency of findings in that literature is also likely due to the influence of multiple underlying factors. Positive and negative affect do influence memory differently in some ways, as was shown by the work reviewed earlier investigating the directionality of emotion. However, positive and negative stimuli also have properties in common that cause them to influence memory in similar ways. One of these is that negative items are also susceptible to semantic relatedness effects (Dougal & Rotello, 2007; Marotos, Allan, & Rugg, 2000; McNeely, Dywan, & Segalowitz, 2004; Windmann & Kutas, 2001).

The positivity effect illustrates the difficulty, faced in the wider literature, of separating mnemonic and attributional accounts of familiarity. Recent investigations of the pseudoword effect provide another example of this difficulty. Whittlesea and Williams (1998) found that phonologically and orthographically regular nonwords (i.e., pseudowords), such as *hension*, produce much higher rates of false recognition than do both real words and irregular nonwords such as *wicsta*. According to their attributional account, this is due to the fact that the surprising fluency of regular nonwords is easily misattributed to familiarity. However, Cleary, Morris, and Langley (2007) recently failed to reliably replicate this property of stimulus regularity in a systematic investigation of many types of materials. They noted that the regular nonwords used by Whittlesea and Williams were more phonologically similar to other stimulus items than were either words or irregular nonwords, suggesting that interstimulus similarity may have been behind the increase in false alarms. Consistent with this, in their own experiments, the pseudoword effect emerged only under conditions that drew attention to the phonological properties of the items. Thus, mnemonic factors make for a persuasive alternative account of the pseudoword effect. The attributional approach to memory reflects a growing appreciation of the role played by decision making and heuristics. In drawing attention to the possibility that mnemonic effects may be wrongly identified as attributional effects, the intent is not to detract from attribution theory but, rather, to focus its investigation.

AUTHOR NOTE

Correspondence concerning this article should be addressed to M. F. Verde, School of Psychology, University of Plymouth, Plymouth PL4 8AA, England (e-mail: michael.verde@plymouth.ac.uk).

