

The perception of a familiar face is no more than the sum of its parts

Jason M. Gold · Jarrett D. Barker · Shawn Barr · Jennifer L. Bittner · Alexander Bratch · W. Drew Bromfield · Roy A. Goode · Mary Jones · Doori Lee · Aparna Srinath

© Psychonomic Society, Inc. 2014

Abstract Why do faces become easier to recognize with repeated exposure? Previous research has suggested that familiarity may induce a qualitative shift in visual processing from an independent analysis of individual facial features to analysis that includes information about the relationships among features (Farah, Wilson, Drain, & Tanaka *Psychological Review*, 105, 482–498, 1998; Maurer, Grand, & Mondloch *Trends in Cognitive Science*, 6, 255–260, 2002). We tested this idea by using a “summation-at-threshold” technique (Gold, Mundy, & Tjan *Psychological Science*, 23, 427–434, 2012; Nandy & Tjan *Journal of Vision*, 8, 3.1–20, 2008), in which an observer’s ability to recognize each individual facial feature shown independently is used to predict their ability to recognize all of the features shown in combination. We find that, although people are better overall at recognizing familiar as opposed to unfamiliar faces, their ability to integrate information across features is similar for unfamiliar and highly familiar faces and is well predicted by their ability to recognize each of the facial features shown in isolation. These results are consistent with the idea that familiarity has a quantitative effect on the efficiency with which information is extracted from individual features, rather than a qualitative effect on the process by which features are combined.

Keywords Visual perception · Face perception and recognition · Bayesian modeling · Perceptual learning

Electronic supplementary material The online version of this article (doi:10.3758/s13423-014-0632-3) contains supplementary material, which is available to authorized users.

J. M. Gold (✉) · J. D. Barker · S. Barr · J. L. Bittner · A. Bratch · W. D. Bromfield · R. A. Goode · M. Jones · D. Lee · A. Srinath
Department of Psychological and Brain Sciences, Indiana University,
1101 East 10th Street, Bloomington, IN 47405, USA
e-mail: jgold@indiana.edu

The ability to accurately recognize the faces of others plays an essential role in human social interactions. As such, the psychological processes that underlie face perception have been the source of a great deal of scientific interest (Farah, et al., 1998; Johnston & Edmonds, 2009). Not surprisingly, many experiments have shown that our ability to identify faces significantly improves with repeated exposure and practice (Bi, Chen, Weng, He, & Fang, 2010; de Heering & Maurer, 2013; Dolan, et al., 1997; Gold, Bennett, & Sekuler, 1999b; Gold, Sekuler, & Bennett, 2004; Hussain, Sekuler, & Bennett, 2009, 2011). But what is the nature of the underlying psychological changes that take place as faces become more familiar to us?

Gold, Bennett, and Sekuler (1999b) and Gold, Sekuler, and Bennett (2004) addressed this question using a combination of external noise masking techniques. This approach allowed them to trace the effects of practice on two general classes of perceptual mechanisms: 1) the magnitude of internal noise, such as the random firing of neural mechanisms (Croner, Purpura, & Kaplan, 1993) and random fluctuations in decision criteria (Benjamin, Diaz, & Wee, 2009); and 2) internal signal strength, which is related to the efficiency of other non-noisy aspects of information processing, such as the tuning of receptive fields or templates (Lu & Doshier, 2004). They found that observers’ ability to identify a set of initially unfamiliar faces improved by a factor of 5 or more over the course of 6 days of training, and that these learning effects were attributable to increases in signal strength rather than reductions in internal noise.

Although Gold et al.’s (1999b, 2004) results place important constraints on the possible mechanisms that mediate the changes occurring as faces become more familiar, there are many different processes that can lead to overall increases in signal strength. In a follow-up experiment, Gold et al. (2004) explored whether changes in performance with familiarity could be accounted for by improvements in the tuning of a

linear template.¹ They again added noise to face stimuli in order to estimate the linear templates used by observers as they learned to discriminate between two initially unfamiliar faces. By correlating observers' estimated linear templates with those of an ideal discriminator, they found that the changes in template tuning that took place with training could generally account for the corresponding improvements in recognition performance.

