VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 1 (46-117)

Spatial Vision, Vol. , No. , pp. 1 – 17 (2007)

Also available online - www.brill.nl/sv

Portraits and perception: configural information in creating and recognizing face images

BENJAMIN J. BALAS * and PAWAN SINHA

Department of Brain and Cognitive Sciences, MIT, 77 Massachusetts Avenue, Room 46-4089, Cambridge, MA 02139, USA

Received **????**; accepted **????**

Abstract—Configural information has long been considered important for face recognition. However, traditional portraiture instruction encourages the artist to use a 'generic' configuration for faces rather than attempting to replicate precise feature positions. We examine this intriguing paradox with two tasks designed to test the extent to which configural information is incorporated into face representations. In Experiment 1, we use a simplified face production task to examine how accurately feature configuration can be incorporated in the generated likenesses. In Experiment 2, we ask if the 'portraits' created in Experiment 1 are discriminable from veridical images. The production and recognition results from these experiments show a consistent pattern. Subjects are quite poor at arranging facial features (eyes, nose and mouth) in their correct locations, and at distinguishing erroneous configurations from correct ones. This seeming insensitivity to configural relations is consistent with artists' practice of creating portraits based on a generic geometric template. Interestingly, the frame of reference artists implicitly use for this generic template — the external face contour — emerges as a significant modulator of performance in our experimental results. Production errors are reduced and recognition performance is enhanced in the presence of outer contours. We discuss the implications of these results for face recognition models, as well as some possible perceptual reasons why portraits are so difficult to create.

Keywords: Face recognition; configural information; portraiture.

1. INTRODUCTION

Portrait artists must implicitly know a great deal about face recognition. To create a convincing likeness, the artist must replicate enough relevant features to make the identity of the model obvious to the observer. However, despite decades of face recognition research, our scientific understanding of how faces are represented by the visual system is still fairly limited. It is unclear what aspects of faces humans

*To whom correspondence should be addressed. E-mail: bjbalas@mit.edu

use in order to be able to rapidly and accurately identify familiar faces under a wide range of viewing conditions. Computational and psychophysical studies of face recognition offer many different ways to address this issue, but do not capture the implicit understanding that portrait artist brings to bear when creating a subject's likeness on canvas.

It appears likely that in exploring what features are important for recognizing a face, it may be helpful to examine the instructions and advice given to budding portrait artists. Specifically, are there aspects of the face that artists are instructed to meticulously duplicate? Conversely, are there facial features that artists are told not to belabor? If so, how do these instructions relate to the perception of faces by non-artists?

Considering face recognition in the context of these questions may yield new insights regarding the nature of facial likeness. In the current study, we examine the role of face configuration in both the production and recognition of familiar faces. By 'configuration' we mean the relative placement of the eyes, nose and mouth for a given individual. These are often referred to as 'second-order' configural features, as opposed to the 'first-order' configuration all faces share of two eyes above the nose, and nose above mouth (Diamond and Carey, 1986). In what follows, we shall use the phrase 'configural features' as shorthand to refer to second-order configural relations, unless stated otherwise.

A substantial number of psychophysical tasks underscore the importance of configural features in human face perception (Haig, 1984; Hosie *et al.*, 1988; Young *et al.*, 1985). The relative spacing of facial features can influence personality judgments (Brunswik and Reiter, 1937), and it has been widely suggested that the face inversion effect (Yin, 1969) results from a failure to recover configural information from the inverted face (Bartlett and Searcy, 1993; Collishaw and Hole, 2002; Leder and Bruce, 2000; Lewis and Johnston, 1997; Rhodes *et al.*, 1993). Moreover, observers are extremely sensitive to small disruptions of configural features, especially in familiar faces (Bredart and Devue, 2006; Ge *et al.*, 2003; O'Donnell and Bruce, 2001). This sensitivity to subtle differences in face configuration may depend critically on early visual experience (Le Grand *et al.*, 2001; Mondloch *et al.*, 2002), as assessed by examining featural and configural discrimination abilities in observers who had congenital cataracts removed shortly after birth. Taken together, these results support a model of human face recognition in which the configuration of facial features plays a substantial role.

