BRIEF REPORT

Judging the familiarity of strangers: does the context matter?

Samantha A. Deffler · Alan S. Brown · Elizabeth J. Marsh

© Psychonomic Society, Inc. 2014

Abstract Context affects face recognition, with people more likely to recognize an acquaintance when that person is encountered in an expected and familiar place. However, we demonstrate that a familiar context can also incorrectly lead to feeling that a stranger is known. More specifically, we asked whether a familiar place can increase the belief that a stranger has been encountered outside of the experimental context (e.g., in the news or in real life). Novel faces were paired with novel places, famous places (landmarks), or neutral (solid color) backgrounds, and participants rated the pre-experimental familiarity of each novel face. Across four experiments, participants misinterpreted the familiarity of the landmark backgrounds as evidence of knowing the faces outside of the experimental context. This effect only disappeared when participants failed to integrate the face with the place, judging that the two did not fit together. Our findings suggest that familiarity is not judged in isolation; rather, people are unable to completely disentangle the familiarity of the individual components of a scene.

Keywords False memory · Familiarity · Face recognition · Context

Imagine walking into your favorite café and noticing what appears to be a familiar face, only to realize you don’t actually know the person. What role did the familiar café play in this error? Would the same error have occurred in the new café across town, in which it would be equally plausible to encounter someone familiar? Or, does the familiarity of a place influence how familiar a person in that location seems? We know that context affects face recognition; many experiments have supported Mandler’s (1980) “butcher-on-the-bus” example, with recognition suffering when a person is encountered outside of their normal context (see Memon & Bruce, 1985, for a review). Such research highlights the idea that faces and their contexts are assessed together at test. However, in Mandler’s example a mismatched contextual cue degrades recognition, whereas we examine how context may inappropriately increase perceived familiarity.

Jacoby’s attributional account of memory allows predictions about what will happen when a stranger is encountered in your favorite café. This framework assumes that feelings of familiarity need to be interpreted, often as evidence of “pastness” (Jacoby, Kelley, & Dywan, 1989). Usually this heuristic works, but it leads people astray when novel objects or experiences elicit feelings of familiarity for other reasons. For example, briefly glancing at a (nonstudied) word before deciding whether it is old or new increases the likelihood that one will incorrectly judge it as having been studied earlier in the experiment (Jacoby & Whitehouse, 1989). The literature is full of examples of people misinterpreting the fluency or ease of processing of words or symbols as evidence of familiarity (see Yonelinas, 2002, for a review).

There are two key differences between the example that we propose and the existing demonstrations of familiarity misattributions. First, our interest is in whether context can influence feelings of pre-experimental familiarity, in which a feeling of familiarity is misattributed to experience outside of the laboratory, as opposed to during an earlier study phase. These two errors are very different, with the threshold for what makes something “familiar” being expected to increase when “familiar” is defined as having encountered something outside the experimental context. Second, we ask whether people are able to disentangle the sources of familiarity in situations in

S. A. Deffler (✉) · E. J. Marsh
Department of Psychology and Neuroscience, Duke University, Durham, NC, USA
e-mail: sad26@duke.edu

A. S. Brown
Department of Psychology, Southern Methodist University, Dallas, TX, USA

Published online: 20 November 2014
which some things are familiar and others are unfamiliar. That is, most situations in life contain people, objects, and places, all of which vary in their familiarity. Can an observer ascertain that her feeling of familiarity is driven solely, for example, by a familiar setting, or can this feeling also be wrongly attributed to a stranger?

A familiar context sometimes influences the belief that a face has been seen earlier in an experiment. Consider a study by Gruppuso, Lindsay, and Masson (2007) that focused on the subjective experience accompanying memory for faces; critically for the present purposes, both old and new faces were tested in new versus previously seen contexts. Participants were more likely to incorrectly claim that a new face was “old” if it was tested with a studied context. This phenomenon has been noted in several studies (e.g., Davies & Milne, 1982; Rainis, 2001), albeit normally as an aside. Such misattributions to an earlier phase of the experiment are interesting, but this misattribution would be subjectively different and not cause the same nagging feeling of “Where do I know her from?” that likely occurs when one believes that someone is known from a time and place outside of the experiment.