The results of Gold et al.'s (1999b, 2004) experiments suggest that the changes that take place when we learn to recognize a new set of faces are largely due to purely linear changes in processing (e.g., adjustments to the relative weights assigned to different stimulus locations or features); however, there is some recent evidence suggesting that a second-order relation-based or "configural" process may be invoked when perceiving familiar faces but not unfamiliar faces (Lobmaier & Mast, 2007; McKone, Brewer, MacPherson, Rhodes, & Hayward, 2007; Megreya & Burton, 2006). For example, Megreya and Burton (2006) had observers perform a face-matching task with unfamiliar faces, and found that performance with upright and inverted versions of the faces was highly correlated, a relationship that is not found with familiar faces (Valentine, 1988). Because impaired performance with inverted in comparison with upright faces is often taken as a hallmark of "configural" processing (Farah, Tanaka, & Drain, 1995; Maurer et al., 2002), Megreya and Burton (2006) suggested that one interpretation of their results is that familiar faces are processed configurally, whereas unfamiliar faces are processed in a manner that is more akin to simple pattern (i.e., linear template) matching.

In a related experiment, Lobmaier and Mast (2007) had observers recognize unfamiliar and familiar faces that were either spatially scrambled or spatially blurred. They argued that spatially scrambling the features of a face should affect configural processing (which relies on the relationships among features) but not featural processing (which relies only on the features themselves, not their relationships), whereas spatially blurring a face should affect featural but not configural processing. They found that performance was better for scrambled than for blurred faces when the faces were unfamiliar and that this difference was reversed for familiar faces. Along these lines, several additional studies have found that unfamiliar faces from a race other than one's own show less of an inversion effect than faces from one's own race (Megreya & Burton, 2007; Rhodes, Tan, Brake, & Taylor, 1989; Tanaka, Kiefer, & Bukach, 2004). Tanaka et al. (2004) found that other-race inversion effects are more pronounced for observers who have had relatively more experience

viewing faces of the other race; and McKone et al. (2007) found that a small amount of training with other-race faces yielded inversion effects for other-race and same-race faces that were equivalent.

Despite the results of these experiments, several other studies have found no differences between familiar and unfamiliar faces with respect to configural processing using a variety of measures, including inversion effects (Ellis, Shepherd, & Davies, 1979; Hole, 1994; Lobmaier & Mast, 2007; Schwaninger, Lobmaier, Wallraven, & Collishaw, 2009). Part of this confusion may be due to the use of widely varying tasks and stimuli across different experiments (Johnston & Edmonds, 2009). However, another, more fundamental issue is that the concept of "configural" processing has often been defined in purely descriptive terms (however, for several notable exceptions, see Amishav & Kimchi, 2010; Boremanse, Norcia, & Rossion, 2013; Richler, Gauthier, Wenger, & Palmeri, 2008; Sergent, 1984a, 1984b; and Wenger & Ingvalson, 2002). As a result, it can be difficult to interpret the results of experiments designed to test for configural properties.

Recently, Gold et al. (2012) introduced a new approach that allowed them to offer a mathematical definition of configurality and make quantitative predictions about its effects on performance. In their experiment, they had observers identify individual facial features shown in isolation (i.e., left eye only, right eye only, nose only, or mouth only) and in combination (i.e., all the features shown together). Six faces were used in the experiment, all of which were unfamiliar to the observers. They measured each observer's contrast sensitivity (i.e., the reciprocal of their RMS contrast threshold) in each of the four isolated feature conditions as well as for the combination of all four features. Using these sensitivities, it can be shown (Nandy & Tjan, 2008) that an observer's squared contrast sensitivity for the combination of features will equal the sum of her squared sensitivities across the individual features if the context within which the features are presented has no impact on her ability to make use of information. That is, if an observer's performance with the combination can be directly predicted from her performance with each of the features shown in isolation, then her squared sensitivity to the combination should equal the sum of her squared sensitivities to the isolated features. This prediction can be expressed mathematically by computing an integration index Φ :

$$\Phi = \frac{S_{combined}^2}{S_{left\ eye}^2 + S_{right\ eye}^2 + S_{mouth}^2 + S_{nose}^2} \quad (1)$$

where S is contrast sensitivity, and $\Phi = 1$. This is referred to as "optimal" integration, because it is the integration index that is predicted for a Bayesian observer who integrates across

¹ A linear template match refers to a purely first-order analysis of the stimulus. In the spatial domain, this typically amounts to a direct point-by-point comparison of the template and the stimulus. A nonlinear analysis involves higher-order comparisons, including second-order features such as the distances and correlations between spatial regions.

features in an optimal manner (Nandy & Tjan, 2008). If, however, an observer derives some additional benefit from viewing the features in combination that she did not derive when the features were shown in isolation (e.g., relational or “configural” processing), Φ should exceed 1. This is referred to as “superoptimal” integration, because it exceeds the integration index predicted by an optimal Bayesian integrator.² Alternatively, if there is a significant cost to processing all of the features when they are shown in combination that exceeds the cost of processing each of the features when they are shown in isolation, Φ should be less than 1. This is referred to as “suboptimal” integration, because it is less than the integration index predicted by an optimal Bayesian integrator.