While there is much behavioral evidence that the configuration of facial features is important for recognition, computational investigations of such features are equivocal in supporting the idea that second-order relations contain much diagnostic information. It has been shown that although recognition can be carried out using purely 'geometric' features similar to our conception of configural information, performance with gray-scale templates that incorporate feature appearance is generally superior (Brunelli and Poggio, 1993). While differences in feature configuration across individuals often require computational systems to map a deformable template of features to each new image, the ultimate arrangement of features is not usually used (Wiskott *et al.*, 1997). Measurements carried out at fiducial points appear to carry more diagnostic information than the relative locations of these points.

The difference between the importance of facial configuration in behavioral tasks and computer vision systems presents an interesting conundrum. While psychophysicists have suggested that configural information is a key aspect of the facial identity signature, computational investigations indicate that this information can be dispensed with, with little loss of performance.

Confronted with this paradox, we turn to the art of portraiture for inspiration. To what extent are artists told to replicate configural relationships between facial features accurately? Does a good portrait depend critically on the precision with which these measurements are made and set down on canvas?

2. INSTRUCTIONS FOR CAPTURING A LIKENESS IN A PORTRAIT

In sharp contrast to the commonly assumed complexity of the task of drawing a portrait, the recommended steps for capturing a likeness are actually quite simple. Here is a synopsis of the instructions, culled from many different sources (Blake, 1981; Graves, 1974, 1984; Hogarth, 1965):

Step 1. Draw a vertical oval with the long axis approximately 1.5 times the short axis.

Step 2. Taper and shape the oval to create a jaw-line and sides of cheeks.

Step 3. Mark off the horizontal axis of the oval at three points to divide it into four equal pieces. The first and the third points serve as iris locations.

Step 4. Sketch in a nose whose length from top to tip is a third of the horizontal axis.

Step 5. Mark the mouth at a third of the distance from the nose tip to the chin.

Step 6. Draw the hair line and general shape of the hair by looking at the model. It is helpful to compare the areas of the hair and face while sketching in the hair shape.

Step 7. Sketch the neck while looking at the model to see where the neck intersects with the jaw line.

Step 8. Look at the model to see where the outer corners of the left and right eye are. Mark these and then divide the line joining these two points into three equal pieces. The outer two pieces constitute the extent of the eyes.

Step 9. Look at the model to determine the extent and height of the eye-brows.

Step 10. Mark the bulbous sides of the nose to have the same extent at the central third determined in step 8.

Step 11. Look at the model to compare the width of the mouth with the width of the lower nose, and also to determine the margins of the upper and lower lips.

Step 12. Refine the outer shape of the hair and the hair-line across the forehead.

Step 13. Refine the shape of the eyes.

Step 14. Refine the shape of the nose.

Step 15. Refine the shape of the mouth.

It is interesting to note that the placement of the basic features (Steps 1–5) is based on a generic geometric template. Additonally, the outer head contour plays a privileged role in the construction of the facial likeness, since it provides a crucial frame of reference for subsequent steps.

Given the instructions above, it appears that portrait artists are not instructed to make particular use of face configuration in their work, save for the obvious need to preserve first-order configuration and make sure second-order features are physically plausible. What then are we to make of the wide range of psychophysical tasks suggesting that configuration is deeply important to human face recognition? Surely the arrangement of features is encoded to some extent by human observers, as evidenced by subjects' ability to discriminate between faces that have subtle differences in inter-ocular distance or nose-to-mouth distances. That said, though it is clear that changes in configuration can be detected, we do not have a good estimate of the fidelity of configural encoding, or even the necessity of high-fidelity configural information for the purpose of recognition.

