Published in final edited form as:
J Mem Lang. 2004 February ; 50(2): 196–211. doi:10.1016/j.jml.2003.09.001.

Illusions of face memory: Clarity breeds familiarity☆

Heather M. Kleider[*] and **Stephen D. Goldinger**
Department of Psychology, Arizona State University, Box 871104, Tempe, Arizona 85287-1104, USA

Abstract

When people perform a recognition memory task, they may avail themselves of different forms of information. For example, they may recall specific learning episodes, or rely on general feelings of familiarity. Although subjective familiarity is often valid, it can make people vulnerable to memory illusions. Research using verbal materials has shown that "old" responses are often increased by enhancing *perceptual fluency*, as when selected words are shown with relatively higher contrast on a computer. Conversely, episodic memory can create an erroneous sense of perceptual advantages for recently studied words. In this investigation, symmetric fluency effects were tested in face memory, a domain that is often considered neurologically and psychologically unique. In eight experiments involving over 800 participants, we found consistent memorial and perceptual illusions—fluency created feelings of familiarity, and familiarity created feelings of fluency. In both directions, these effects were manifested as response biases, suggesting effects based on memorial and perceptual attributions.

Keywords

Face recognition; Fluency; Heuristics; Recognition memory

When recalling information from memory, what decision-making processes are used to distinguish passing thoughts from true experiences? Recently, Whittlesea and Leboe (2000; Whittlesea & Williams, 2001a; 2001b) suggested that people use *memory decision heuristics* when evaluating recollections. By this hypothesis, the act of remembering (particularly recognition) entails two stages: (1) the production of mental responses to stimuli, and (2) evaluation of those responses. For example, suppose you encounter a famous person in an unexpected place, such as a neighborhood restaurant. In the first stage, the memory prompt (famous face) activates prior memory traces, as conceived in many theories (e.g., Hintzman, 1986). In the second stage, the source of this activation must be evaluated: for example, you may instantly recognize the person. Alternatively, you may only achieve a nagging feeling of familiarity, without ever achieving recognition.

☆This article is based on a doctoral dissertation by H.M. Kleider done at Arizona State University under the supervision of S.D. Goldinger. The work was supported by NRSA grant F32-MH63521-01 to H.M. Kleider and NIDCD Grant R01-DC04535-02 to S.D. Goldinger. We are grateful to Michelle Balhorn, Michael Shaw, and Alice Kirk for their help in data collection. We also thank Don Homa, Stan Parkinson, Michael McBeath for helpful comments during the research. Colleen Kelley, Don Read, Bruce Whittlesea, and James Bartlett all provided helpful comments on prior drafts, and Larry Jacoby provided invaluable feedback, greatly improving the paper.

[*] Corresponding author. heatherkleider@asu.edu (H.M. Kleider).

Jacoby and Dallas (1981) originally proposed that recognition decisions entail attributions, building upon a theory by Mandler (1980). According to Mandler, people can make recognition decisions using different forms of information, either retrieval of specific encoding events or general feelings of familiarity. In fact, these classifications of experience later formed the response options in the "remember/know" paradigm (Tulving, 1985). Although most people equate remembering with the former experience (episodic retrieval), Jacoby and Dallas (1981; also Kelley, Jacoby, & Hollingshead, 1989) argued that "old" responses often reflect familiarity. The different reliance on retrieval or familiarity is driven by task or stimulus factors—when retrieval is made difficult (e.g., by limiting rehearsal), people rely more on familiarity cues. When those cues are directly manipulated, people may experience a false sense of memory.

In essence, the memory-attribution framework suggests that recognition often requires a person to decide that a target stimulus feels "old," although its specific study episode cannot be recalled. Without this critical cue, people behave in a manner consistent with signal-detection theory: Some strength of evidence (familiarity) is evoked by a stimulus, which is then evaluated against an internal criterion. Thus, recognition is often inferential. Returning to the previous example, it is generally uncommon to encounter celebrities in daily life, so most people would never resolve their nagging sense of familiarity. By contrast, when dining in Hollywood, people may interpret every tingle of familiarity as a brush with fame—the change of venue is used as a "rule of thumb," creating a more liberal criterion. Without absolute criteria for discriminating true and false recognition, people rely on *memory decision heuristics*. Whittlesea and Leboe (2000) described three such heuristics, called *generation, resemblance*, and *fluency*. The present study focused on the fluency heuristic, as originally described by Jacoby and Dallas (1981).

The Fluency heuristic

By the fluency heuristic, Jacoby and Dallas (1981) and Jacoby, Kelley, and Dwyan (1989) suggest that, when familiarity is the major determinant of recognition, people often use the fluency (ease) of perceptual processing as a memory cue. Many data suggest that perceptual processing is enhanced when target stimuli are more familiar (Jacoby & Dallas, 1981; Logan & Etherton, 1994). People seem to implicitly assume this relationship, as suggested by "memory illusions" created by manipulations of fluency. That is, when stimulus perception is enhanced, feelings of familiarity often arise, leading to increased "old" recognition judgments. Although this effect occurs among old items, it is generally larger for new items, because familiarity is their only available cue. When fluency increases familiarity (appropriately or not), people will show a liberal criterion shift in recognition.

Prior studies have shown that fluency can create illusions of memory. For example, Jacoby and White-house (1989) showed participants a study word list, followed by a standard recognition test. During the test, all words were preceded by subliminal primes (either related or unrelated to the targets). Related primes evoked more "old" responses (increasing both hits and false-alarms) than unrelated primes. The authors suggested that related primes facilitate lexical access—this enhanced perception is experienced as familiarity. When participants were made aware of the priming words, the effect was eliminated. In experiments combining word identification (in noise) with recognition judgments, small improvements in signal-to-noise ratios often elicit more “old” judgments. This has been shown in both the visual (Whittlesea, Jacoby, & Girard, 1990) and auditory (Goldinger, Kleider, & Shelley, 1999) domains. Moreover, Whittlesea (1993) showed that variations in *conceptual* fluency also create familiarity illusions. In one experiment, people judged whether target words were semantically related to any words in previous study lists. “Conceptual fluency” was manipulated by presenting target words in either predictive or neutral sentences. Words in predictive sentences evoked more (correct and incorrect) “old” responses than words in neutral sentences. Whittlesea suggested that contextually supported words have a processing advantage that feels like familiarity.

The foregoing studies show that perceptual fluency can affect memory judgments. Others have shown the complementary effect—i.e., that memory can affect perceptual judgments. For example, Witherspoon and Allan (1985) showed people study words, followed later by new and old test words. In a duration judgment task, participants consistently gave longer time estimates to previously studied words, suggesting that recent memory facilitated perception, creating a false sense of bottom-up support (see also Whittlesea et al., 1990). In the auditory domain, Jacoby, Allan, Collins, and Larwill (1988) played old and new sentences to listeners. These were mixed with varying levels of white noise; participants made recognition judgments and subjective noise estimates. Old sentences gave the impression of greater perceptual clarity (less noise), even when listeners believed the sentences were new (see also Goldinger et al., 1999).

Face recognition

In Whittlesea and Leboe’s (2000) framework, the fluency heuristic is portrayed as a general principle relating memory and perception. However, fluency effects have typically been tested using linguistic stimuli, such as words or sentences (Goldinger et al., 1999; Whittlesea et al., 1990). This limited test-bed raises a potential concern: Although reading is a highly practiced perceptual process, it is a learned behavior. As such, it may be particularly vulnerable to fluency manipulations, relative to more ingrained perceptual processes. By contrast, face recognition is a natural ability, present at birth (Pascalis, Petit, Kim, & Campbell, 1999; Segerstrale & Molnar, 1997). Many data suggest that infants pay special attention to faces. By just 30 min of age, infants track moving faces farther than other moving patterns of comparable contrast, complexity, and spatial frequency (Johnson, Dziurawiec, Ellis, & Morton, 1991). Neuropsychological data suggest that recognition memory for faces relies on unique brain areas, separate from those for words (Cousins, Hanley, Davies, Turnbull, & Playfer, 2000; Farah, Klein, & Levinson, 1995). As such, face recognition is an interesting testing ground for the ubiquity of fluency effects in memory retrieval.

In the present research, we investigated fluency effects in face recognition.[1] Following prior studies (e.g., Whittlesea et al., 1990), we assessed symmetric bottom–up and top–down

---

[1] In the present article, we apply the term “face recognition” in an experimentally constrained manner. Specifically, we intend a situation wherein people view novel faces at study, and later discriminate those faces (identical photographs) from distracters in a recognition memory test. Thus, “face recognition” refers to an experimental task (old-new discrimination), rather than person memory.

illusions. That is, we assessed perceptual fluency effects on memory judgments, and memorial effects on perceptual judgments. Creating illusions of face memory through perceptual manipulation may seem unlikely, given the robust nature of face recognition in healthy young adults. However, the literature provides several reasons to predict such an effect. First, perceptual fluency effects are rather ubiquitous, occurring in domains beyond word or sentence memory. For example, the fluency heuristic leads to increased fame judgments for previously seen, nonfamous names (Jacoby et al., 1989, 1989), increased truth judgments for statements (Begg & Armour, 1991), and biased judgments of other people's performance (Jacoby & Kelley, 1987).