In their experiments, Gold et al. (2012) found that the average integration index for upright faces was not significantly different from 1. Further, they found that the average integration index was significantly less than 1 for inverted faces. Thus, their results were inconsistent with the idea that observers were using qualitatively different strategies for upright and inverted faces (i.e., that they were using the relationships among features when viewing upright but not inverted features). Rather, their results indicated that the difference in performance between inverted and upright facial features was due to quantitative differences in information integration efficiency.

As mentioned above, the face stimuli used by Gold et al. (2012) were completely unfamiliar to all of the observers who participated in their experiments. Thus, it remains a possibility that the inability to find evidence for configural processing of faces in their experiments was due to their use of unfamiliar faces. That is, it is possible that learning and familiarity serve to promote the use of relational strategies, as some previous experiments have suggested (Ellis et al., 1979; Hole, 1994; Lobmaier & Mast, 2007; Schwanger et al., 2009). Given this possibility, the purpose of the present experiment was to explicitly test whether the results reported by Gold et al. (2012) are restricted to unfamiliar faces or if they extend to highly familiar faces.

We tested the above possibility by recruiting two groups of observers who were highly familiar with the faces of the members of their own group yet entirely unfamiliar with the faces of the members of the other group. We used the faces of the observers themselves to generate two separate sets of stimuli, and tested each group with both sets of faces. We also tested observers in three successive sessions with each set of face stimuli, to trace the shorter-term effects of training on integration efficiency. If familiarity serves to shift observers from using a parts-based to a relations-based strategy, we should find unfamiliar face features to be integrated optimally

and familiar face features to be integrated superoptimally. Further, we might also expect to find a systematic increase in integration efficiency for the unfamiliar sets of faces across training sessions.

Method

Observers

Two groups of seven participants served as observers in the experiment. The participants in both groups were approximately the same age (~21 years). Group 1 (four females, three males) were close friends, and had known each other for approximately 6 years; Group 2 (four males, three females) were also close friends, and had known each other for approximately 4 years. One participant from each group (J.D.B., male; M.J., female) was an author; the remaining six participants in each group were experimentally naive.

Stimuli

Stimuli were generated using the general approach described in Gold et al. (2012). The six naive participants from each group (three males and three females) also served as actors, and were photographed making neutral expressions. Only the six naive observers from each group were used as models, because the seventh participant in each group was also an experimenter, and thus was familiar with the face of the experimenter from the other group. Neither experimenter was familiar with any of the remaining six participants in the other group.

The face photographs were transformed into 256×256 pixel ($4.18^\circ \times 4.18^\circ$) grayscale images and normalized to range between 0 and 255. Next, the pixels in each image were converted into values of contrast by subtracting the mean value within the set of six images and then dividing by this mean value. Once the images were expressed in values of contrast, the four facial features (left eye, right eye, nose, and mouth) were isolated from each image by multiplying by four two-dimensional Gaussian windows, with each window centered at the general location where an individual feature appeared in the faces (see the [Supplementary Material](#) for a table with the dimensions and locations of the Gaussian windows applied to each feature in each face set). These dimensions were fixed for all faces within each set. The above procedure generated five groups of six images for each set of faces, shown in Figure 1 and Figure 2 (right eye only, left eye only, nose only, mouth only, and all features combined).

Unlike in Gold et al. (2012), pixel noise was not added to the stimuli during the experiment. This was done in order to maximize the range of contrast values that could be shown when measuring observers' contrast thresholds, and in particular to produce more reliable threshold estimates for

² Note that superoptimal integration is only possible for an observer who does not use all of the available information in one or more of the isolated-feature conditions. Superoptimal integration results when the combination allows the observer to make use of available information that she was unable to use when the feature(s) were shown in isolation.