Though they have rarely been explored, there have been several demonstrations of misinterpreting familiarity or fluency as pre-experimental experience. The best example is the false fame paradigm, where exposure to nonfamous names (e.g., Sebastian Weisdorf) increases later belief that the name refers to a famous person. However, the mechanism driving this effect is different from the one that we are interested in; false fame occurs when the experimental source is forgotten, time passes, or attention is divided (Jacoby, Kelley, Brown, & Jasechko, 1989; Jacoby, Woloshyn, & Kelley, 1989; see Brown & Marsh, 2008, for a similar demonstration with places). In contrast, our research does not involve forgetting of a recent experience (e.g., source amnesia), but instead explores how people misinterpret simultaneous familiarity signals stemming from the current experience. The two paradigms may yield the same misattributions of familiarity, but for different reasons.

What literature there is suggests that it is difficult to disentangle multiple simultaneous familiarity signals. People are unable to ignore the familiarity of an irrelevant stimulus flanking a target (Ste-Marie & Jacoby, 1993) or behind a target (Anderson, Jacoby, Thomas, & Balota, 2011). In one of the few studies using complex stimuli, participants were tested on their memories of line drawings of scenes; critically, the test included new scenes that differed in content but contained the same spatial configuration as the original scenes. Participants often misinterpreted the layout’s familiarity as past experience with the depicted scenes. For instance, having seen a plane approaching a runway increased the perceived familiarity of an unstudied but configurally similar picture of a dragonfly landing on a lily pad (Cleary, Ryals, & Nomi, 2009). Combined, these laboratory demonstrations suggest that the familiarity of part of a scene can be misattributed to another component of the same scene.

We conducted four experiments examining misattributions of the familiarity of one part of a scene to another, by asking whether participants would misattribute the familiarity of a setting to a stranger superimposed on it. Critically, we focused on pre-experimental familiarity. There was no study phase; each face only appeared once during the experiment when it was judged for its pre-experimental familiarity. To preview, participants were more likely to claim pre-experimental familiarity for strangers paired with famous backgrounds than for those paired with novel scenes (Exps. 1A and 1B). In Experiments 2 and 3, we investigated the mechanism underlying this effect.

### Experiments 1A and 1B

**Method**

**Participants** Thirty-seven Southern Methodist University undergraduates (SMU; Exp. 1A) and 31 Duke University undergraduates (Exp. 1B) participated.

**Materials** Black-and-white yearbook photos of strangers (the critical items) and celebrities were obtained through Internet searches. Each photo was posed face-forward, showing the person’s face and upper torso. Five undergraduate research assistants rated 132 celebrities and 156 strangers on a 6-point scale (1 = very unfamiliar, 6 = very familiar). Sixty celebrities ranging across the familiarity scale (range: 1.8–5) and 120 strangers rated as being low on baseline familiarity (M ≤ 2.8) were selected; half of the pictures depicted females. The celebrity faces (fillers) were interspersed to ensure that participants used the full range of the familiarity scale; these were not necessarily recognizable, given that many were decades old. The celebrity faces were not considered part of the experimental design.

The backgrounds consisted of color photographs of famous landmarks and novel scenes obtained through Internet searches, as well as a set of solid-color backgrounds. The same undergraduates rated the familiarity of each of 147 landmarks and 130 unfamiliar scenes on the same 6-point scale described above. Sixty landmark backgrounds rated as being high in familiarity (M ≥ 4.6) and 60 novel backgrounds rated as being low in familiarity (M ≤ 3.0) were selected. The 30 color backgrounds (used twice each per stimulus set) varied in saturation and hue.

The faces were divided into three sets, each containing 40 critical novel faces and 20 celebrity filler faces. These sets were matched on the mean familiarity of both the critical and celebrity faces and contained equal numbers of men and women. All participants saw all faces. Across participants,
each set was rotated through the three background conditions (see Fig. 1). Each face always appeared in the same location on the screen (out of 11 possible). To avoid obscuring the scene, a face was never centered on the screen. The stimuli are available from the first author.

Procedure The participants rated each of a series of faces for its real-world familiarity, defined as its familiarity prior to the experiment (seen previously in real life, in the news, or on TV). Familiarity was rated using a 1 (very unfamiliar) to 6 (very familiar) scale. It was stressed that the participants need not remember where they had seen a person, but should simply make a quick judgment of each person’s familiarity. To explain the presence of the different backgrounds, participants were told that the screen would change to help them prepare for the appearance of each face, and that they might be asked questions about the faces and the backgrounds later.

Participants completed three practice trials before completing 180 experimental trials. Each background appeared for 500 ms before the face photo was superimposed on it (see Fig. 1). The face and background appeared together for 1,500 ms. Immediately following the offset of the pictures, the rating scale appeared until the participant responded.