Second, many behavioral and neuropsychological data suggest that face perception is "special," engaging more holistic processing than occurs with other stimuli (Farah, Wilson, Drain, & Tanaka, 1998). That is, faces are not typically processed as collections of parts, but as entire configurations. As a result, face memory is likely to rely upon overall familiarity, rather than semantic elaboration. Indeed, popular memory improvement books (e.g., Lorayne & Lucas, 1996) typically advise readers to enhance face memory by verbally rehearsing whatever distinctive features they may detect. Presumably, this advice exists in self-help books because people do not adopt this strategy spontaneously. According to Jacoby and Dallas (1981), when stimuli are less amenable to elaboration, they are more susceptible to perceptual fluency effects. Jacoby and Witherspoon (1982) found evidence consistent with this claim: recognition memory for real words was relatively unaffected by perceptual fluency, whereas memory for nonwords was strongly affected. Jacoby and Witherspoon suggested that people cannot rely on elaborative strategies for nonwords, making their decisions more fluency-based. We expected faces to follow this pattern.

Given its practical importance, face recognition has been extensively studied. Much of the literature relates to eyewitness memory and its systematic errors. Although fluency effects have not been directly tested, two prior studies relate to the present research: Bartlett, Strater, and Fulton (1991) compared face recognition (albeit indirectly) in younger and older adults. Their investigation was motivated by findings that older people tend to generate more false-alarms than younger people. Bartlett et al. cited a model of face recognition (Bruce & Young, 1986) that resembles Mandler's (1980) theory, positing that face recognition entails two retrieval processes. The first computes similarity of probe faces to stored traces (i.e., familiarity); the second involves explicit recall of personal information. Bartlett et al. hypothesized that explicit recollection wanes with aging, leaving familiarity as the main basis for recognition and increasing the potential for false-alarms. To test this, they conducted two "exclusion" experiments, one assessing judgments of face-viewing recency, the other assessing false-fame judgments (Jacoby, Woloshyn, & Kelley, 1989). Among older participants, familiar faces increased both recency and fame judgments.

In another study, Busey, Tunnicliff, Loftus, and Loftus (2000, Experiment 3) examined yet another possible dichotomy of processes in face-recognition. Specifically, Busey et al., sought to determine whether recognition accuracy and confidence derive from a common

---

When their focus is person memory (as in applied studies), researchers often present slightly different photographs of the same people across study and test.

source (e.g., the strength of activating a stored memory trace), or whether they have separate underlying sources. Their hypothesis was that recognition is affected by different sources (e.g., initial encoding time, rehearsal), reflecting both stimulus familiarity and elaboration. Confidence, however, was hypothesized to reflect mainly the retrieval of specific details. In their experiment, people studied and were later tested on faces that were either dim or bright, in a fully crossed design. The findings supported the dual-process framework: Recognition confidence was not isomorphic with accuracy. Of primary interest to the present research, Busey et al. found that bright test faces led to *attributions* of accuracy. That is, true accuracy followed a pattern predictable from *encoding specificity* (Tulving & Thompson, 1973)—faces studied under dim conditions were better recognized when tested dim. However, when originally dim faces were tested bright, accuracy decreased while confidence increased. Fluent encoding apparently feels like memory, even when leading one astray.

Overview

In the present study, fluency effects were tested by creating perception-driven memorial illusions (Experiments 1–5) and memory-driven perceptual illusions (Experiments 6–7). Experiments 1–5 all followed the same basic procedure: a series of faces were presented for study, followed by a distracter task and a test phase. During recognition tests, participants saw new and old faces, embedded in varying levels of visual white noise. Following previous studies, we expected “old” responses to increase to clear faces, as evidenced by liberal bias shifts. This hypothesis was tested within each experiment, and by comparison to a baseline condition (Experiment 1). Experiments 1–4 tested whether the anticipated fluency effect was sensitive to noise level variation and contextual changes. Experiment 5 extended the test across modality, pairing faces with auditory noise. Experiments 6–7 tested whether previously seen faces influence subjective judgments of clarity or presentation time.

Experiments 1–5: Perceptual effects on memory

People are typically quite accurate at remembering faces, even if study exposure is brief. Experiment 1 was conducted to assess baseline recognition accuracy using our materials and procedures, without any stimulus manipulations. A series of photographs were shown, followed by a distracter number/letter classification task. Participants then received a face recognition test, with all faces clearly presented.

Method

Participants—Forty-seven Arizona State University students, both undergraduates and graduates, participated in Experiment 1. All later experiments included introductory psychology students exclusively. The undergraduates received partial course credit, and all participants (in all experiments) were treated in accordance with APA ethical standards. Both male (19) and female (28) students participated, in proportions that reflect psychology course enrollments.

Materials—All experiments used a common pool of 36 black-and-white photographs of male faces. The photos, taken specifically for this study, were front-facing head and shoulder views of young men (ages 16–19 years), all with the same background. The original photos were digitized and equated for image size using Adobe Photoshop, then loaded on PC-compatible computers for display. All computers were equipped with four-button response boxes.

Design and procedure—This experiment contrasted only photo type (old versus new), with photos counterbalanced across participants. Participants were tested in groups of 5–8, seated in separate booths equipped with computers and response boxes. Response buttons were marked “new,” “old,” “letter,” and “number.” Participants were familiarized with the response buttons, and received both oral and written instructions. The study session involved 18 photos, presented serially for 2 s each. Students were instructed to memorize the faces, without making any responses. The distracter task followed, composed of 32 trials of a letter/number forced-choice classification task. The stimuli were displayed for 50 ms each, and students responded as quickly and accurately as possible. In the recognition test, 36 photos (18 new, 18 old) were presented serially for 2 s each. Immediately after each photo, a response prompt “*new or old*?” appeared, remaining until the participant responded or 3 s elapsed. Participants were instructed to respond as accurately as possible. Once a response was recorded, the screen cleared. All programming was done using ERTS software. The procedures in this and all subsequent experiments followed Arizona State University IRB guidelines.

Results and discussion

In Experiments 1–5, our analyses focused on signal-detection measures of discrimination and bias. For each participant, we calculated four indices: The first two were signal-detection measures for sensitivity (d') and the intersection bias measure C. As explained by Feenan and Snodgrass (1990), C is preferable to β for recognition memory, as β may be correlated with discrimination. C is centered around zero: positive values represent a conservative bias; negative values represent a liberal bias. The other two indices were P_r and B_r, the sensitivity and bias measures from the “two-high threshold” model of recognition. P_r is a common accuracy score, representing the difference of hits and false-alarms. B_r is defined as the probability of responding “old” despite uncertainty ($B_r = FA = (1 - P_r)$), and is centered around .5: Values lower than .5 reflect a conservative bias; values above .5 reflect a liberal bias.

Table 1 displays all hit and false-alarm rates for Experiments 1–5. In this straight recognition task without stimulus manipulations, these rates were 84 and 17%, respectively. Participants were fairly accurate at recognizing faces ($d' = 1.952$; $P_r = .67$). More important (for comparison to later experiments), participants showed no evidence of bias ($C = -.002$; $B_r = .501$) in either direction. Similar criteria have been observed for faces studied individually (Brown & Lloyd-Jones, 2002; O’Toole, Deffenbacher, Valentin, & Abdi, 1994) and faces studied in pairs (Winograd & Rivers-Bulkeley, 1977).[2] These results were used as a benchmark for comparison in Experiments 2–5.

---

[2] Because Experiment 1 was brief and engaging, we used it (same materials and procedure) as a distraction-task for another study ($N = 58$). Despite the situational differences, the results were remarkably similar to those of Experiment 1 (83.6% hits, 16.4% false-alarms, $P_r = .683$, $B_r = .495$), providing a replication.

Experiment 2: Clear and 30% distortion

When making recognition decisions, people often rely on familiarity. When words are used as stimuli, familiarity illusions can be created by increasing the ease of perceptual processing. In Experiment 2, a visual mask (30% distortion) was added to half the faces during the test phase, making them relatively difficult to process. In doing so, we expected that clear faces would seem comparatively easy to process, creating a false sense of familiarity and thus a liberal bias.

Method

Participants—Ninety-three students participated for partial course credit.

Materials—The photos from Experiment 1 were used again. For each photo, a second version was created wherein Gaussian noise was added using Adobe PhotoShop. The visual noise was added at a 30% level, giving the photos a slightly “fuzzy” appearance (see Fig. 1).