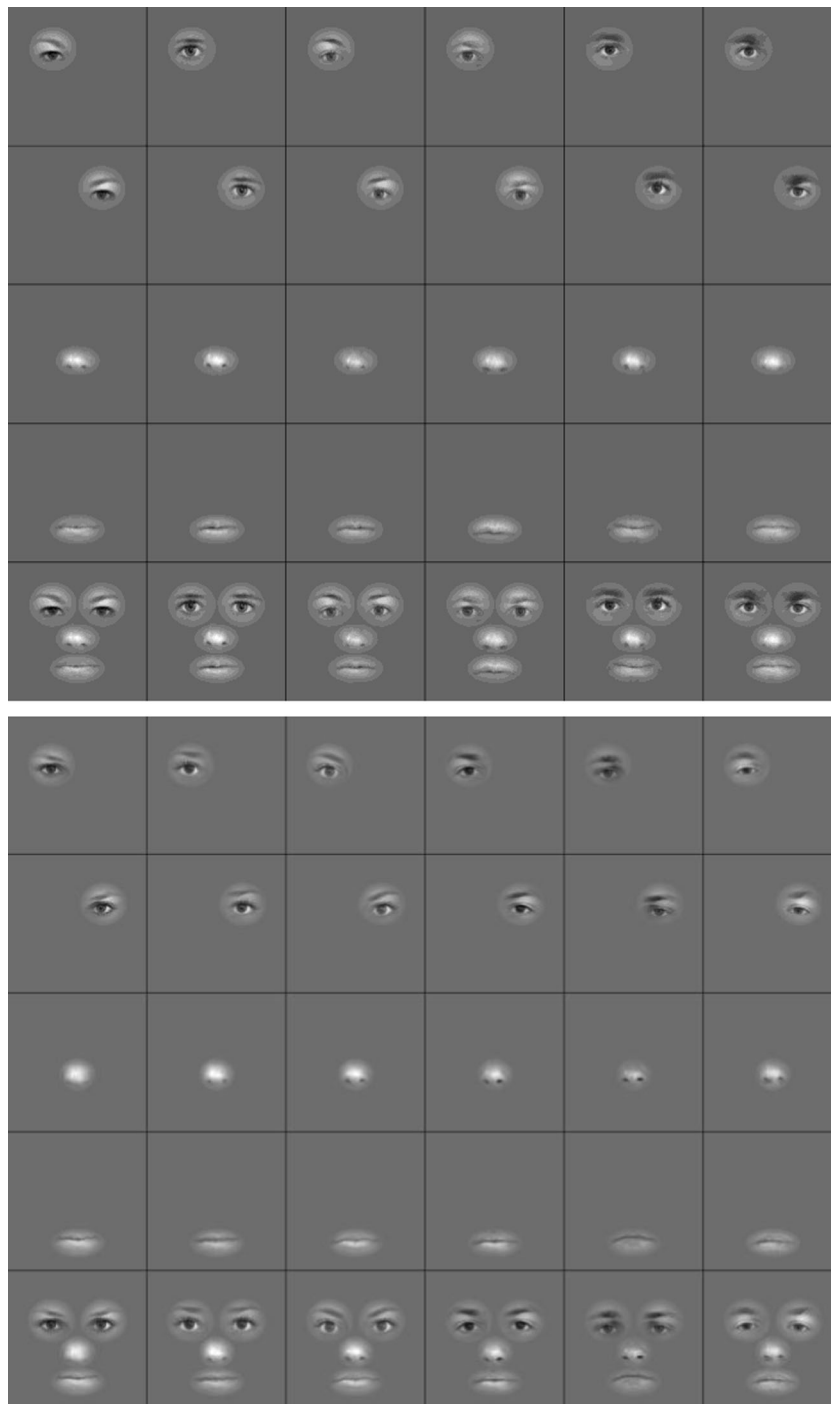


Fig. 1 The stimuli used in the experiments. The top panel shows the stimuli from Face Set 1 and the bottom panel shows the stimuli from Face Set 2. Within each panel, the columns correspond to individual participants/actors and the rows correspond to different feature conditions

individual features, such as the nose and mouth, that typically have relatively higher thresholds.

Apparatus

Observers were tested on an Apple iMac computer and the images were displayed on a Sony Trinitron G520 CRT (frame rate: 85 Hz; resolution: 1024×768 pixels, 38.25×28.25 cm).

The display was calibrated using a Minolta LS-100 photometer. Stimulus presentation was conducted within the MATLAB programming environment (version 2008b), using both in-house software and the extensions provided by the psychophysics toolbox (Brainard, 1997). The CRT provided the only source of illumination in the room. Viewing distance was binocular with natural pupils, and was fixed at 130 cm using a head/chin rest.