A natural way to address this issue is to determine how accurately observers can estimate the proper configuration of facial features for highly familiar individuals. If subjects can successfully determine the true distances and angles between facial features, we can conclude that configuration is strongly encoded within the human visual system. On the other hand, if performance is poor, we can conclude that the information concerning second-order relations is perhaps not being used. While one could attempt to measure this through the use of a discrimination task (original faces *vs.* altered faces, for example) we will instead employ a face production paradigm in this experiment. The use of a production paradigm offers several advantages. First, subjects are free to make errors in any number of ways. In many configural encoding studies, specific configural features (such as inter-ocular distance) are studied exhaustively, while others are completely ignored. This is necessary in the context of a controlled discrimination experiment, but substantially limits the range of behavior being studied. Second, we suggest that production tasks tap into recognition mechanisms in a particularly elegant manner. As opposed to asking subjects to verify the correctness of a given image, asking them to produce a correct image removes the need to limit presentation time. By allowing subjects unlimited time to create the most accurate image possible, our estimates of encoding error can be derived from a scenario in which the subjects' performance was not limited by artificial task constraints. Finally, given that so little research effort has been directed towards understanding the creation of facial likeness, this experiment is interesting in its own right. Though recognition is easy, drawing and painting are very hard for most people. The use of a production task may provide us with some insights as to what the difference is between being able to recognize a face well and being able to create a convincing likeness.

Most people are not artists, so we cannot simply ask all our subjects to draw portraits of the same set of familiar faces. Instead, we present our subjects with a simple 'production system' that is meant to isolate configural encoding and does not rely on artistic ability. We present subjects with a set of celebrity faces in which the two eyes, nose, and mouth can be freely translated around the image. Our subjects' task is to place these four features in the correct configuration, both with and without the aid of the head outline. This procedure allows us to measure the accuracy of feature placement in the presence or absence of 'external' information from the head contour. The external contour of the head is a very strong cue for recognition, as evidenced by the strength of the 'Clinton–Gore' illusion (Sinha and Poggio, 1996). Here we manipulate the availability of the contour to determine whether or not configural encoding depends upon or incorporates measurements defined by the relationship between internal and external features.

By examining these 'portraits' we hope to gain some insight into the nature of configural encoding in human observers. To help build a bridge between the domains of vision and art, we also compare the production abilities of one group of subjects to the recognition abilities of another. In our second task, we ask whether or not the veridical image of each celebrity face can be picked out from the portraits constructed by the participants in our production task. Given that our production task removes the need for special motor skills and other training required to be a competent portrait artist, we might expect that the ability to reproduce feature configuration may only be limited by the fidelity with which it can be perceived. In this case, veridical images should be very difficult to discriminate from the distracter images. However, if there are additional perceptual factors that make production harder than recognition, we would expect that selecting the correct face should be possible. Furthermore, just as we compare face production accuracy in the presence or absence of external features, we also measure recognition accuracy with and without the outline of the head. We suggest that these two experiments provide an example of how production and perception paradigms can be combined to provide deeper insights into a particular topic than either methodology alone. Our investigation into the relationship between face recognition and portraiture provides an intriguing test case for this approach, and demonstrates that studying the intersection between vision and art raises many interesting questions for further research.

3. METHODS

3.1. Experiment 1: Face production

3.1.1. Subjects. 10 subjects (4 men, 6 women) were recruited from the MIT community to participate in this task. Subject age ranged from 21–50 years. All

VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 6 (345-387)

subjects reported normal or corrected-to-normal acuity and were familiar with all of the celebrities used in the study.

3.1.2. Stimuli. Twelve images of celebrities were used in this experiment. The celebrities depicted were: Cher, Bill Clinton, Cindy Crawford, Billy Crystal, Harrison Ford, Mel Gibson, Anthony Hopkins, Jennifer Lopez, Julia Louis-Dreyfuss, Julia Roberts, Winona Ryder and Robin Williams. All images were frontal views of the celebrities faces and were presented at sizes ranging from approximately 6 degrees to 10 degrees of visual angle. Examples of the stimuli are displayed in Fig. 1.

For use in the feature placement task, each image was binarized and edited in Adobe Photoshop such that the eyes, nose, and mouth of each face could be easily segmented and the remaining interior of the face was uniform. Each facial feature was then segmented from the background and placed in a separate image layer so that it could be manipulated independently of the other features. Facial features were then removed from their proper locations in each face and placed along the bottom edge of the image (Fig. 2).