Design *and Procedure*. Experiment 2 comprised a 2 × 2 design, crossing photo type (old, new) by noise level (clear, 30% distorted). Procedures were identical to those of Experiment 1, with 18 photos shown at study and 36 shown at test. In the test phase, half the photos were clear, and half were distorted. The assignment of photos to conditions (old, new, clear, and noisy) was counterbalanced across participants. Before the test phase, participants were instructed that some faces would appear “fuzzy,” as if recorded by a low-quality security camera.

Results and discussion

As shown in Table 1, hit rates to clear and noisy faces were 85 and 68%, respectively. False-alarm rates to clear and noisy faces were 24 and 26%, respectively. In discrimination, performance to clear faces ($d' = 1.725$; $P_r = .605$) exceeded that for noisy faces ($d' = 1.105$; $P_r = .418$), a reliable difference [for P_r, $F(1, 92) = 45.4$; MSe = .036; $\eta^2 = .330$].[3] This difference in sensitivity was certainly expected, reflecting the relatively poor quality of the noise-obscured photos. Of greater interest were the bias measures, which were slightly liberal for clear faces ($C = -.16$; $B_r = .611$), and slightly conservative for noisy faces ($C = .089$; $B_r = .447$). This difference was also reliable [for B_r, $F(1, 92) = 24.9$; MSe = .040; $\eta^2 = .212$].

Our expectation for Experiment 2 was that, by conferring relative perceptual fluency, clear faces would induce a liberal bias shift, such that fluency is misattributed to familiarity. Although we observed the expected liberal shift for clear faces, we also observed a *conservative* shift to noisy faces. This may reflect probability matching by our participants—having answered “old” to more than half the clear faces, they may have compensated by disproportionately answering “new” to noisy faces, thus evenly distributing their responses. No matter the source, these divergent biases complicate interpretation of the anticipated

---

[3] As described in Results section of Experiment 1, we calculated two sets of signal-detection measures for each participant. We also conducted ANOVAs on both sets. As one might expect, the ANOVA results, in terms of reliable and null effects, were identical across tests. Therefore, we only report analyses based on the two-high threshold indices, P_r and B_r.

liberal shift. To further assess its reliability, we next compared the clear trials from Experiment 2 to the results of Experiment 1. This is an ideal comparison, as the materials and procedures were identical; the only difference concerned the presentation context for clear faces in Experiment 2.

Regarding sensitivity, performance in Experiment 1 ($d' = 1.952$; $P_r = .67$) exceeded that for the clear trials in Experiment 2 ($d' = 1, 725$; $P_r = .605$), although this difference was not reliable [$F(1, 138) = 1.93$; MSe = .051; $\eta^2 = .014$]. As expected, the bias in clear faces was liberal in Experiment 2 ($C = -.16$; $B_r = .611$), relative to that in Experiment 1 ($C = -.002$; $B_r = .501$). This difference was reliable [$F(1, 138) = 5.43$; MSe = .057; $\eta^2 = .038$].

Previous articles have suggested that feelings of familiarity result from an evaluation process that integrates not only perceptual fluency, but also the *expectation* of processing difficulty in trying circumstances. For example, Jacoby and Whitehouse (1989) increased perceptual fluency by subliminally priming target words. Because priming was subliminal, fluency for these target words could not be attributed to any obvious perceptual cause, and participants interpreted fluency as familiarity. They suggested that the difference between participants' expectation and experience produced the illusion of familiarity. More recently, Whittlesea and Williams (1998, 2001a) proposed a *discrepancy-attribution hypothesis*, wherein "familiarity" arises from the perceived *coherence* of processing episodes, rather than fluency alone. That is, feelings of familiarity reflect both perception and comprehension of stimulus events. As people integrate various aspects of an event for evaluation on a single dimension (e.g., familiarity), they naturally compare their current experience to others in the same domain (see discussion of "relative fluency" in Jacoby & Dallas, 1981). In the present case, clear faces may inappropriately boost familiarity because general expectations of fluency have been lowered. If so, we may further investigate the effect by holding clear faces constant, while further varying the context in which they appear.

Experiments 3A and 3B: Context and vigilance

In Experiment 2, we observed a liberal criterion shift to clear faces, relative to both noisy faces and clear faces from Experiment 1. Nevertheless, the effect was relatively small (about an 11% shift). In Experiment 3A, we added a third noise level, hoping to further diminish participants' overall expectations of clarity. By doing so, clear faces should present greater discrepancy between expected and perceived fluency, leading to higher levels of false familiarity.

In Experiment 3B, we further tested the discrepancy-attribution hypothesis by specifically warning participants about our manipulation. That is, we informed them that noise levels had no bearing on the "old-new" status of any photograph. If an attribution process is truly involved in creating false familiarity, increasing participants' vigilance to the manipulation may attenuate the fluency effect. Note that our vigilance manipulation differs from that in previous studies, wherein perceptual fluency was manipulated by more subtle procedures (e.g., Jacoby & Whitehouse, 1989; Whittlesea et al., 1990, Experiment 3). In those studies, participants were originally unaware that their perceptual experiences were subject to manipulation, and thus could not attribute fluency to anything other than familiarity. Given our stimuli, there was little doubt that participants could appreciate differences between clear and noisy. Our instructions were meant to guard against either of two conscious processes: (1) an assumption by participants that clear photos (which bore greater resemblance to study photos) are more likely “old,” and (2) an assumption by participants that the experimenter *wanted* them to say “old” more frequently to clear photos.

Method

Participants—Experiment 3A included 213 students, and Experiment 3B included 192 students; all received course credit. Nobody participated in more than one experiment.

Materials—In Experiment 2, we used two sets of photographs, clear and 30% noise. For Experiments 3A and 3B, we created a third set with 60% noise added. For discussion purposes, these noise levels are hereafter called *clear*, *moderate noise*, and *heavy noise*. Fig. 1 shows an example of one stimulus photo at all three noise levels.

Design and procedure—Both Experiments 3A and 3B used 2 × 3 within-subject designs, with photo type (new, old) and noise level (clear, moderate, and heavy) as the manipulated variables. The only difference between Experiments 3A and 3B concerned the instructions to participants. The procedures were identical to Experiment 2, except new and old faces were equally divided among 3 noise levels (fully counterbalanced, as before). The faces were presented in a random order.

Results and discussion

Experiment 3A—As shown in Table 1, the hit rates in Experiment 3A were 85, 69, and 60%, respectively, to faces in clear, moderate, and heavy noise. The corresponding false-alarm rates were 33, 26, and 34%. Turning first to sensitivity, we observed a predictable falling trend as noise levels increased. Across conditions, the discrimination measures were: clear ($d' = 1.447$; $P_r = .511$), moderate noise ($d' = 1{:}156$; $P_r = .435$), and heavy noise ($d' = 0{:}659$; $P_r = .257$). Combined, the conditions revealed a powerful main effect of noise [$F(1, 212) = 94{:}48$; MSe=.073; $\eta^2 = .308$]. Planned comparisons revealed that all three conditions reliably differed from one another (all $p < .01$, with Bonferroni correction).

In the criterion measures, we again observed a liberal shift for clear trials, coupled with small conservative shifts for noisy trials. Across conditions, the criterion measures were: clear ($C = -.294$; $B_r = .684$), moderate noise ($C = .072$; $B_r = .456$), and heavy noise ($C = .079$; $B_r = .459$). Combined, the conditions revealed another main effect of noise [$F(1, 212) = 56{:}41$; MSe = .073; $\eta^2 = .210$]. Of greater interest, planned comparisons revealed that the clear condition reliably differed from both noise conditions (each $p < .01$), but the noise conditions were statistically equivalent ($p = .775$).

As before, we compared the clear trials from Experiment 3A to the results of Experiment 1. Consulting Table 1, hit rates were quite comparable across relevant conditions, differing by only 1%. However, false-alarms in Experiment 3A doubled those in Experiment 1. As a result, the experiments reliably differed in both sensitivity [$F(1, 258) = 9.75$; MSe=.089; $\eta^2 = .036$] and bias [$F(1, 258) = 12{:}18$; MSe = .081; $\eta^2 = .045$]. Finally, we compared the bias value from the clear condition of Experiment 3A to that of Experiment 2, to assess whether adding a third noise level resulted in a reliably larger criterion shift. Although people were more liberal to clear faces in Experiment 3A, the result was only marginal [$F(1, 304) = 3.27$; MSe = .24; $p = .06$; $\eta^2 = .021$].

Experiment 3B—Hit rates in Experiment 3B were 88, 70, and 60%, respectively, to faces in clear, moderate, and heavy noise. The corresponding false-alarm rates were 30, 28, and 31%. In sensitivity, we again observed a falling trend as noise increased. Across conditions, the discrimination measures were: clear ($d' = 1:728$; $P_r = .586$), moderate noise ($d' = 1:117$; $P_r = .423$), and heavy noise ($d' = 0.725$; $P_r = .281$). Combined, the conditions revealed a powerful main effect of noise [$F(1, 191) = 140.66$; MSe = .064; $\eta^2 = .424$]. Planned comparisons revealed that all three conditions reliably differed from one another (all $p < .001$).