Results and discussion

As we predicted, participants rated novel faces paired with landmark backgrounds as being more familiar than those appearing on novel backgrounds. Table 1 shows the data for both experiments. Repeated measures analyses of variance (ANOVAs) showed that background type affected the rated familiarity of novel faces in both Experiments 1A, $F(2, 72) = 4.33, p = .017, \eta^2 = .107$, and 1B, $F(2, 60) = 3.34, p = .042, \eta^2 = .100$. Faces paired with landmarks were rated as being more familiar than those paired with novel backgrounds [1A: $t(36) = 2.47, p = .018, \eta^2 = .145$; 1B: $t(30) = 2.10, p = .044, \eta^2 = .129$]. The background effect was weaker when landmarks were compared to color backgrounds; this difference was marginally significant in Experiment 1A, $p = .081$, and not significant in Experiment 1B.

The data showed that context can affect one’s belief that strangers have been encountered prior to the experiment. It is not concerning that the effect was weaker with the color backgrounds; the pattern was similar, and the novel backgrounds are the more ecologically valid control. Participants failed to separate the familiarity of the person from the context; instead, the familiarity of one part of the scene bled over to another part. In Experiment 2, we tested whether participants could separate these two sources of familiarity when required to judge the face and the background separately. Relative to participants who only rated the familiarity of the face (single-rating), the question was whether participants directed to rate both the face and place (two-ratings) would better be able to separate the familiarity of the background from the face. To ensure that judgment order did not contaminate the results, we counterbalanced the order of the face and place ratings.

Fig. 1 Presentation procedure and sample pairings of a novel face with different backgrounds. In Experiments 1A and 1B, we used all three types of background; in Experiments 2 and 3, we used only the landmark and novel backgrounds.

![Fig. 1](image-url)
Experiment 2

Method

Participants A total of 77 students from Duke and SMU participated. Because no difference between schools had been found in Experiments 1A and 1B, students participated concurrently on both campuses. Forty-five of the participants rated just the faces (single-rating), and 55 rated both the faces and the backgrounds (two-ratings).

Materials The same materials were used as in the experiments above, except without the color backgrounds. Since all colors would be familiar to the participants, it was not logical to have participants judge the colors’ familiarity in the two-ratings condition. Because of the reduced number of backgrounds, faces were randomly selected for removal, reducing the number of faces to 120 (80 novel, 40 celebrity). Faces and backgrounds were paired in the same manner as before.

Procedure In the single-rating condition, participants only judged the familiarity of the faces. In the two-ratings condition, participants rated both the faces and the backgrounds for familiarity, with half of the participants rating the place first and the others rating the face first. On each trial, the 6-point familiarity scale appeared with a sentence indicating which item to rate.

Results and discussion

The main effect of the order of the familiarity judgments was not significant, $F(2, 97) < 1, p > .9$, nor was it involved in any significant interactions, $F$s $< 1, ps > .1$, so the following analyses are collapsed over order. The full data appear in Table 2 (where the second column shows the analyzed ratings). Did judging both the face and the place help participants to separate the sources of familiarity, as compared to participants who rated only the faces? A 2 (number of judgments: 1 or 2) × 2 (background: novel, famous) ANOVA on the average rated familiarity of the novel faces showed that faces paired with landmarks were rated as being more familiar ($M = 2.06, SE = 0.08$) than faces paired with novel backgrounds ($M = 1.91, SE = 0.07$), $F(1, 98) = 14.83, p < .001, \eta^2 = .131$. Making two judgments did not help participants to separate the two possible sources of familiarity. We observed no main effect of the number of judgments, nor an interaction of number of judgments with background, $ps > .05$. Finally, we note that the place ratings showed that two-ratings participants followed the instructions, rating landmarks ($M = 5.26, SE = 0.06$) as being more familiar than novel places ($M = 2.42, SE = 0.12$), $F(1, 54) = 607.83, p < .001, \eta^2 = .918.$

Across experiments, the data suggested that participants spontaneously separate the two familiarity signals, since the background effect was relatively small across studies (even though the familiarity signal from the famous backgrounds should have been very strong). To measure the extent to which the familiarity of the face and place were separated versus integrated, in Experiment 3 we created an orienting task to capture whether or not participants combined the two sources of the familiarity (for a similar logic, see Baddeley & Woodhead, 1982; Memon & Bruce, 1985). Participants in the experimental condition answered the question “Does this person fit in this place?” prior to judging face familiarity. A “yes” answer provided evidence that the face and place were integrated, whereas a “no” response was interpreted as a lack of integration. To ensure that nothing was special about “yes” versus “no” responses, we included a control in which participants also answered a question before judging face familiarity (“Is this person female?”). Critically, responses in the control condition should not be aligned with whether or not the face and place were integrated.