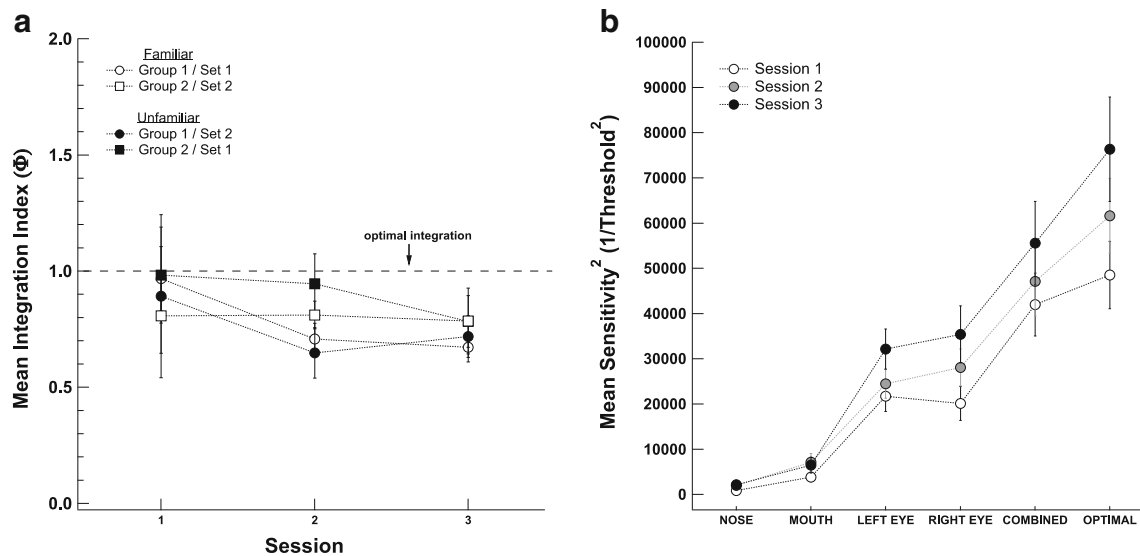


Fig. 2 A) Mean integration indices for each observer group and face set combination, plotted as a function of session. B) Mean sensitivities for each isolated feature, the combined features, and the predicted sensitivity

Threshold measurement

Thresholds and corresponding sensitivities were measured in each condition by varying the contrast of the images across trials using a 2-down 1-up adaptive staircase procedure. As described in Gold et al. (2012) and in further detail by Nandy and Tjan (2008), the contrast of an image on a given trial was set by adjusting the RMS contrast (i.e., square root of the mean across contrast values) of the combined image to the desired level and then removing all of the unwanted features (e.g., removing the left eye, nose, and mouth from the combined image in the right-eye-only condition).³ Weibull psychometric functions were fit to the staircase data in each condition in order to estimate 50 % correct RMS contrast thresholds. These contrast thresholds were then converted into squared sensitivities (where sensitivity is defined as 1/RMS contrast) in each condition.

Procedure

Observers were initially exposed to 2 min of dark adaption at the beginning of each experimental session. On each trial of the experiment, a thin box that surrounded the perimeter of the stimulus region appeared for an initial ~500 ms (43 frames), followed by the stimulus presentation for ~500 ms. Afterward, a selection window appeared that contained high-contrast versions of all six of the images that corresponded to the condition from which the image was drawn (e.g., if a right

³ This definition of contrast is called *nominal* contrast, and is necessary for the computation of the integration index given in Equation 1 (see Nandy & Tjan, 2008, for a formal treatment of the integration index and nominal contrast).

based on the sum of the isolated features for each session. Error bars on all symbols correspond to ± 1 SEM

eye was shown, all six right-eye-only images were shown). The observer made a selection by clicking on the corresponding image in the selection window. Accuracy feedback was given in the form of a high (correct) or low (incorrect) beep.

Design

Each observer completed three successive sessions on three separate days for each face set, with face set order counterbalanced across the six naive observers in each group. The seventh observer in Group 1 completed Face Set 1 first; the seventh observer in Group 2 completed Face Set 2 first. Within each session, thresholds were measured in all five conditions simultaneously, using five randomly interleaved adaptive staircases. Thus, observers did not know from trial to trial which condition they would be tested on next. Also, the identity of the face from which the stimulus was generated was chosen randomly on each trial, with uniform probability. There were 125 trials in each condition, for a total of 625 trials per session.

