Short Report

PERCEPTION

Perception
0(0) 1–10

DOI: 10.1177/0301006620946265
journals.sagepub.com/home/pec
SAGE

Face Familiarity and Image-Specific Memory

Robin S. S. Kramer and Chloe J. Berry
University of Lincoln, UK

Alex L Jones
Swansea University, UK

Georgina Gous
University of Lincoln, UK

Abstract

Face familiarity produces advantages for both memory and matching. By developing an internal representation through repeated experience, viewers extract identity-specific information that aids subsequent recognition. However, researchers have recently argued that this process may also result in a familiarity *disadvantage*, whereby specific instances of the face are more difficult to remember, perhaps due to this process of prioritising identity- over image-specific information. Although previous experiments found no evidence of this disadvantage in working memory, initial research has demonstrated an effect in longer term storage. Here, we attempted to replicate this finding by focussing on the ability to learn images of a single (un)familiar identity. Our results failed to demonstrate a familiarity disadvantage while replicating the finding that familiarity influences response bias. As researchers continue to investigate how familiarity alters both internal representations and associated processes, it is important to establish which processes may or may not be affected.

Keywords

face familiarity, face recognition, image memory, sensitivity, response bias

Date Received: 3 March 2020; accepted: 9 July 2020

Corresponding author:

Robin S. S. Kramer, School of Psychology, University of Lincoln, Lincoln LN6 7TS, UK.
Email: remarknibor@gmail.com

Introduction

The benefits of being familiar with a face are now well established within the literature. When asked to determine whether different photographs depict the same person or not, substantial difficulties arise when the faces are unfamiliar to the viewer (e.g., Burtone et al., 2010; Fysh & Bindemann, 2018; Kemp et al., 1997; Kramer et al., 2019) but the task becomes trivial for familiar viewers (e.g., Bruce et al., 2001; Megreya & Burton, 2007). This pattern of a familiarity advantage is also observed in tests of memory and recognition (e.g., Klatzky & Forrest, 1984; for a review, see Johnston & Edmonds, 2009).

Familiarity with a face results in an internal representation that appears to extract the identity-specific information (e.g., Burton et al., 2005, 2016). As such, image-level properties (e.g., quality, lighting, viewing angle, etc.) have little effect on matching and recognition performance for familiar faces (e.g., Burton et al., 1999). In contrast, given that viewers have no internal representation prior to their first experience with an unfamiliar face, their inherently limited representation is bound much more closely to the visual properties of the particular image(s) being viewed (Hancock et al., 2000; Megreya & Burton, 2006).

Given these representational differences are the result of differing levels of familiarity, researchers have recently identified the potential for a familiarity *disadvantage* (Armann et al., 2016). While increased familiarity produces benefits in recognition and matching (see earlier), it may be the case that remembering specific images is more difficult for familiar faces. One possibility is that the formation of an internal representation that prioritises identity-level information (allowing better generalisation and recognition in new encounters, etc.) may, as a result, sacrifice information regarding image- or instance-level properties. Indeed, storing idiosyncratic information about particular images could even harm later recognition if the details are common to several identities.

Recent work by Armann et al. (2016) has provided initial support for this disadvantage. In a memory task, participants were shown face photographs during a learning phase, and were subsequently tested on these, along with new photographs. Even when instructed to "remember the exact pictures" (p. 4) at learning, participants were less accurate when asked if they had seen the picture before with new images of previously seen familiar faces than for previously seen unfamiliar faces (where the correct response was "no" in all cases). It is worth noting, however, that for images which *had* previously been seen during learning (requiring a "yes" response), participants were more accurate with previously seen familiar, in comparison with unfamiliar, faces. The latter result appears to contradict the idea that image-specific memory in general is worse for familiar faces. A second issue, regarding the task itself, is that new images of previously seen unfamiliar faces were likely to be viewed by participants simply as unseen people. Recognising an unfamiliar face across images is error-prone (see earlier) and so this particular condition is difficult to interpret.