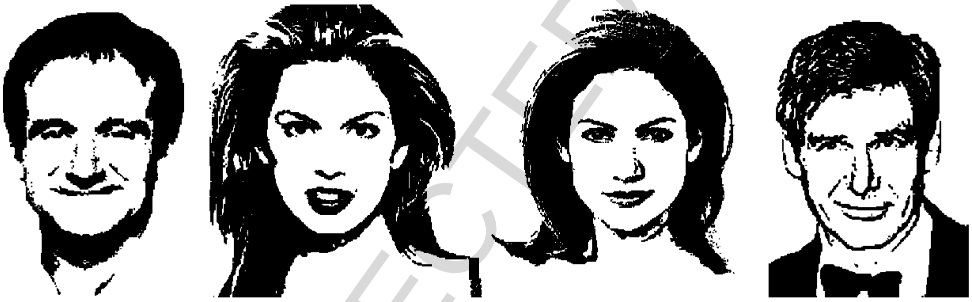

Figure 1. Examples of the original stimuli used as the basis for Experiments 1 and 2. From left, the celebrities depicted are: Robin Williams, Cindy Crawford, Jennifer Lopez and Harrison Ford.

Figure 2. An example of how the face and its features were presented to subjects in the 'outline present' condition. The face pictured here is Cindy Crawford's.

3.1.3. Procedure. Subjects were presented with the stimuli opened in Adobe Photoshop. They were shown that the name of each celebrity was listed in the title bar above each image and instructed to move the eyes, nose and mouth around on the blank background until they felt they had positioned them in the correct configuration. Subjects were instructed that they were to make the best likeness possible, first without the aid of the head shape and then with the head shape included. Observers were not shown any veridical images of the celebrities before carrying out this task, forcing them to rely on their memory for each celebrity's appearance. When the facial features had been placed to their satisfaction, subjects were asked to save their final composition.

Each subject first positioned the facial features of all twelve celebrities without the outline of the head, and then repeated this procedure with the outline included. While this introduces an ordering confound, we deemed it necessary to present the 'no outline' images before their counterparts. It would be difficult to completely forget the appearance of the head outline if it was presented first and subsequent visualization of its appearance could guide the placement of features even when it was absent.

3.2. Experiment 2: Face recognition

3.2.1. Subjects. An additional 24 subjects volunteered to participate in this task. Subject age ranged from 19–45 years. All subjects reported normal or corrected-to-normal acuity and were familiar with all of the celebrities used in the study.

3.2.2. Stimuli. The original images of each celebrity and the accompanying 10 versions of each face created by the participants in the 'outline included' condition in Experiment 1 were used in this task. Each celebrity face and its replicas were printed in an array on standard printer paper. Each image was approximately $2'' \times 1.5''$ in size, and the positions of the target face and the distracters were randomized on each page. The name of each celebrity was listed at the bottom of the page.

Twelve subjects were shown the facial features with the external outline of the head included, and an additional twelve subjects were shown the facial features with no external outline. We note that the distracter images in both of these cases came from the portraits created by subjects who had access to the external outline in Experiment 1. The only difference between conditions in Experiment 2, therefore, is the presence or absence of the head outline. There is no difference in the arrangement of the facial features. This is crucial if we are to meaningfully compare recognition ability for facial configuration subject to the availability of external features.

3.2.3. Procedure. Subjects were told that each page of their packet contained a veridical image of a celebrity along with 10 distracter images created by a previous

VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 8 (453-502)

group of subjects. They were then asked to circle the face on each page that they believed was the original or 'correct' face. As in Experiment 1, participants were not shown veridical images of the celebrities before viewing the test stimuli. Subjects were informed that target images and distracters could only differ in the arrangement of the eyes, nose and mouth. All participants were given unlimited time to complete the task. Typically, the task required 20 minutes to complete.