Results and discussion

Figure 2a plots the mean integration indices for each group with each set of faces, as a function of training session (see the [Supplementary Material](#) for all individual observer sensitivities and integration indices, broken down by group, face set, and testing session). In this plot, the dashed horizontal line corresponds to the index that would be predicted by optimal integration. The open symbols correspond to familiar faces and the closed symbols correspond to unfamiliar faces. These

data exhibit three striking patterns. First, none of the mean integration indices exceeded optimal integration, despite the extensive long-term familiarity observers had with one set of faces and the short-term training they received with both. Second, long-term familiarity did not appear to have a consistent effect on integration efficiency. And third, short-term training appears to have *reduced* integration efficiency rather than having increased it.

We tested the significance of these observations by carrying out a 2 (Group: 1, 2) \times 3 (Session: 1, 2, 3) \times 2 (Face Set: 1, 2) mixed, repeated-measures ANOVA. There were no significant main effects or interactions [Group: $F(1,11) = .59$, $p = .46$, $\eta_p^2 = .05$; Session: $F(2,10) = 2.06$, $p = .15$, $\eta_p^2 = .16$; Face Set: $F(1,11) = .27$, $p = .61$, $\eta_p^2 = .02$]. This also included a nonsignificant Face Set \times Group interaction [$F(1,11) = .13$, $p = .72$, $\eta_p^2 = .01$], indicating that familiarity had no significant effect on integration efficiency. We followed up on these initial analyses by conducting a series of one-sample t -tests (two-tailed) to determine whether the mean index for each group, face set, and session combination significantly differed from the prediction of an index of 1. After we used a Bonferroni correction for multiple comparisons, only one index significantly differed from optimal [Group 1, Set 1, Session 3: $t(6) = -5.15$, $p = 0.002$, Cohen's $d = 1.94$; see the [Supplementary Material](#) for a table with the complete t -test results for all combinations of group, set, and session].

We further explored the effects of short-term learning by computing the mean squared sensitivity across both groups and face sets for each isolated feature condition and the combined feature condition within each training session (Figure 2b). We also computed the squared sensitivity that would be predicted by optimal summation for each session (i.e., the summed sensitivity across the individual feature conditions; rightmost data points in Figure 2b). The increasing discrepancy between combined and optimal sensitivity across sessions highlights the negative effect short-term learning has on integration efficiency. But more importantly, these data reveal that there was very little improvement with the isolated noses and mouths, whereas the combined improvements with the isolated eyes actually exceeded the improvement when all the features were shown in combination. Thus, the decreases in the integration indices across sessions can be attributed almost exclusively to systematic increases in sensitivity to the eyes when they are shown in isolation as opposed to when they appear within the context of the other facial features. That is, the learning that takes place when observers are trained to recognize individual eye features shown in isolation does not appear to transfer fully to complete face recognition performance. These results are consistent with previous experiments that have found that the eyes receive disproportionate weight during recognition (Gold et al., 2004; Schyns, Bonnar, & Gosselin, 2002; Sekuler, Gaspar, Gold, & Bennett, 2004; Vinette, Gosselin, & Schyns, 2004) and that this disproportionate weighting becomes more exaggerated with

training (Vinette et al., 2004). Further, there is electrophysiological evidence that eyes produce larger cortical responses (as indexed by the magnitude of the N170) when presented in isolation than when shown in the context of a complete face (Bentin, Allison, Puce, Perez, & McCarthy, 1996).

So how might we reconcile our findings with the results of previous experiments that have shown evidence for configural processing of faces? One possibility may lie in the fact that we have explicitly restricted our definition of configural processing to those cases for which an observer benefits from viewing all features in combination (i.e., when the integration index is greater than 1). However, a more general definition of configurality might also include cases in which there are costs associated with viewing all features in combination (i.e., when the integration index is less than 1). There are a few reasons why we have chosen to adopt the more restrictive definition of configural processing as superoptimal integration. First, this definition naturally maps on to the common intuition that configural processing supports rather than hinders an observer's ability to integrate information across the features within a face. Second, it is straightforward to interpret an integration index greater than 1: In such a case, processing all of the parts in combination allows the observer to make better use of information than when the parts are shown in isolation. From this, we can infer that there must be some process at work that allows the whole to be perceived better than would be predicted from the individual parts. On the other hand, interpreting an index that is less than or equal to 1 is more difficult. In the case of an index that is less than 1, we can conclude that there is a cost to presenting the parts of a face in combination. This may be for a multitude of reasons, such as lateral masking or limitations on the ability to distribute attention across all features simultaneously. However, it also remains a possibility that the beneficial effects of configural processing remain at work, but are obscured by the greater costs associated with other processes. Similarly, an integration index of 1 could result from a perfect balance between the benefits of configural processing and the costs of other processes (although this would be a decidedly unparsimonious account of the data). Thus, our definition of configurality is somewhat conservative, in that it only allows us to identify those cases in which the benefits associated with configural processing outweigh the costs associated with processing all of the features of a face simultaneously.