In criterion measures, we again observed a liberal shift for clear trials, and small conservative shifts for noisy trials. Across conditions, the criterion measures were: clear ($C = -.333$; $B_r = .720$), moderate noise ($C = .025$; $B_r = .484$), and heavy noise ($C = .121$; $B_r = .437$). Combined, the conditions revealed another main effect of noise [$F(1, 191) = 83:30$; MSe = .070; $\eta^2 = .304$]. Of greater interest, planned comparisons revealed that the clear condition reliably differed from both noise conditions (each $p < .001$), but the noise conditions were statistically equivalent ($p = .112$).

We again compared the clear trials from Experiment 3B to the results of Experiment 1. Consulting Table 1, hit rates were slightly superior in Experiment 3B, differing by 5%. This improvement, however, was offset by a 13% increase in false-alarms. The experiments differed in sensitivity, but not to a reliable degree [$F(1, 237) = 3:13$; MSe = .068; $p = .078$]. They also differed in bias, which was reliable [$F(1, 237) = 14:48$; MSe = .084; $\eta^2 = .058$]. Finally, as in Experiment 3A, we compared the bias values from Experiment 3B and Experiment 2, to assess whether a third noise level resulted in a larger criterion shift. In this case, people were more liberal to clear faces in Experiment 3B, and the result was reliable [$F(1, 283) = 5:15$; MSe = .08; $\eta^2 = .039$].

The results of Experiments 3A and 3B were partly expected and partly surprising. In Experiment 3A, we anticipated that adding a third noise level would reduce participants' expectations of clarity, thus increasing the “familiarity” response to clear faces. This prediction was confirmed, supporting the discrepancy-attribution hypothesis. As suggested by Whittlesea and Williams (2001a, 2001b; Jacoby & Whitehouse, 1989), people are not impressed by fluency alone when making recognition decisions. Instead, there must be a difference between actual and expected fluency, given the context. When clear faces “pop out” in the context of noisy faces, perceptual fluency may be attributed to familiarity.

In this light, the results of Experiment 3B were not anticipated. In similar experiments with verbal materials, providing observers with alternate attributions (e.g., making prime words supraliminal) removed the fluency-familiarity linkage. When participants know that fluency does not reliably cue familiarity, effects such as those in Experiments 2 and 3A should no longer occur.[4] Nevertheless, the results of Experiment 3B were nearly identical to those of

Experiment 3A. Moreover, we repeated this procedure with more detailed warnings, testing over 50 participants, and the same pattern was observed. Accordingly, we are confident that the persistent fluency effect in Experiment 3B is a valid observation, rather than a type II error. (On a positive note, Experiment 3B suggests that demand characteristics did not create the previous results.)

Thus far, the similarities of Experiments 2 and 3A with prior research would suggest that fluency effects with faces and words are fundamentally similar. Experiment 3B, however, suggests that perceptual processing of faces may be somewhat encapsulated (Fodor, 1983), such that momentary intentions or warnings have little effect. Indeed, studies of face recognition suggest that faces receive more holistic processing than other, less integral stimuli (Farah et al., 1998; Rhodes, 1988; Sergent, 1984). In this light, Experiment 3B may demonstrate that people cannot separately attend to faces and added noise, and therefore cannot compensate for their perceptual experience after-the-fact.

Experiment 4: Reduced noise levels

In Experiment 3A, adding a heavy noise condition increased the liberal criterion shift to clear faces. In Experiment 4, we replicated Experiment 3A with a less extreme range of noise levels. We considered this important for two main reasons: first, a concern from Experiment 3A was that the 60% noise level may have been too hard for participants, possibly discouraging their best efforts. Second, we worried that increased false-alarms to clear faces may have reflected the extreme contrast to noisy faces—replicating the pattern with less extreme values would alleviate this concern.

Method

Participants—Eighty students participated for partial course credit.

Materials—The materials were the same as Experiment 3A, except the heavy (60%) noise level was replaced with a *high-moderate* (40%) level, and the moderate (30%) level was replaced with a *soft* (20%) level. Subjectively, the clear and soft noise levels appeared quite similar; the high-moderate level still appeared quite fuzzy.

Design and procedure—The design and procedure were identical to those of Experiment 3A.

Results and discussion

As shown in Table 1, the hit rates in Experiment 4 were 86, 72, and 68%, respectively, to faces in clear, soft, and high-moderate noise. The corresponding false-alarm rates were 30, 23, and 28%. Turning first to sensitivity, we again observed a falling trend as noise levels increased. Across conditions, the discrimination measures were: clear ($d' = 1.628$; $P_r = .566$), soft noise ($d' = 1.30$; $P_r = .483$), and high-moderate noise ($d' = 1.046$; $P_r = .398$).

---

[4] Notably, explicit warnings are not always effective against illusions, even in experiments with verbal materials. Whittlesea and Williams (1998) reported a fluency illusion that survived such warnings, suggesting that some effects are cognitively impenetrable. This may be the case with faces in noise. In a later study, Whittlesea and Williams (2001b) eliminated the illusion, but only when people processed the stimuli for a new purpose.

Combined, the conditions revealed a main effect of noise [$F(1,79) = 15.65$; MSe = .072; $\eta^2 = .165$]. Planned comparisons revealed that the clear and high-moderate noise conditions reliably differed ($p < .001$), but the soft noise condition was statistically equivalent to both other conditions.

In the criterion measures, we again observed a robust liberal shift for clear trials and small conservative shifts for noisy trials. Across conditions, the criterion measures were: clear ($C = -.283$; $B_r = .686$), soft noise ($C = .077$; $B_r = .451$), and high-moderate noise ($C = .063$; $B_r = .463$). Combined, the conditions revealed another main effect of noise [$F(1, 79) = 36.98$; MSe = .068; $\eta^2 = .319$]. Of greater interest, planned comparisons revealed that the clear condition reliably differed from both noise conditions (each $p < .001$), but the noise conditions were statistically equivalent to each other ($p = .494$).

As before, we compared the clear trials from Experiment 4 to the results of Experiment 1. Consulting Table 1, hit rates were comparable across the relevant conditions, differing by 3%. However, false-alarms in Experiment 4 were 13% higher those in Experiment 1. As a result, the experiments again differed in both sensitivity [$F(1, 125) = 4.70$; MSe = .058; $\eta^2 = .036$] and bias [$F(1; 125) = 15.01$; MSe = .070; $\eta^2 = .107$]. Finally, we again compared the bias values across Experiments 4 and 2, to assess the impact of adding a third noise level. As before, people had a larger liberal shift to clear faces in Experiment 4 [$F(1, 171) = 5.31$; MSe = .063; $\eta^2 = .072$].

Experiment 4 closely replicated Experiment 3A, despite the changes in overall noise levels. Indeed, Experiments 2, 3A, 3B, and 4 all produced similar results, despite stimulus and procedural differences. Fig. 2 provides an overview of the data to this point: as shown in the left panel, sensitivity (P_r) was strongly and uniformly affected by noise (note that P_r increased in Experiment 4, reflecting lower noise levels). Most important, the right panel of Fig. 2 shows that participants set more liberal criteria for clear faces in every experiment using noise. This is shown in two ways: first, reliable liberal shifts occurred to clear faces in every experiment using noise, all relative to Experiment 1. Second, and perhaps more important, similar liberal shifts occurred to clear faces *within* every experiment using noise: when clear and noisy faces were intermixed during test, people set liberal criteria for clear faces, and slightly conservative criteria for obscured faces. Notably, whereas sensitivity varied continuously with noise levels, bias was applied in a more absolute manner. We consider this further in the General discussion.

Experiment 5: Auditory noise

Together, Experiments 1–4 suggest that increases in perceptual fluency lead to increases in face familiarity. Before turning to the opposite direction of effect (memorial influences on perception), we conducted a final control experiment to assess a possible cognitive-workload account of the previous data. Given our results thus far, we hypothesized that relative perceptual fluency—the ease of seeing clear faces, relative to noisy faces—caused our effects. However, clear trials may have conferred a secondary benefit to participants, such that all cognitive processes may have been (or felt) easier. If so, the criterion shifts may reflect variations in cognitive workload, rather than face perception. In similar research using spoken words, Goldinger et al. (1999) found that subjective impressions of noise generally increased when participants performed a dual-task, suggesting that workload and subjective perceptual fluency are connected.

In Experiment 5, we maintained the general design and procedures of all prior experiments, but we replaced the visual noise with two levels of auditory noise. All photos were clearly shown, with one-third accompanied by silence, one-third by a burst of moderate white noise, and one-third by a burst of relatively loud white noise. The noise (especially the louder noise) was set at levels that the investigators (and research assistants) found distracting and annoying. We thus introduced a distracting stimulus without dramatically compromising the similarity to Experiments 1–4. If the previous “fluency” effects were actually due to variations in cognitive workload, similar results would be anticipated in Experiment 5.