To further investigate the possibility of a face familiarity disadvantage, Dunn et al. (2019) carried out a series of experiments exploring the initial encoding of image-specific information in working memory (rather than the longer term memory tested in Armann et al., 2016). In all cases, their results showed no differences in performance for familiar versus unfamiliar faces. As such, the researchers concluded that, if any disadvantage existed, it was the result of longer term storage mechanisms rather than the process of encoding.

Given this lack of clarity as to the existence of a familiarity disadvantage in relation to image-specific memory, the aim of the current study was to investigate this issue using a task with comparable memory demands to those featured in Armann et al. (2016). In order to address the potential issues with interpreting their results (identified earlier), we chose to employ a simpler experimental design, which also allowed for the calculation of signal detection measures. Participants were shown multiple images of a single identity, who was previously familiar or unfamiliar to them. At test, if familiar participants were less able to identify which specific images they had or had not seen during learning, this would provide clear evidence of a disadvantage. The results of Armann et al. (2016) predict that familiar participants should perform worse when shown new, previously unseen images in particular when compared with unfamiliar participants. However, a lack of differences in performance due to our familiarity manipulation would argue against the existence of a familiarity disadvantage, in line with recent work by Dunn et al. (2019).

Methods

Participants

One hundred and forty-seven volunteers ($M_{\text{age}} = 24.8$ years, $SD_{\text{age}} = 9.4$ years; 59% women; 92% self-reported as White) gave informed, online consent before participating in the experiment and were debriefed onscreen upon completion. Recruitment took the form of sharing the experiment's weblink via email and social media, as well as approaching students and staff on campus and asking if they would be willing to participate.

The data from eight additional participants were excluded before analyses due to their level of familiarity with the model featured in our stimuli (see later).

The sample size for this experiment was informed by a priori calculations using GPower 3.1 software (Faul et al., 2007). The key finding in Armann et al. (2016) was produced through a simple main effects analysis, showing a significant advantage for unfamiliar over familiar faces in their "picture task." Their η^2_p effect sizes were 0.15 (Experiment 1) and 0.30 (Experiment 2), resulting in Cohen's f values of 0.38 and 0.62, respectively. Choosing an α of .05 and with power ($1–\beta$) set to 0.80, a total sample size of between 30 and 69 was required for our one-way between-subjects analyses of variance (ANOVA). However, we deliberately oversampled during data collection.

The experiment presented here was approved by the University of Lincoln's ethics committee (#846) and was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki.

Stimuli

We created a database of 44 colour photographs of a Dutch celebrity, Chantal Janzen, using Google Images searches. Pilot testing with a small sample of participants that did not take part in the experiment itself showed that this number of images resulted in a task of medium difficulty, with the use of fewer images risking ceiling effects that might mask any differences across conditions. The 44 images were cropped to contain only the head and neck and were resized to 380 × 570 pixels. Finally, all background colours were removed. Example stimuli can be seen in Figure 1.

Procedure

The experiment was carried out online using the Gorilla experiment builder (gorilla.sc). Participants were assigned to one of three experimental conditions based upon their current location. Those living in the Netherlands (98% self-reported as Dutch) participated in the "familiar—images" condition ($n = 50$), while participants living in the United Kingdom

Figure 1. Example Stimuli Used in the Experiment. Image attributions from left to right: Mariskad G (Own work; CC BY-SA 4.0), Rob van Hilten (Own work; CC BY-NC-SA 2.0), and Paul and Menno Ridderhof (Own work; CC BY-NC-SA 2.0).

(98% self-reported as British) were assigned at random to one of two conditions: "unfamiliar—images" ($n = 47$) and "unfamiliar—person" ($n = 50$).

In all conditions, participants first completed the learning phase of the experiment. Twenty-two images of Janzen (initially selected at random from the original set of 44 images but then kept constant for all participants, allowing for subsequent image-level comparisons) appeared onscreen on a white background, one after the other, with each presented for 1,500 ms and with an interstimulus interval of 500 ms (white screen). Presentation order was randomised for each participant. No responses were given during this phase.