---

Table 1 Mean rated familiarity of novel faces—Experiments 1A and 1B

<table>
<thead>
<tr>
<th>Experiment</th>
<th>N</th>
<th>Color</th>
<th>Novel</th>
<th>Landmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A: SMU</td>
<td>37</td>
<td>1.97 (0.72)</td>
<td>1.91 (0.70)</td>
<td>2.11 (0.81)</td>
</tr>
<tr>
<td>1B: Duke</td>
<td>31</td>
<td>1.67 (0.59)</td>
<td>1.65 (0.57)</td>
<td>1.80 (0.72)</td>
</tr>
<tr>
<td>Both</td>
<td>68</td>
<td>1.83 (0.68)</td>
<td>1.80 (0.65)</td>
<td>1.97 (0.78)</td>
</tr>
</tbody>
</table>

Standard deviations are in parentheses. Ratings were made on a scale from 1 (very unfamiliar) to 6 (very familiar).

Table 2 Mean rated familiarity of novel faces by number of ratings—Experiment 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Novel</th>
<th>Landmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-rating</td>
<td>45</td>
<td>1.91 (.74)</td>
<td>2.07 (.91)</td>
</tr>
<tr>
<td>Two-ratings</td>
<td>55</td>
<td>1.91 (.65)</td>
<td>2.04 (.73)</td>
</tr>
<tr>
<td>Rate face first</td>
<td>27</td>
<td>1.90 (.70)</td>
<td>2.00 (.75)</td>
</tr>
<tr>
<td>Rate face second</td>
<td>28</td>
<td>1.91 (.61)</td>
<td>2.10 (.72)</td>
</tr>
</tbody>
</table>

Standard deviations are in parentheses. Ratings were made on a scale from 1 (very unfamiliar) to 6 (very familiar).
Experiment 3

Method

Participants Fifty-four Duke University undergraduates participated, with 27 per condition.

Materials The same materials were used as in Experiment 2.

Procedure Participants were divided into two conditions. In both, they first made a yes/no response while the face and background were visible, followed by the face familiarity rating. In the experimental condition, participants judged whether or not each face fit with the background, and in the control condition, participants judged whether or not each face depicted a female. These conditions resulted in a 2 (background: novel or landmark) × 2 (orienting task: experimental or control) design. Six practice trials were included to ensure that participants understood the procedure and made their yes/no responses while each face/place pair was still on the screen.

Results and discussion

The key question was whether the background effect depended on “yes” versus “no” answers in the experimental condition (in which those responses aligned with whether or not the face and place were integrated), but not in the control condition, in which participants judged gender. A 2 (background) × 2 (orienting task) × 2 (response) ANOVA was computed on the ratings of face familiarity. The data appear in Table 3. We found a significant three-way interaction between background (novel or landmark), orienting task (experimental or control), and response (“yes” or “no”), F(1, 52) = 7.35, p = .009, that precluded the other significant effects. To probe this interaction, we examined whether the 2 (background) × 2 (response) interaction was significant for each of the two orienting tasks (Table 3). Critically, in the experimental condition we found an interaction between background and participants’ answers to the orienting questions, F(1, 26) = 8.90, p = .006, η² = .208. As compared to novel backgrounds, landmarks increased the perceived familiarity of novel faces only when the face was rated as fitting with the place (a “yes” answer), t(26) = 2.81, p = .009, η² = .232, but not when the face and place were not integrated (a “no” answer), t(26) = 1.16, p = .256, η² = .024. Numerically, there was a background effect in the control condition, but it did not reach significance, F(1, 26) = 2.28, p = .143, η² = .078, perhaps because the “female” judgment distracted control subjects from the background and prompted some separation of the familiarity streams. Regardless, we found no interaction between background and answers to the orienting question in the control condition, F(1, 26) = 0.26, p = .618.

General discussion

In four experiments, strangers were judged as being more familiar if they were paired with landmarks rather than novel scenes. Familiarity was defined as pre-experimental familiarity, and judgments were made without prior experimental exposure to the faces. Given that there was no study phase in the experiment and no source to forget, the underlying mechanism could not be source amnesia (Schafer, Harbluk, & McLachlan, 1984; as in the false-fame effect: Jacoby, Woloshyn, et al. 1989). Rather, the results inform us more generally about how people make familiarity judgments. Familiarity is not judged in isolation—our participants were unable to judge the faces’ familiarity without being influenced, to some extent, by the context in which they appeared.