Method

Participants—One-hundred-one students volunteered for partial course credit.

Materials—All clear photos were used; previous levels of visual noise were replaced with three levels of auditory white noise: clear (no noise), moderate noise (approximately 75 db SPL) or loud noise (approximately 90 db SPL). The noise was created using a Matlab white noise generator and was disseminated over Senneheiser HD250 headphones.

Design and procedure—The design and procedure were the same as Experiment 4, except that participants wore headphones throughout the experiment (although they were only required for the test phase). Participants were told that, during test, some photos would be accompanied by white noise. They were instructed to make face recognition judgments, disregarding the noise if possible. The timing of noise onset and offset was the same as visual noise in the previous experiments.

Results and discussion

As shown in Table 1, the hit rates in Experiment 5 were 80, 76, and 75%, respectively, to faces accompanied by quite, moderate, and loud noise. The corresponding false-alarm rates were 25, 23, and 25%. In sensitivity, we observed little effect as noise levels increased. Across conditions, the discrimination measures were: quiet (d' = 1:506; P_r = .547), moderate noise (d' = 1:444; P_r = .529), and loud noise (d' = 1:321; P_r = .491). Although smaller than prior effects, the combined conditions did reveal a main effect of noise [F(1, 100) = 6:85; MSe = .085; η^2 = .041]. However, planned comparisons revealed that no conditions reliably differed from one another (all $p > .35$).

In the criterion measures, we observed little evidence for criterion shifts in any condition. Across conditions, the criterion measures were: quiet (C = −.08; B_r = .553), moderate noise (C = .027; B_r = .482), and loud noise (C = .00; B_r = .50). Unlike prior experiments, there was no main effect of noise, nor any significant differences among conditions. Finally, we compared the quiet trials from Experiment 5 to the results of Experiment 1. Consulting Table 1, both hit rates and false-alarms were considerably worse in Experiment 5, falling and rising by 4 and 9%, respectively. As a result, the experiments reliably differed in sensitivity

[$F(1, 146) = 4.59$; MSe = .088; $\eta^2 = .027$]. However, the analysis revealed no differences in criteria ($p > .85$).

The main finding of Experiment 5 was that the pattern from Experiments 2–4 was not replicated— performance was similar to all photos, regardless of auditory noise levels. Thus, auditory noise (intended to manipulate workload) did not behave like visual noise (intended to manipulate perceptual fluency). This is obviously a null result, and must therefore be interpreted with caution. However, its worrisome nature is mitigated by two facts. First, direct comparisons to Experiment 1 verified that the auditory noise manipulation was effective, reducing overall sensitivity. As such, the null bias result should not simply reflect a weak manipulation. Second, the results of Experiment 5 are stable: we repeated this experiment three more times, using different auditory signals and instructions to participants. Similar results were observed in every replication. Given these facts, we suggest that the fluency effects seen in Experiments 2–4 were likely caused by variations in face perception, not cognitive workload.

Experiments 6–7: Memorial effects on perception

It has been frequently suggested that memory and perception are fundamentally linked, that neither process can be examined separately from the other (Johnson, 1983; Roediger, 1996; Whittlesea, 1997). One line of evidence is that perceptual fluency affects memory judgments for printed and spoken words (Goldinger et al., 1999; Witherspoon & Allan, 1985). Experiments 2–4 showed similar results in face recognition. Previous studies have also shown the *opposite* direction of effect, with memory for words affecting the subjective experience of their later perception (Whittlesea et al., 1990). In Experiments 6–7, we assessed whether prior exposure to faces would similarly affect subjective perceptual experience (see Reber, Winkielman, & Schwarz, 1998, for a similar approach).

Experiment 6: Noise judgments

In Experiment 6, we tested whether prior experience with faces would induce the subjective impression of greater perceptual clarity. Given the photos and noise levels from Experiment 4, participants gave subjective noise ratings, followed by a recognition test. If face memory affects perceptual experience, previously seen faces should seem clearer than new faces, at least in the two noisy conditions.

Method

Participants—Fifty-one students received partial course credit for their participation.

Materials—The stimulus materials were the same as Experiment 4.

Design and procedure—The design was the same as Experiment 4, crossing photo type (old, new) and noise level (clear, moderate or high-moderate). The subjective clarity data were also analyzed as a function of recognition accuracy. The study and distracter procedures were the same as Experiment 4, including the counterbalancing of photos across participants. In the test phase, old and new photos were shown with either 0, 20, or 40%

visual noise. However, after each photo, a 5-point scale appeared for subjective noise judgments. The scale was shown on the computer, ranging from clear (1) to noisy (5), and remained visible until a response was entered. The same scale was marked on the computer keyboard. Following noise judgments, participants were prompted to indicate whether the photo was old or new, using marked keys.

Results and discussion

One participant was removed prior to analysis due to failure to follow instructions. In the noise judgments, a main effect of noise level verified that judgments followed actual noise levels. Judged noise levels to the 0, 20, and 40% conditions were 1.31, 3.09, and 4.25, respectively [$F(2, 48) = 798.3$, MSe = .225; $\eta^2 = .97$]. More importantly, a small but significant effect of photo type showed that old photos ($M = 2.83$) were judged clearer than new photos [$M = 2.96$; $F(1, 49) = 18.87$, MSe = .073; $\eta^2 = .28$]. This result suggests that prior experience facilitated perceptual processing. The noise level X photo type interaction was null [$F(2, 48) = .60$, ns].

We also analyzed noise judgments as a function of recognition accuracy. This analysis was modeled after Goldinger et al. (1999), who found that people experienced a greater sense of perceptual fluency when their memory judgments were correct. To avoid empty cells in the ANOVA matrix, we first collapsed across actual noise levels (an effect that was not in question), then removed 10 participants who made no errors in one condition of the remaining design. A main effect of accuracy showed that noise judgments were higher in trials leading to errors ($M = 3.10$) than in trials leading to correct recognition [$M = 2.72$; $F(1, 39) = 12{:}60$, MSe = .471; $\eta^2 = .24$]. However, this tendency did not vary across old and new photos; the photo type X accuracy interaction was null [$F(1, 39) = .44$, ns].

Participants’ noise judgments were influenced by prior experience, such that old photos seemed clearer than new photos. The effect was quite small, but was consistent across individuals. This supports a view that perception and memory are different expressions of a common processing system, and it demonstrates another parallel between face and word recognition. Interestingly, the effect was not dependent on participants correctly discriminating old and new faces (see also Goldinger et al., 1999). This observation is important because it suggests that demand characteristics did not motivate the lower noise ratings to old photos. Even when overt recognition judgments were erroneous, noise judgments provided indirect evidence of memory.

Experiment 7: Duration judgments

Experiment 6 suggested that memory facilitates face perception. When asked to directly judge clarity, participants apparently attributed memory-driven perceptual fluency to actual clarity. However, as noted above, the effect of photo type (old versus new) was quite small. Experiment 7 was therefore conducted to conceptually replicate the effect, extending it to the subjective perception of presentation duration. Participants made subjective “speed” judgments to old and new faces, all shown in the clear. Following Witherspoon and Allan’s (1985) findings with words, we expected face familiarity to enhance perceptual processing, creating a subjective impression of longer presentations.

Method

Participants—Fifty-nine students participated for partial course credit.

Materials—The clear photos from Experiment 1 were used in Experiment 7.

Design and procedure—Experiment 7 was a 2 × 3 within-subject design, crossing photo type (new, old) and presentation duration (short, moderate, and long). The duration estimate data were also analyzed as a function of recognition accuracy. The study and distracter procedures were the same as Experiment 4, including the counterbalancing of photo durations across participants. In the test phase, new and old photos were shown serially for one of three presentation durations: 1000, 1500, and 2000 ms. After each photo, participants were shown a 5-point scale, labeled from “slow” to “fast,” and made subjective judgments of presentation duration. Because duration judgments are relative, practice trials were given before the test, demonstrating the range of possible presentation durations. In this demonstration, two rounds of words (“one,” “two,” etc.) were presented for five presentation durations; the actual test durations formed the endpoints and midpoint of the range. In the test phase, participants made duration judgments, followed by recognition judgments.

Results and discussion

Results for two participants were eliminated due to failure to follow instructions. In the duration judgments, a main effect of actual duration was observed [$F(2, 55) = 386.9$, MSe = .197; $\eta^2 = .93$]; mean duration judgments were 1.95, 3.00, and 3.92 to the short, moderate, and long durations, respectively. Of greater interest, an effect of photo type showed that new photos ($M = 2.92$) evoked shorter duration estimates than old photos [$M = 3.01$; $F(1, 56) = 7.14$, MSe = .089; $\eta^2 = .11$]. Although this was another small effect, it further suggested that memory can affect subjective perception.