Following this learning, participants completed the test phase, in which all 44 images of Janzen were presented in a random order for each participant. Participants were required to respond with either "old" or "new" using the mouse. Onscreen instructions prior to this phase explained that the former response referred to images seen during learning while the latter response referred to new images. Responses during this phase were self-paced.

The learning and test phases were identical across the three conditions with a few exceptions. First, the instructions provided prior to the learning phase differed across conditions. For the "familiar—images" and "unfamiliar—images" conditions, participants were instructed to "learn the images" that would follow. However, for the "unfamiliar—person" condition, participants were instructed to "learn the person." In all three conditions, participants were informed that the images to be learned were of the same person and that they would be tested afterwards (although the nature of the test was not specified). Second, for the "familiar—images" condition (for our sample living in the Netherlands), all instructions were displayed in both English and Dutch onscreen. These translations were developed and refined through a process of back-translation involving two native Dutch speakers.

Upon completion of the test phase, participants in all conditions were asked, "Before taking part in this experiment, how familiar were you with the woman in the photographs (Chantal Janzen)?" Responses were given using a 0 (*totally unfamiliar*) to 10 (*highly familiar*) scale.

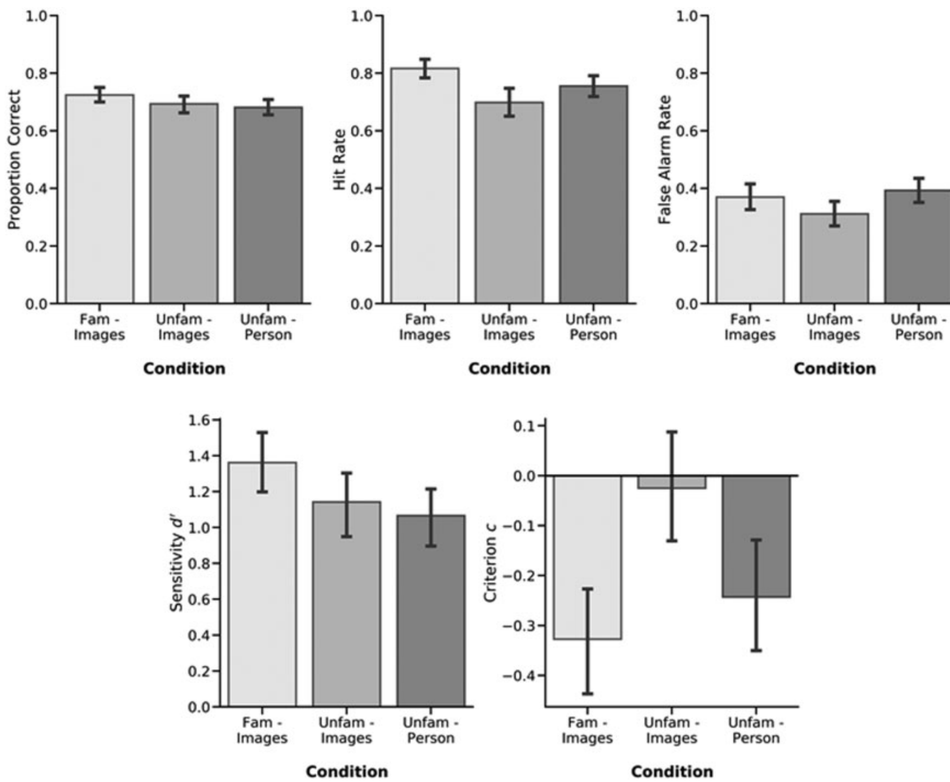

Figure 2. A Summary of the Results for the Three Conditions. Error bars represent 95% confidence intervals. Fam = familiar, Unfam = unfamiliar.