REFERENCES

- ARNDT, J., & HIRSHMAN, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory & Language*, **39**, 371-391.
- BORNSTEIN, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin*, **106**, 265-289.
- BRADLEY, M. M., & LANG, P. J. (1999). Affective norms for English words (ANEW). Gainesville: University of Florida, NIMH Center for the Study of Emotion and Attention.
- CENTERBAR, D. B., SCHNALL, S., CLORE, G. L., & GARVIN, E. (2008). Affective incoherence: When affective concepts and embodied reactions clash. *Journal of Personality & Social Psychology*, **94**, 560-578.
- CLARK, S. E., & GRONLUND, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, **3**, 37-60.
- CLAYPOOL, H. M., HALL, C. E., MACKIE, D. M., & GARCIA-MARQUES, T. (2008). Positive mood, attribution, and the illusion of familiarity. *Journal of Experimental Social Psychology*, **44**, 721-728.
- CLEARY, A. M., MORRIS, A. L., & LANGLEY, M. M. (2007). Recognition memory for novel stimuli: The structural regularity hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 379-393.
- CLORE, G. L., & SCHNALL, S. (2005). The influence of affect on attitude. In D. Albarracín, B. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes* (pp. 437-489). Mahwah, NJ: Erlbaum.
- COOK, G. I., HICKS, J. L., & MARSH, R. L. (2007). Source monitoring is not always enhanced for valenced material. *Memory & Cognition*, **35**, 222-230.
- CORNEILLE, O., MONIN, B., & PLEYERS, G. (2005). Is positivity a cue or a response option? Warm glow vs. evaluative matching in the familiarity for attractive and not-so-attractive faces. *Journal of Experimental Social Psychology*, **41**, 431-437.
- CRISS, A. H., & SHIFFRIN, R. M. (2004). Context noise and item noise jointly determine recognition memory: A comment on Dennis and Humphreys (2001). *Psychological Review*, **111**, 800-807.
- DEWHURST, S. A., & ANDERSON, S. J. (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition*, **27**, 665-673.
- DOUGAL, S., & ROTELLO, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, **14**, 423-429.
- FEINGOLD, G. A. (1915). Recognition and discrimination. *Psychological Monographs*, **18**(2, Whole No. 78).
- GARCIA-MARQUES, T., MACKIE, D. M., CLAYPOOL, H. M., & GARCIA-MARQUES, L. (2004). Positivity can cue familiarity. *Personality & Social Psychology Bulletin*, **30**, 585-593.
- GILLUND, G., & SHIFFRIN, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.
- GRIDER, R. C., & MALMBERG, K. J. (2008). Discriminating between changes in bias and changes in accuracy for recognition memory of emotional stimuli. *Memory & Cognition*, **36**, 933-946.
- HARMON-JONES, E., & ALLEN, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from psychophysiological and individual differences approaches. *Personality & Social Psychology Bulletin*, **7**, 889-898.
- HINTZMAN, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, **95**, 528-551.
- JACOBY, L. L., & DALLAS, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, **110**, 306-340.
- JACOBY, L. L., & WHITEHOUSE, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, **118**, 126-135.
- KAPUCU, A., ROTELLO, C. M., READY, R. E., & SEIDL, K. N. (2008). Response bias in "remembering" emotional stimuli: A new perspective on age differences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **34**, 703-711.
- KUCERA, F., & FRANCIS, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- MACMILLAN, N. A., & CREELMAN, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MANDLER, G., NAKAMURA, Y., & VAN ZANDT, B. J. (1987). Nonspecific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **13**, 646-648.
- MAROTOS, E. J., ALLAN, K., & RUGG, M. D. (2000). Recognition memory for emotionally negative and neutral words: An ERP study. *Neuropsychologia*, **38**, 1452-1465.
- MCNEELY, H. E., DYWAN, J., & SEGALOWITZ, S. J. (2004). ERP indices of emotionality and semantic cohesiveness during recognition judgments. *Psychophysiology*, **41**, 117-129.
- MICKES, L., WIXTED, J. T., & WAIS, P. E. (2007). A direct test of the unequal-variance signal detection model of recognition memory. *Psychonomic Bulletin & Review*, **14**, 858-865.
- MONAHAN, J. L., MURPHY, S. T., & ZAJONC, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, **11**, 462-466.
- MONIN, B. (2003). The warm glow heuristic: When liking leads to familiarity. *Journal of Personality & Social Psychology*, **85**, 1035-1048.
- MURPHY, G. L., & MEDIN, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.
- NOSOFSKY, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.
- PHAF, R. H., & ROTTEVEEL, M. (2005). Affective modulation of recognition bias. *Emotion*, **5**, 309-318.
- RATCLIFF, R., SHEU, C., & GRONLUND, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, **99**, 518-535.
- REBER, R., WINKIELMAN, P., & SCHWARZ, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, **9**, 45-48.
- SCHNALL, S., & LAIRD, J. D. (2003). Keep smiling: Enduring effects of facial expressions and postures on emotional experience. *Cognition & Emotion*, **17**, 787-797.
- SCHNALL, S., & LAIRD, J. D. (2007). Facing fear: Expression of fear facilitates processing of emotional information. *Social Behavior & Personality*, **35**, 513-524.
- SCHWARZ, N., & CLORE, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality & Social Psychology*, **45**, 513-523.
- SCHWARZ, N., & CLORE, G. L. (1988). How do I feel about it? Informative functions of affective states. In K. Fiedler & J. Forgas (Eds.), *Affect, cognition and social behavior* (pp. 44-62). Toronto: Hogrefe International.
- SEAMON, J. G., MCKENNA, P. A., & BINDER, N. (1998). The mere exposure effect is differentially sensitive to different judgment tasks. *Consciousness & Cognition*, **7**, 85-102.
- SHIFFRIN, R. M., HUBER, D. E., & MARINELLI, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 267-287.
- VERDE, M. F. (2004). Associative interference in recognition memory: A dual-process account. *Memory & Cognition*, **32**, 1273-1283.
- VERDE, M. F., MACMILLAN, N. A., & ROTELLO, C. M. (2006). Measures of sensitivity based on a single hit rate and false alarm rate: The accuracy, precision, and robustness of d' , A_z , and A' . *Perception & Psychophysics*, **68**, 643-654.
- VERDE, M. F., & ROTELLO, C. M. (2003). Does familiarity change in the revelation effect? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 739-746.
- WESTERMAN, D. L. (2008). Relative fluency and illusions of recognition memory. *Psychonomic Bulletin & Review*, **15**, 1196-1200.
- WHITTLESEA, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1235-1253.
- WHITTLESEA, B. W. A. (1997). Production, evaluation and preservation of experiences: Constructive processing in remembering and performance tasks. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 37, pp. 211-264). New York: Academic Press.
- WHITTLESEA, B. W. A. (2002). Two routes to remembering (and another to remembering not). *Journal of Experimental Psychology: General*, **131**, 325-348.

- WHITTLESEA, B. W. A., & WILLIAMS, L. D. (1998). Why do strangers feel familiar, but friends don't? A discrepancy-attribution account of feelings of familiarity. *Acta Psychologica*, **98**, 141-165.
- WINDMANN, S., & KUTAS, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, **13**, 577-592.
- 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 & Social Psychology*, **81**, 989-1000.
- WINKIELMAN, P., SCHWARZ, N., FAZENDEIRO, T. A., & REBER, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation* (pp. 189-217). Mahwah, NJ: Erlbaum.
- ZAJONC, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality & Social Psychology*, **9**(Suppl.), 1-27.
- distinguished from *relative bias*, measured by indexes such as c and β , which describe criterion placement relative to the *old* and *new* distribution means.
2. Emotional valence and arousal are often confounded, and many of the studies reviewed earlier did not control for arousal. Because the purpose of the present study was to conceptually replicate previous demonstrations of the positivity effect, we likewise did not control for arousal.
3. A change in c_a (which measures location relative to the distribution means) can indicate a shift in absolute bias (which refers to a point on the evidence axis). In the top panel of Figure 1, c_a takes a lower value for C_L than for C . Note, however, that if subjects adopt the criterion C across the different conditions depicted in the three panels, absolute bias remains constant but c_a takes on a lower value in the lower two panels, due to the shift in evidence distributions.
4. However, Grider and Malmberg (2008) noted a problem with the accuracy index used in this study.

NOTES

1. The term *bias* refers here to *absolute bias*, the placement of the recognition criterion at a specific level of evidence. This should be

(Manuscript received March 30, 2009;
revision accepted for publication August 5, 2009.)