As in Experiment 6, we also analyzed duration judgments as a function of recognition accuracy. To do so, we first collapsed across actual durations. To eliminate any remaining zero cells, data from another 10 participants were removed. No main effect of accuracy was observed [$F(1, 46) = 2.28$, ns]. Overall, participants gave equivalent duration judgments, regardless of eventual recognition accuracy. However, there was a reliable interaction of accuracy X photo type [$F(1, 46) = 4.81$, MSe = .308 $\eta^2 = .10$]. In trials with correct recognition decisions (either hits or correct rejections), duration judgments were equivalent to old and new photos. However, when participants made recognition errors, actual old–new status affected duration estimates, with old photos (misses) leading to longer estimates than new photos (false-alarms). Once again, this suggests a memory dissociation: even when people incorrectly claimed that new photos were old (false-alarms), their subjective experience was that the photos appeared very briefly. When they incorrectly believed that old photos were new (misses), their subjective experience was that the photos appeared for longer periods.

Taken together, Experiments 6 and 7 complement recent word-memory studies, suggesting that face memory affects subjective perception. Old photos were consistently judged perceptually “easier,” whether clarity or duration judgments were elicited. An interesting facet to these results is an apparent disconnection between subjective perceptual judgments (which may constitute a form of implicit memory) and overt recognition. In both experiments, old photos evoked more favorable perceptual judgments, even when participants believed those photos were new. A similar pattern was observed by Goldinger et al. (1999) in noise judgments to spoken words.

General discussion

It has been frequently suggested that face recognition is "special." Neuropsychological data show that face recognition can be selectively impaired, implying brain areas distinct from those for general object perception (Farah et al., 1998). As mentioned earlier, similar neurological dissociations exist for faces and words, the materials used in previous fluency experiments (Cousins et al., 2000). In addition, primate studies have located brain cells that respond specifically to faces, and differentially to particular faces (Desimone, 1991). Regarding orientation, most objects are difficult to recognize upside down, but inversion makes faces dramatically harder to recognize (Valentine, 1988). These (and other) findings suggest that face processing is unique. It is therefore reasonable to speculate that face memory might behave differently than memory for other perceptual objects, such as words. In the present investigation, we assessed perceptual fluency effects on face recognition, and face memory effects on feelings of perceptual fluency.

Although faces differ from other stimuli in many regards, the present results generally resemble those from similar studies testing printed words, spoken words, and sentences. In Experiments 2–4, manipulations of photo clarity consistently increased "old" responses to clearly shown faces. Unlike prior studies with words (e.g., Whittlesea et al., 1990), explicitly warning participants to disregard clarity had no effect (Experiment 3B; see also Whittlesea & Williams, 1998, 2001b). In Experiment 5, visual noise was replaced by auditory noise. The noise was intended to increase cognitive workload, and it apparently did. However, no "fluency" effect was observed for quiet trials. Finally, in Experiments 6–7, the direction of effect was reversed, such that memory for faces created impressions of greater perceptual fluency.

Heuristics in perception and memory

According to Whittlesea and Leboe (2000), people use various heuristics across situations, but all fall into two general categories, the *information* and *quality-of-processing* heuristics. These refer to the "what" and "how" of mental processing—what information is activated when stimuli are engaged for some purpose, and how fluently such activation occurs. Both are related to the experience of familiarity. The information heuristic uses prior mental content (stimulus names or images) to help identify stimuli. The quality-of-processing heuristic is based on the speed, cohesiveness, and vividness of information retrieval—perceptual fluency falls under this domain (Jacoby & Dallas, 1981).

In the present study, fluency effects were revealed by creating conflicts between perceptual and memorial cues. When face clarity was relatively high, it often induced (both appropriate and often inappropriate) "old" responses. In our experiments, all photographs depicted anonymous young men, rather than familiar people (e.g., celebrities). As such, we provided participants little opportunity to apply elaborative cues, and likely increased their reliance on perceptual information (Jacoby & Dallas, 1981; Mandler, 1980). Jacoby and Witherspoon (1982) hypothesized that recognition memory for familiar stimuli (e.g., words) is less reliant on perceptual cues, relative to unfamiliar stimuli (e.g., nonwords). Words can benefit from elaboration, but nonwords require more bottom–up familiarity—a feeling of “having seen this before.” This reliance on perceptual information makes nonwords, and perhaps novel faces, especially susceptible to perceptual manipulations.

In Experiments 2–4, manipulations of visual clarity were intended to assess perceptual fluency effects in recognition memory. However, the data suggest that fluency cannot be the entire story. Consider Experiments 3A, 3B, and 4, which all included three noise levels. By the most direct application of a fluency hypothesis (i.e., clarity equals familiarity), one would expect a liberal criterion to clear faces, a slightly more conservative criterion to moderately obscured faces, and an even more conservative criterion to the most obscured faces. Indeed, before conducting our experiments, we anticipated such a result. However, our results did not follow this pattern. Instead, the data were more compatible with a *fluency-attribution* process: As shown in Fig. 2, increasing noise levels created predictable, continuous effects on sensitivity. However, criterion settings were more binary: Clear photos induced an “old” bias, and all noisy photos were treated equivalently. This suggests that perceptual clarity triggers an attribution process (Jacoby et al., 1989, 1989), in the form of an “all-or-none” criterion setting.

Note that some aspects of our results are compatible with encoding specificity (Tulving & Thompson, 1973), although it does not easily predict the full results without elaboration. In our designs, clear photos were shown during study. When clear photos appeared at test, they surely conferred perceptual benefits, but they also provided the greatest overlap with encoding conditions. As such, people may have set their criteria based on processing overlap, rather than perceptual clarity. Note that, although this hypothesis derives from encoding specificity, it remains an attribution process. By the most straightforward application of encoding specificity, a natural prediction emerges for Experiments 2–4: when clear faces are shown during test, overall performance should improve, meaning more hits and fewer false-alarms. Referring to Fig. 2, it is apparent that sensitivity was superior in the clear conditions. While this is consistent with encoding specificity, the observed bias shifts are not. In a somewhat similar design, Busey et al., 2000 found that manipulations of face luminance created bias shifts, expressed as “old–new” confidence ratings. They also noted that their accuracy data were consistent with encoding specificity, but an attribution process best explained the dissociation with confidence.

By way of comparison, an experiment by Sporer (1993) directly examined encoding specificity in face recognition. Sporer hypothesized that, when people memorize faces from photographs, any visible clothing may serve as contextual cues to benefit later recognition. In an incidental learning task, participants saw a series of faces, half shown with visible clothing, half with the clothing obscured. In a surprise recognition test, clothing was manipulated to create contextual matches and mismatches. In contrast to the present results (and Busey et al., 2000), Sporer found that contextual matches improved both hits and false-alarms—sensitivity increased, with no change in bias.

In the present study, clear faces—randomly interspersed among obscured faces—had an obvious perceptual advantage. Participants seemed to naturally mistake this for familiarity, despite forewarnings that clarity and familiarity were orthogonal. Although our experimental situation is obviously contrived, similar effects may arise in everyday contexts. For example, imagine scanning a crowded cafeteria for a friend who is already seated. In this situation, you may scan the entire room, implicitly expecting your friend to "pop-out" from the unfamiliar crowd. This bottom–up strategy will often succeed, especially if your friend helps by waving. However, it also creates an opening for false familiarity cues. While scanning the cafeteria, perhaps you make eye contact with someone. Although you would not mistake this person for your friend, a momentary "tingle" of familiarity may occur, perhaps giving rise to a false start. And the unexpected presence of a *different* familiar person (e.g., an acquaintance from work) would likely trigger a strong "hit" response.

Whittlesea and Leboe (2000) noted that fluency effects are contextually defined, that the *expectation* of fluency is a key component of familiarity illusions. As their discrepancy-attribution hypothesis states, ". . .feelings of familiarity are not based directly on the fluency of processing, but instead are based on the discrepancy between that fluency and the fluency that could normatively be expected for that item in that task and context (page 85)." Continuing the foregoing example, imagine scanning for the same friend, now at a company Christmas party. Because many people would be familiar, it seems unlikely that a casual acquaintance would still draw significant attention. Intuition suggests that, in this context, people naturally refine their scanning process, resisting general familiarity in favor of more specific features.

Testing the discrepancy-attribution hypothesis more formally, Whittlesea and Williams (2001a) designed three sets of items, familiar words (e.g., *daisy*), regular nonwords (easy to pronounce; e.g., *hension*), and irregular nonwords (difficult to pronounce; e.g., *stofwus*). After pronouncing half the stimuli at study, participants received a recognition test. Regular nonwords elicited far more false-alarms (37%) than real words (16%), which elicited more than irregular nonwords (9%). Whittlesea and Williams suggested that regular non-words incurred symmetric penalties, relative to both other stimulus types. From the bottom–up, their greater perceptual fluency makes them seem familiar, relative to irregular nonwords. From the top–down, their lack of semantic associations makes them harder to reject, relative to real words. Thus, new regular nonwords inappropriately benefit from the fluency heuristic, and the information heuristic cannot offset that effect. In Experiments 2–4 of the present investigation, new clear faces also apparently engaged the fluency heuristic, which could not be offset by prior associations.