Results

Comparing Performance Across Conditions

We first considered participants' responses to the familiarity check. For those required to be familiar with Janzen ("familiar—images"; those living in the Netherlands), we excluded data from seven participants who rated their familiarity as 5 or less on our scale. The remaining participants' responses ranged from 6 to 10 ($M = 8.67$). For the participants who were required to be unfamiliar with Janzen (the two "unfamiliar" conditions), we excluded data from one participant who gave a rating of 5. The remaining participants' responses ranged from 0 to 2 ($M = 0.05$). As such, data from eight participants were excluded prior to the following analyses.

In order to determine whether image memory differed across conditions, we carried out several one-way between-subjects ANOVAs, with a summary of the data presented in Figure 2. It is interesting to note that this was a difficult task for participants, with low performance accuracies in all conditions. Here, signal detection measures were considered because we were interested in the extent to which representations in memory contained image-specific details (Dunn et al., 2019), and unlike Armann et al. (2016), the current design was well-suited to the calculation of these measures. As such, we calculated sensitivity

(d') using the following: *Hit*—participant responded “old” to an old image; *False alarm*—participant responded “old” to a new image.

Considering overall performance measures, we found no statistically significant differences between conditions when comparing both the proportion of correct responses, $F(2, 136) = 2.58$, $p = .080$, $\eta^2_p = 0.04$, and sensitivity d', $F(2, 136) = 2.92$, $p = .057$, $\eta^2_p = 0.04$. For hit rates, we found differences across conditions, $F(2, 136) = 7.97$, $p = .001$, $\eta^2_p = 0.11$, with “familiar—images” participants showing higher values than “unfamiliar—images” participants ($p < .001$; pairwise comparisons were Bonferroni corrected here and below). The remaining comparisons were not significant (all ps $> .119$). For false alarm rates, we also found differences across conditions, $F(2, 136) = 3.80$, $p = .025$, $\eta^2_p = 0.05$, with “unfamiliar—images” participants showing lower values than “unfamiliar—person” participants ($p = .024$). The remaining comparisons were not significant (all ps $> .196$).

Finally, for criterion c, we found differences across conditions, $F(2, 136) = 7.96$, $p = .001$, $\eta^2_p = 0.11$, with “unfamiliar—images” participants showing less of a tendency to respond “old” in comparison with both “unfamiliar—person” ($p = .015$) and “familiar—images” participants ($p = .001$). The remaining comparison was not significant ($p = .845$). As Figure 2 illustrates, participants in the “unfamiliar—images” condition showed a smaller response bias (where a value of 0 means no bias) in comparison with participants in the other two conditions.

Individual Differences in Our Familiar Sample

Although participants in our “familiar—images” condition were expected to be familiar with Janzen, our familiarity check included responses ranging from 0 to 10 ($M = 7.76$, $SD = 2.71$). Indeed, recent research has highlighted the nature of familiarity as a continuum of experience rather than a binary concept (Kramer et al., 2018). Therefore, within this sample alone, we were able to investigate whether familiarity was associated with image-specific memory.

Including data from only the 50 participants in this condition, we correlated familiarity ratings with the performance measures discussed earlier. We found significant relationships with hit rate, $r_s(48) = .30$, $p = .033$, and sensitivity d', $r_s(48) = .32$, $p = .023$, only (all other ps $> .097$). Reported significance levels were not corrected for multiple tests and so we suggest the need for further research to confirm these associations. However, the moderate, *positive* correlations were in the opposite direction to those expected if increasing familiarity with a face resulted in poorer image-specific memory.

Comparing Memory for Specific Images

As well as comparing participants’ performance measures across the three conditions, we considered performance at the image-level. As mentioned earlier, participants in all conditions were asked to learn the same subset of 22 images of Janzen, allowing us to investigate responses for these images, as well as the 22 “new” images, in each condition. Previous research has shown that photographs of different faces varied in their memorability (Bainbridge et al., 2013), and that “iconic” images of famous faces were better recognised (Carbon, 2008). Here, all images depicted the same person, allowing us to ask whether familiarity with a face influenced the specific images that people remembered.