Another possible reason why we failed to find evidence for superoptimal integration in our experiments is that configural processing may be taking place at a more subordinate level than is involved with the features we have defined in our experiment. That is, our experiment was designed to reveal only whether the parts of the faces that we have defined as features (left eye, right eye, nose, and mouth) are processed more efficiently when shown in combination than would be expected from how they are processed in isolation. As a result,

any relational dependencies that might be taking place *within* each of the parts of the face that we have defined as a single feature (e.g., the relationship between an eyebrow and a pupil with an individual eye) would not be revealed by our summation-at-threshold technique.

In fact, there are several lines of evidence that are consistent with the idea that the individual regions we have chosen to define as features in our experiment might exhibit within-feature relational dependencies. First, individual facial features similar to those that we have defined in our experiments have been shown to exhibit inversion effects akin to those of complete faces (Rakover & Teucher, 1997; however, see Rhodes, Brake, & Atkinson, 1993). Second, recent results have shown that a part of the brain thought to respond selectively to faces (the fusiform face area or FFA) responds similarly to face parts shown in isolation and in combination (Arcurio, Gold, & James, 2012). And third, Moscovitch, Winocur, and Behrmann (1997) found that a patient suffering from a pronounced form of object agnosia (i.e., an inability to accurately recognize objects) nevertheless exhibited a normal ability to recognize whole faces as well as isolated facial features. Although the results of our experiment cannot rule out the possibility that configural processing is taking place within individual facial features, future experiments could easily explore this possibility by applying our summation-at-threshold approach to facial features defined at a more subordinate level.

Another aspect of our experiment worth considering is the possibility that the generalizability of our results might be limited by certain stimulus- and task-related factors, such as the use of relatively small sets of stimuli under near-threshold contrast conditions. With respect to the issue of stimulus set size, our experiment is certainly not alone in the use of relatively small sets of stimuli to explore face recognition (see, e.g., Farah et al., 1998; Frowd, Bruce, McIntyre, & Hancock, 2007; Gold et al., 1999a; Megreya & Burton, 2006; Schwaninger et al., 2009; Schyns et al., 2002; Tanaka & Farah, 1993). Nevertheless, it is entirely possible that our results might not extend to larger or unbounded sets of face stimuli. Our use of only six face images within a set was primarily motivated by two factors: 1) the difficulty of finding two groups of participants who were highly familiar with the faces of the other participants within their group, entirely unfamiliar with the participants in the other group, and willing to serve as participants in the experiment; and 2) the limited number of combined stimuli that can be simultaneously displayed within a selection window on a computer screen. Although both of these factors are somewhat flexible, they do place a relatively low practical limit on the number of stimuli that can be used with our summation-at-threshold technique in the context of face recognition.

With respect to the use of near-threshold contrast stimuli, there are several reasons to think that our results are not limited to low-contrast conditions. First, the recognition task

that observers were asked to perform was difficult enough under all conditions to place the stimulus contrast at a level that was significantly above detection threshold and thus highly visible. Second, previous experiments have shown that recognition thresholds for faces and objects increase linearly as a function of the contrast of an externally added noise (Gold et al., 2004; Pelli, Burns, Farell, & Moore-Page, 2006; Tjan, Braje, Legge, & Kersten, 1995). This linear relationship between threshold and external noise contrast is exactly what one would predict if observers were adopting a strategy that was independent of the contrast of the external stimulus (Pelli, 1990).

Conclusions

Our goal was to test explicitly whether the optimal feature integration reported by Gold et al. (2012) with unfamiliar faces extends to familiar faces. We found that highly familiar and completely unfamiliar faces yielded similar mean integration indices, and none exceeded what would be predicted by optimal integration. Further, we found that short-term training tended to *decrease* integration efficiency rather than increase it, and that this effect was almost exclusively a result of improvements in observers' ability to recognize eyes when shown in isolation as opposed to within the context of other facial features. Whether this surprising negative effect of short-term learning on integration efficiency is specific to our task and stimuli or is instead indicative of a more general property of face identification is an interesting and open question that is worth pursuing in future research. Regardless, our results support the conclusion that, as found previously with unfamiliar faces, the perception of a familiar face appears to be no more than the sum of its constituent parts.