Conclusions

In the present research, perceptual fluency affected face recognition decisions, and face memory affected judgments of perceptual quality. It has been argued that memory tasks, such as old–new recognition, represent an artificial use of human memory (e.g., Glenberg, 1997). However, a framework of memory attribution heuristics (Jacoby & Dallas, 1981; Whittlesea & Williams, 2001a, 2001b) views such challenging tasks in essentially the same manner as "normal" memory. The only differences arise in the *degrees* to which different processes are utilized. In recognition, some items will be clearly familiar, some clearly novel. However, many will fall in the “gray area,” forcing evaluation using whatever information is available. When pressed in this manner, our participants often evaluated perceptual fluency as familiarity.

Two of our experiments produced potentially surprising results. In Experiment 3B, we still observed fluency effects, despite explicitly cautioning the participants. As mentioned earlier, this was a null effect (in a manner of speaking), and therefore requires caution. However, presuming the noneffect is real, it challenges our characterization of the attribution process —given direct advice, participants should have discounted fluency in their decisions. Its persistent effect may reflect cognitive impenetrability in face perception (Farah et al., 1998). People may have special difficulty “correcting” their perceptual experience per instructions, creating the null effect. This hypothesis finds support from our other potentially surprising result. In Experiment 5, auditory noise had clear detrimental effects on performance, but created no “fluency” effect in quiet trials. Auditory noise should be easily attributed to an irrelevant source, having no special effect on face evaluation.

Following prior studies, the present investigation suggests that perception and memory are intimately connected: manipulations of either component changes the other, whether words or faces are evaluated. At first glance, such vulnerability to perceptual and memorial illusions may suggest a flaw in human cognition—heuristic processing seems unwise when survival may be at stake. However, Gigerenzer and Todd (1999) take a different view, arguing that heuristic processing makes human cognition extraordinarily adaptive. Similarly, Roediger and McDermott (2000) suggest that memory illusions are natural by-products of intelligent cognitive systems. The intimate connection of perception and memory allows people to build elaborate episodes from simple stimulus events.

References

Bartlett JC, Strater L, Fulton A. False recency and false fame of faces in young adulthood and old age. Memory & Cognition. 1991; 19:177–188. [PubMed: 2017041]

Begg I, Armour V. Repetition and the ring of truth: Biasing comments. Canadian Journal of Behavioural Science. 1991; 23:195–213.

Brown C, Lloyd-Jones TJ. Verbal overshadowing in a multiple face presentation paradigm: Effects of description instruction. Applied Cognitive Psychology. 2002; 16:873–885.

Bruce V, Young A. Understanding face recognition. British Journal of Psychology. 1986; 77:305–327. [PubMed: 3756376]

Busey TA, Tunnicliff J, Loftus GR, Loftus EF. Accounts of the confidence-accuracy relation in recognition memory. Psychonomic Bulletin & Review. 2000; 7:26–48. [PubMed: 10780019]

Cousins R, Hanley JR, Davies A, Turnbull C, Playfer J. Understanding memory for faces in Parkinson’s disease: The role of configural processing. Neuropsychologia. 2000; 38:837–847. [PubMed: 10689058]

Desimone R. Face-selective cells in the temporal cortex of monkeys. Journal of Cognitive Neuroscience. 1991; 3:1–8. [PubMed: 23964801]

Farah MJ, Klein KL, Levinson KL. Face perception and within-category discrimination in prosopagnosia. Neuropsychologia. 1995; 33:661–674. [PubMed: 7675159]

Farah MJ, Wilson KD, Drain M, Tanaka JN. What is “special” about face perception? Psychological Review. 1998; 105:482–498. [PubMed: 9697428]

Feenan K, Snodgrass JG. The effect of context on discrimination and bias in recognition memory for pictures and words. Memory & Cognition. 1990; 18:515–527. [PubMed: 2233264]

Fodor, JA. The modularity of mind. Cambridge, MA: IT Press; 1983.

Gigerenzer, G., Todd, PM. The ABC Research Group. Simple heuristics that make us smart. Oxford University Press; 1999.

Glenberg AM. What memory is for? Behavioral and Brain Sciences. 1997; 20:1–55. [PubMed: 10096994]

Goldinger SD, Kleider HM, Shelley E. The marriage of perception and memory: Creating two-way illusions with words and voices. Memory & Cognition. 1999; 27:328–338. [PubMed: 10226442]

Hintzman DL. "Schema abstraction" in a multiple-trace memory model. Psychological Review. 1986; 93:411–428.

Jacoby LL, Allan LG, Collins JC, Larwill LK. Memory influences subjective experience: Noise judgments. Journal of Experimental Psychology: Learning, Memory and Cognition. 1988; 14:240–247.

Jacoby LL, Dallas M. On the relationship between autobiographical memory and perceptual learning. Journal of Experimental Psychology: General. 1981; 110:306–340. [PubMed: 6457080]

Jacoby LL, Kelley CM. Unconscious influences of memory for a prior event. Personality and Social Psychology Bulletin. 1987; 13:314–336.

Jacoby, LL., Kelley, CM., Dwyan, J. Memory attributions. In: Roediger, HL., III, Craik, FIM., editors. Varieties of memory and consciousness: Essays in honour of Endel Tulving. Hillsdale, NJ: Erlbaum; 1989. p. 391-422.

Jacoby LL, Whitehouse K. An illusion of memory: False recognition influenced by unconscious perception. Journal of Experimental Psychology: General. 1989; 118:126–135.

Jacoby LL, Witherspoon D. Remembering without awareness. Canadian Journal of Psychology. 1982; 36:300–324.

Jacoby LL, Woloshyn V, Kelley CM. Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. Journal of Experimental Psychology: General. 1989; 118:115–125.

Johnson, MK. A multiple-entry, modular memory system. In: Bower, GH., editor. The psychology of learning and motivation. Vol. 17. New York: Academic Press; 1983. p. 81-123.

Johnson MH, Dziurawiec S, Ellis HD, Morton J. Newborn's preferential tracking of face-like stimuli and its subsequent decline. Cognition. 1991; 40:1–19. [PubMed: 1786670]

Kelley CM, Jacoby LL, Hollingshead A. Direct versus indirect tests of memory for source: Judgments of modality. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1989; 15:1101–1108.

Logan GD, Etherton JL. What is learned during automatization? The role of attention in constructing an instance. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1994; 20:1022–1050.

Lorayne, H., Lucas, J. The Memory Book. Ballantine Books; 1996.

Mandler G. Recognizing: The judgment of previous occurrence. Psychological Review. 1980; 87:252–271.

O'Toole AJ, Deffenbacher KA, Valentin D, Abdi H. Structural aspects of face recognition and the other-race effect. Memory & Cognition. 1994; 22:208–224. [PubMed: 8035697]

Pascalis O, Petit O, Kim JH, Campbell R. Picture perception in primates: The case of face perception. Cahiers de Psychologie Cognitive/Current Psychology of Cognition. 1999; 18:889–921.

Reber R, Winkielman P, Schwarz N. Effects of perceptual fluency on affective judgments. Psychological Science. 1998; 9:45–48.

Rhodes G. Looking at faces: First-order and second-order features as determinates of facial appearance. Perception. 1988; 17:43–63. [PubMed: 3205669]

Roediger HL III. Memory illusions. Journal of Memory and Language. 1996; 35:76–100.

Roediger HL, McDermott KB. Tricks of memory. Current directions in Psychological Science. 2000; 9:123–137.

Segerstrale, CU., Molnar, P. Nonverbal communication: Where nature meets culture. Hillsdale, NJ: Lawrence Erlbaum Associates; 1997.

Sergent J. An investigation into component and configural processes underlying face perception. British Journal of Psychology. 1984; 75:221–242. [PubMed: 6733396]

Sporer SL. Clothing as a contextual cue in facial recognition. German Journal of Psychology. 1993; 17:183–199.

Tulving E. Memory and consciousness. Canadian Psychologist. 1985; 21:1–12.

Tulving E, Thompson D. Encoding specificity and retrieval processes in episodic memory. Psychological Review. 1973; 80:352–373.

Valentine T. Upside-down faces: A review of the effects of inversion upon face recognition. British Journal of Psychology. 1988; 79:471–491. [PubMed: 3061544]

Whittlesea BWA. Illusions of familiarity. Journal of Experimental Psychology: Learning, Memory and Cognition. 1993; 19:1235–1253.