Prior to analysis, data from the eight participants mentioned earlier were again excluded. For each condition, we calculated the proportion of correct responses given for each image across participants, with these values representing image-level difficulties. Next, we

Table 1. A Summary of the Relationships Between Image-Level Performances Across the Three Conditions.

Conditions to correlate		*Old* images	*New* images
Familiar—images	Unfamiliar—images	−.02	.75***
Familiar—images	Unfamiliar—person	.43*	.78***
Unfamiliar—images	Unfamiliar—person	.16	.74***

*$p < .05$. ***$p < .001$.

considered the correlations between these values for pairs of conditions, separately for "old" (learned) and "new" images. As Table 1 illustrates, we found a small (and not significant) relationship between image-level responses given by "unfamiliar—images" and "unfamiliar—person" participants to "old" images. In addition, we found a moderate (and significant, although not after correcting for multiple tests) relationship between image-level responses given by "familiar—images" and "unfamiliar—person" participants for "old" images. Finally, for all condition pairs, we found large associations between image-level responses to "new" images.

Discussion

Our results address the recent debate over the existence of a familiarity disadvantage for remembering specific images. Although Armann et al. (2016) provided initial evidence of this effect, a series of experiments by Dunn et al. (2019) found no detriment due to familiarity, at least with regard to working memory. Here, we also find no familiarity disadvantage in longer-term memory storage, comparable with those processes investigated in the original study (learning images followed by a test phase—Armann et al., 2016).

Although Armann et al. (2016) reported a familiarity disadvantage in image memory, their results showed that this was true only for correct rejections (i.e., responding that a "new" image was indeed new). In fact, performance was better for *familiar* faces with regard to hit rates (i.e., responding correctly to a previously seen image), and consideration of overall sensitivity (d') found that familiarity with faces resulted in better (Experiment 1—"remember all these people"), or at least not different (Experiment 2— "remember the exact pictures"), performance in comparison with unfamiliar faces. Taken together, we argue that this original pattern of results failed to provide compelling evidence of a familiarity disadvantage.

Interestingly, our pattern of results mirrored those of Armann et al. (2016). When instructed to learn the images, familiar participants showed significantly higher hit rates in both our work and theirs, as well as higher false alarm rates (although this difference was only significant in Armann et al.). Again, in line with Armann et al. (2016), we found numerically (although not statistically) superior performance in terms of sensitivity for those participants who were familiar with Janzen. Our results also demonstrated a shift in response bias (c), with familiar participants showing a greater bias to respond "old," explaining their increase in both hits rates (significant) and false alarms (nonsignificant) in comparison with those participants in the "unfamiliar—images" condition. This change in response bias as a result of familiarity was also evident in the data collected by Armann et al. (2016).

If increased familiarity with a face resulted in a decreased ability to remember specific images of that face (described as "poor coding of pictorial information"; Armann et al., 2016, p. 4) then this disadvantage should not be limited to false alarms. That familiarity led to better recognition of previously seen images (here and in Armann et al., 2016) is incompatible with the notion that these images were encoded less well. Here, overall performance (both proportion correct and sensitivity) was unaffected by familiarity, while response bias was significantly altered. This shift in bias (also evident in Armann et al., 2016, and Dunn et al., 2019) is capable of explaining the pattern of results described across studies—simply, familiar participants were more likely to think they had previously seen the test images. Given that no overall detriment was found for familiar participants, we are unable to interpret our results as evidence of a familiarity disadvantage.

Considering only our "familiar—images" condition, we found moderately sized, positive correlations between familiarity with Janzen and both hit rates and sensitivity. These patterns were in the opposite direction to those predicted by a familiarity disadvantage, given that here, our results suggested that increasing familiarity with a face may, in fact, produce better image memory.