However, it is not that strangers in familiar places suddenly become as familiar as, say, family members. The background effect was relatively small in our initial experiments and in the single-rating condition of Experiment 2 (Exp. 1A: η² = .107; Exp. 1B: η² = .100; Exp. 2: η² = .094), thus making our replications particularly important. However, similar effects in episodic memory experiments are also small (Bower &

---

Table 3  Mean rated familiarity of novel faces by answer and orienting task—Experiment 3

<table>
<thead>
<tr>
<th>Orienting Task</th>
<th>Answer</th>
<th>Proportion Responses</th>
<th>Background Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Novel</td>
</tr>
<tr>
<td>Integration –</td>
<td>Yes</td>
<td>.36</td>
<td>2.12 (0.86)</td>
</tr>
<tr>
<td>“Does this person fit with this place?”</td>
<td>No</td>
<td>.64</td>
<td>1.92 (0.60)</td>
</tr>
<tr>
<td>(n = 27)</td>
<td>M</td>
<td></td>
<td>1.98 (0.66)</td>
</tr>
<tr>
<td>Control –</td>
<td>Yes</td>
<td>.47</td>
<td>1.91 (0.62)</td>
</tr>
<tr>
<td>“Is this person female?”</td>
<td>No</td>
<td>.53</td>
<td>2.01 (0.57)</td>
</tr>
<tr>
<td>(n = 27)</td>
<td>M</td>
<td></td>
<td>1.96 (0.57)</td>
</tr>
</tbody>
</table>

Standard deviations are in parentheses. Ratings were made on a scale from 1 (very unfamiliar) to 6 (very familiar).
Karlin, 1974), and even novel “unfamiliar” scenes likely contain familiar components (e.g., a mailbox or flagpole; see Fig. 1). Our focus on pre-experimental experience likely contributed to the small effects, since participants should require more evidence to claim that they have encountered someone prior to the experiment. Participants would have known that the experimenters would not have access to pictures of most of the people in their lives.

Furthermore, larger effects were observed in two cases: when people paid more attention to the backgrounds, and when we obtained a better measure of place and face integration. That is, a stronger effect was observed in Experiment 2 when participants explicitly rated the backgrounds as well as the faces (η² = .224); this task likely further drew attention to both components. Similarly, the effect was larger when participants integrated the face with its background in Experiment 3 (i.e., “yes” answers in the experimental condition; η² = .232). When the two were not integrated, the effect of the background was not significant, and the effect size was small (η² = .024). Critically, the effect of “fitting” the background was not due to a response bias associated with “yes” answers, since it did not matter whether participants responded “yes” or “no” in the control condition. This effect was also not due to a bias to claim that more-familiar faces fit with the backgrounds; regression analyses controlling for the baseline familiarity of the novel faces revealed the same pattern of significant results, and a chi-square analysis showed that participants were no more likely to judge the backgrounds as fitting celebrities than as fitting novel faces.2

Overall, the results show that people cannot completely separate the familiarity of a scene’s components; these signals may be combined into a single gestalt, although the components of a scene must be integrated for this bleed-over to occur (such as in Exp. 3, in which face familiarity was not influenced by the background when the two were not integrated). The idea that participants process some kind of combined familiarity signal, as opposed to parsing it to different sources, has a history in the episodic memory literature (“generalization of familiarity”; Bower & Karlin, 1974). However, it is unclear how far a familiarity signal may generalize, particularly in the real world. For example, could a single familiar object (e.g., the same table as at your house) affect the perceived familiarity of an entire scene? This idea has been used to explain déjà vu (Brown & Marsh, 2010). More generally, research from other domains suggests that people are often poor at parsing affective and other subjective signals into separate components, misattributing (for example) racial attitudes to an ambiguous pictograph (Payne, Cheng, Govorun, & Stewart, 2005) and athleticism from one person to another (Förderer & Unkelbach, 2011). The misattribution of familiarity to a stranger is likely an example of a larger class of illusions, whereby people look for cues in their environment in order to interpret ambiguous stimuli.

Author note We thank Meredith Mechanik, Katie Croft, and Lindy Loneman for help with the data collection.

References


---

2 We thank the reviewers for suggesting these analyses.