Whittlesea, BWA. Production, evaluation and preservation of experiences: Constructive processing in remembering and performance tasks. In: Medin, DL., editor. The Psychology of Learning and Motivation. Vol. 37. New York: Academic Press; 1997. p. 221-264.

Whittlesea BWA, Leboe J. The heuristic basis of remembering and classification: Fluency, generation and resemblance. Journal of Experimental Psychology: General. 2000; 129:84–106. [PubMed: 10756488]

Whittlesea BWA, Jacoby LL, Girard K. Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. Journal of Memory and Language. 1990; 29:716–732.

Whittlesea BWA, Williams L. Why do strangers feel familiar, but friends don't. The unexpected basis of feelings of familiarity. Acta Psychologica. 1998; 98:141–166. [PubMed: 9621828]

Whittlesea BWA, Williams L. The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. Journal of Experimental Psychology: Learning, Memory and Cognition. 2001a; 27:3–13.

Whittlesea BWA, Williams L. The discrepancy-attribution hypothesis: II. Expectation, uncertainty, surprise and feelings of familiarity. Journal of Experimental Psychology: Learning, Memory and Cognition. 2001b; 27:14–33.

Winograd E, Rivers-Bulkeley NT. Effects of changing context on remembering faces. Journal of Experimental Psychology: Human Learning and Memory. 1977; 3:397–405.

Witherspoon D, Allan LG. The effect of prior presentation on temporal judgments in a perceptual identification task. Memory & Cognition. 1985; 13:101–111. [PubMed: 4033413]

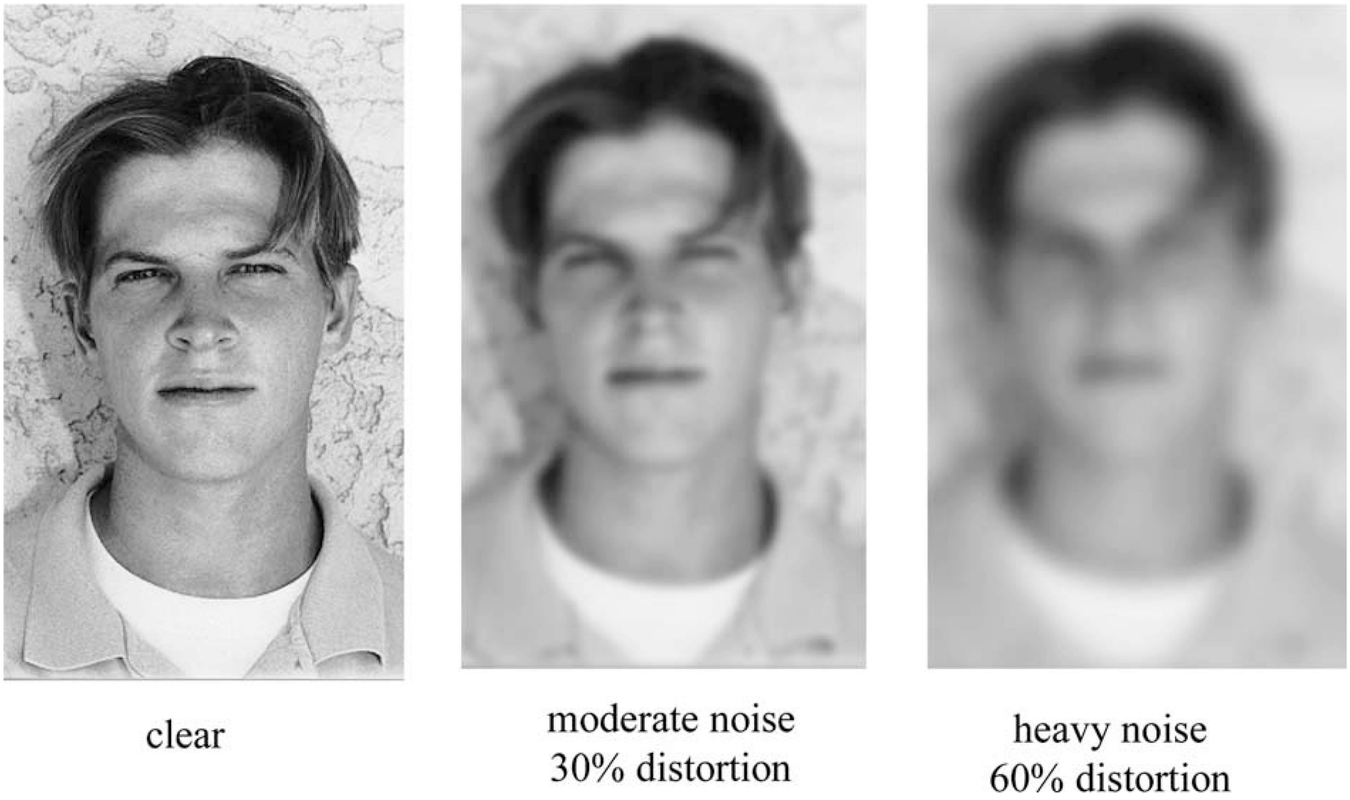

Fig. 1.

Examples of one stimulus photograph shown with 0% noise (clear), 30% noise, and 60% noise.

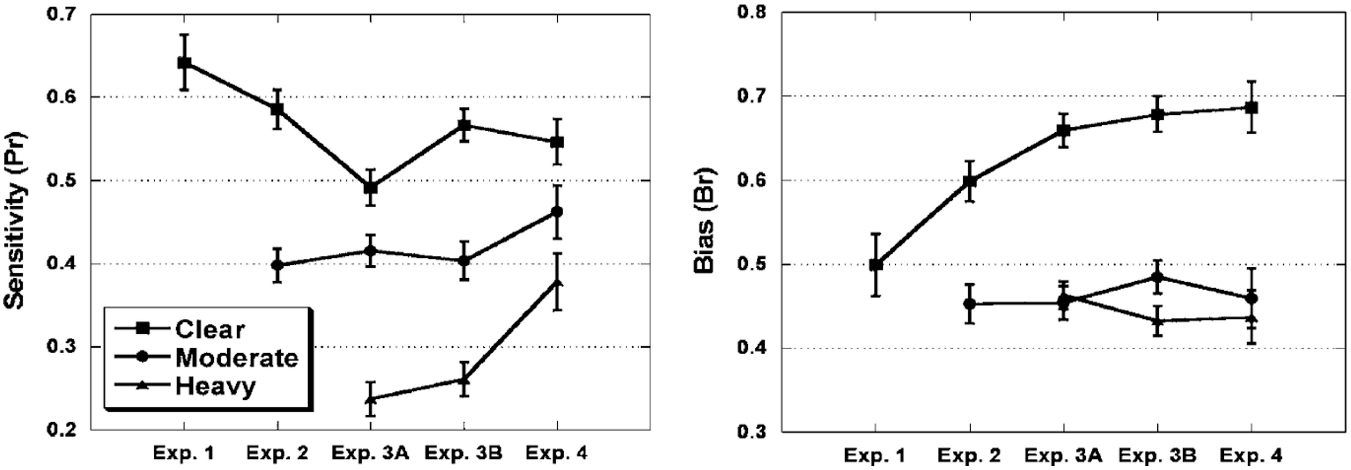

Fig. 2.

Signal-detection parameter estimates (±standard errors) from Experiments 1 through 4. Left and right panels show sensitivity (P_r) and bias (B_r), respectively. Increases in P_r reflect greater sensitivity. B_r is centered around a neutral value of .5; lower values represent a conservative bias and higher values represent a liberal bias.

Table 1

Miss and false-alarm rates (percentages, standard errors) from Experiments 1–5

	Noise level[*]		
	Clear	**Moderate noise**	**Heavy noise**
Experiment 1	(*N* = 47)		
Hits	83.61 (1.95)	–	–
False-Alarms	16.51 (2.08)	–	–
Experiment 2	(*N* = 93)		
Hits	84.68 (1.30)	67.84 (1.71)	–
False-Alarms	24.12 (1.75)	26.03 (1.68)	–
Experiment 3A	(*N* = 213)		
Hits	84.58 (1.15)	69.33 (1.33)	59.86 (1.59)
False-Alarms	33.41 (1.69)	25.75 (1.36)	34.12 (1.52)
Experiment 3B	(*N* = 192)		
Hits	88.45 (1.01)	70.31 (1.53)	59.55 (1.59)
False-Alarms	29.78 (1.64)	27.95 (1.57)	31.42 (1.61)
Experiment 4	(*N* = 80)		
Hits	86.38 (1.71)	71.54 (2.55)	67.71 (2.26)
False-Alarms	29.77 (2.06)	23.33 (2.17)	27.88 (2.61)
Experiment 5	(*N* = 101)		
Hits	79.73 (1.95)	75.66 (2.31)	74.59 (1.91)
False-Alarms	25.05 (2.07)	22.68 (2.01)	25.46 (1.87)

[*] Please consult text for actual noise levels across experiments.