Through analysing responses to specific images, we found that participants in the "familiar—images" and "unfamiliar—person" conditions showed a stronger association in their response patterns than the other condition pairings for "old" images. Although we are cautious to draw any strong conclusions without further investigation, this result is suggestive of the idea that instructing participants to learn the person, rather than the images, produced representations in memory and subsequent behaviours that more closely resembled familiarity with the identity. In addition, participants in all three conditions appeared to respond similarly to "new" images, suggesting that, irrespective of familiarity and learning instruction, specific images of Janzen were more likely to produce false positives than others. This result complements previous research showing that some faces are intrinsically more memorable (Bainbridge et al., 2013) and that particular, iconic images are better recognised for famous faces (Carbon, 2008). Given that our results applied to participants incorrectly "recognising" images they had not previously seen, we might hypothesise that certain "new" images were perhaps more visually similar to the learned images than others, although this idea requires further exploration.

Although the current task represents a conceptual replication of previous work (Armann et al., 2016) by considering longer-term image memory, there remain some differences worth noting. First, the nature of presenting multiple images of a single identity meant that no "familiar—person" condition could be explored. Simply, familiar participants could not be asked to "learn the person" from the images. Second, it is possible that learning a set of more visually similar images (i.e., all depicting the same face) alters the requirements of the task in comparison with learning images of different people. Although our data do not speak to this issue, it is clear that our participants were above chance levels in their performance and so were able to carry out this potentially more demanding task. In any case, we argue that a familiarity disadvantage should also extend to the current work if image-specific memory is influenced by familiarity, in that the same theoretical argument can be applied to the encoding of "one image of multiple identities" and "multiple images of one identity" designs.

In summary, recent research has provided mixed evidence regarding the idea that familiarity with a person may result in a disadvantage when tasked with remembering specific images of that person (Armann et al., 2016; Dunn et al., 2019). Results of the current study fail to evidence this disadvantage while replicating the finding that familiarity influences response bias (Armann et al., 2016). With researchers continuing to explore the nature of our mental representations as we become increasingly familiar with a face, it is important to understand how such representations and their related processes may (or may not) change.

Here, we find no evidence that familiarity with a face causes difficulties with remembering specific instances of that face.

Acknowledgements

The authors thank Bethany Patterson, Sarah Pinder, and Abigail Hill for their help with data collection. They also thank Abi Davis for comments on the manuscript as well as Nienke Muskee and Nienke Flederus for back-translation support.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Robin S. S. Kramer https://orcid.org/0000-0001-8339-8832

References

Armann, R. G., Jenkins, R., & Burton, A. M. (2016). A familiarity disadvantage for remembering specific images of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 571–580.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*, 1323–1334.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied*, *7*, 207–218.

Burton, A. M., Jenkins, R., Hancock, P. J., & White, D. (2005). Robust representations for face recognition: The power of averages. *Cognitive Psychology*, *51*, 256–284.

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science*, *40*, 202–223.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods*, *42*, 286–291.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science*, *10*, 243–248.

Carbon, C.-C. (2008). Famous faces as icons. The illusion of being an expert in the recognition of famous faces. *Perception*, *37*, 801–806.

Dunn, J. D., Ritchie, K. L., Kemp, R. I., & White, D. (2019). Familiarity does not inhibit image-specific encoding of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *45*, 841–854.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fysh, M. C., & Bindemann, M. (2018). The Kent face matching test. *British Journal of Psychology*, *109*, 219–231.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Sciences*, *4*, 330–337.

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, *17*, 577–596.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology*, *11*, 211–222.
Klatzky, R. L., & Forrest, F. H. (1984). Recognizing familiar and unfamiliar faces. *Memory & Cognition*, *12*, 60–70.
Kramer, R. S. S., Mohamed, S., & Hardy, S. C. (2019). Unfamiliar face matching with driving licence and passport photographs. *Perception*, *48*, 175–184.
Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition*, *172*, 46–58.
Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*, 865–876.
Megreya, A. M., & Burton, A. M. (2007). Hits and false positives in face matching: A familiarity-based dissociation. *Perception & Psychophysics*, *69*, 1175–1184.