Author note This research was supported by National Institute of Health Grant No. R01-EY019265 (to J.M.G.). J.M.G., J.D.B., and M.J. generated the stimuli. J.M.G. programmed the experiments and wrote the first draft of the manuscript. J.M.G. and S.B. analyzed the data. J.D.B., R.A.G., and M.J. recruited the naive participants and J.M.G. photographed their faces. All authors contributed to the idea and design of the experiments, the collection of the data, and the editing of the manuscript. We thank Mike Burton for suggesting that we explore the relationship between familiarity and integration efficiency. We also thank two anonymous reviewers for helpful comments and suggestions. The authors declare that they have no conflicts of interest with respect to their authorship or the publication of this article.

References

- Amishav, R., & Kimchi, R. (2010). Perceptual integrality of component and configural information in faces. *Psychonomic Bulletin & Review*, *17*, 743–748.

- Arcurio, L. R., Gold, J. M., & James, T. W. (2012). The response of face-selective cortex with single face parts and part combinations. *Neuropsychologia*, *50*, 2454–2459.
- Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychology Review*, *116*, 84–115.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.
- Bi, T., Chen, N., Weng, Q., He, D., & Fang, F. (2010). Learning to discriminate face views. *Journal of Neurophysiology*, *104*, 3305–3311.
- Boremanse, A., Norcia, A. M., & Rossion, B. (2013). An objective signature for visual binding of face parts in the human brain. *Journal of Vision*, *13*(11).
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Croner, L. J., Purpura, K., & Kaplan, E. (1993). Response variability in retinal ganglion cells of primates. *Proceedings of the National Academy of Sciences USA*, *90*, 8128–8130.
- de Heering, A., & Maurer, D. (2013). The effect of spatial frequency on perceptual learning of inverted faces. *Vision Research*, *86*, 107–114.
- Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, *389*(6651), 596–599.
- Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*, 431–439.
- Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 628–634.
- Farah, M. J., Wilson, K. D., Drain, H. M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, *105*, 482–498.
- Frowd, C., Bruce, V., McIntyre, A., & Hancock, P. (2007). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, *98*, 61–77.
- Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999a). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research*, *39*, 3537–3560.
- Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999b). Signal but not noise changes with perceptual learning. *Nature*, *402*, 176–178.
- Gold, J. M., Mundy, P. J., & Tjan, B. S. (2012). The perception of a face is no more than the sum of its parts. *Psychological Science*, *23*, 427–434.
- Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, *28*, 167–207.
- Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, *23*, 65–74.
- Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision Research*, *49*, 2273–2284.
- Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2011). Superior identification of familiar visual patterns a year after learning. *Psychological Science*, *22*, 724–730.
- Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, *17*, 577–596.
- Lobmaier, J. S., & Mast, F. W. (2007). Perception of novel faces: The parts have it! *Perception*, *36*, 1660–1673.
- Lu, Z. L., & Doshier, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision*, *4*, 44–56.
- Maurer, D., Grand, R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Science*, *6*, 255–260.
- McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G., & Hayward, W. G. (2007). Familiar other-race faces show normal holistic processing and are robust to perceptual stress. *Perception*, *36*, 224–248.
- 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.
- Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, *9*, 555–604.
- Nandy, A. S., & Tjan, B. S. (2008). Efficient integration across spatial frequencies for letter identification in foveal and peripheral vision. *Journal of Vision*, *8*(13:3), 1–20.
- Pelli, D. G. (1990). The quantum efficiency of vision. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 3–24). Cambridge: Cambridge University Press.
- Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646–4674.
- Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception & Psychophysics*, *59*, 752–761.
- Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What’s lost in inverted faces? *Cognition*, *47*, 25–57.
- Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, *80*, 313–331.
- Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 328–342.
- Schwaninger, A., Lobmaier, J. S., Wallraven, C., & Collishaw, S. (2009). Two routes to face perception: Evidence from psychophysics and computational modeling. *Cognitive Science*, *33*, 1413–1440.
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*, 402–409.
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*, 391–396.
- Sergent, J. (1984a). Configural processing of faces in the left and the right cerebral hemispheres. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 554–572.
- Sergent, J. (1984b). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, *75*, 221–242.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A*, *46*, 225–245.
- Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*, B1–B9.
- Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, *35*, 3053–3069.
- Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology*, *79*, 471–491.
- Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It’s in the eyes. *Cognitive Science*, *28*, 289–301.
- Wenger, M. J., & Ingvalson, E. M. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 872–892.