4. RESULTS

4.1. Experiment 1

All of our analyses concerning the placement of facial features are carried out in 'nose-coordinates' obtained by determining the vector connecting the centroid of the nose layer to the centroid of the left eye, right eye, or mouth layer. We characterize errors in feature placement by computing the vector difference between the correct location of each feature relative to the nose and the location selected by the subject. These vectors allow us to quantify errors in the relative placement of the eyes and mouth across subjects. In Fig. 3, we provide a schematic view of how error vectors are computed.

Given the set of error vectors for each celebrity face in the 'no outline' and 'outline' conditions, we ask two questions:

(1) Is there a significant difference in the *magnitude* of error vectors between conditions?

(2) Is there a significant difference in the *direction* of error vectors between conditions?

4.1.1. Error magnitudes. To determine whether or not the presence of the head outline affects error magnitude in our task, we compare the average error magnitude

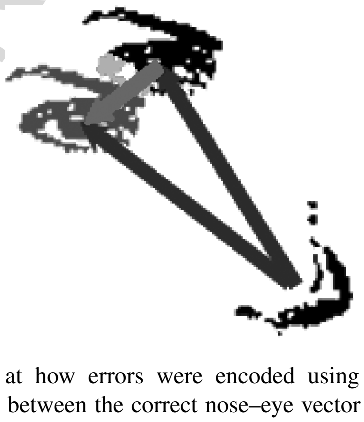

Figure 3. A graphical look at how errors were encoded using vector differences. The red vector represents the difference between the correct nose–eye vector (opaque eye) and the subject's estimate (transparent eye). We examine both the magnitude of these error vectors, and their angular distribution.

VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 9 (502-562)

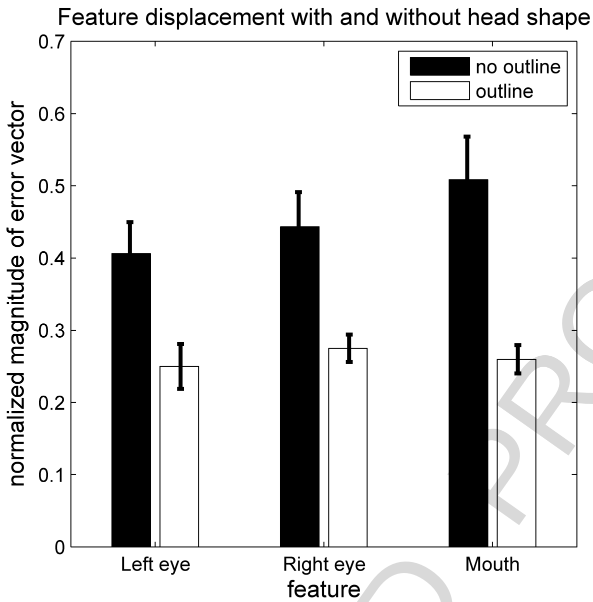

Figure 4. The mean magnitude of error vectors for each feature across subjects in the outline absent (blue bars) and outline present (red bars) tasks. Error bars represent ±1 S.E.M. There is no effect of feature, but a highly significant effect of external outline.

for each subject across items in both conditions. However, since our faces are different sizes we express the error magnitude as a percentage of the magnitude of the relevant 'ground truth' vector. In this way, error magnitude is normalized across images.

A graph of the average error magnitudes across features and outline conditions is displayed in Fig. 4. A two-way within-subjects ANOVA with feature (left eye, right eye and mouth) and outline (absent or present) as factors reveals a highly significant effect of outline ($F(1, 2) = 34.6$, $p < 0.0001$). No other main effects or interactions were significant.

4.1.2. Error directions. Next, we ask whether or not the direction in which subjects make placement errors is affected by the presence of the outline. In particular, we are interested in determining whether or not the absence of the outline induces an 'expansion' of features radially outwards from the nose.

Since we are interested only in the direction of errors in this analysis, the magnitude of each error vector is normalized to unit length. We continue by computing a circular mean for each subject across stimuli. This yields two distributions of mean directions for each feature, one for the 'no outline' condition and one for the 'outline' condition. The distributions of mean error vectors for each feature are displayed in Fig. 5.

To determine whether or not any of the distribution pairs for a given feature are different from each other we use a two-sample Kuiper's test (Batschelet, 1981;

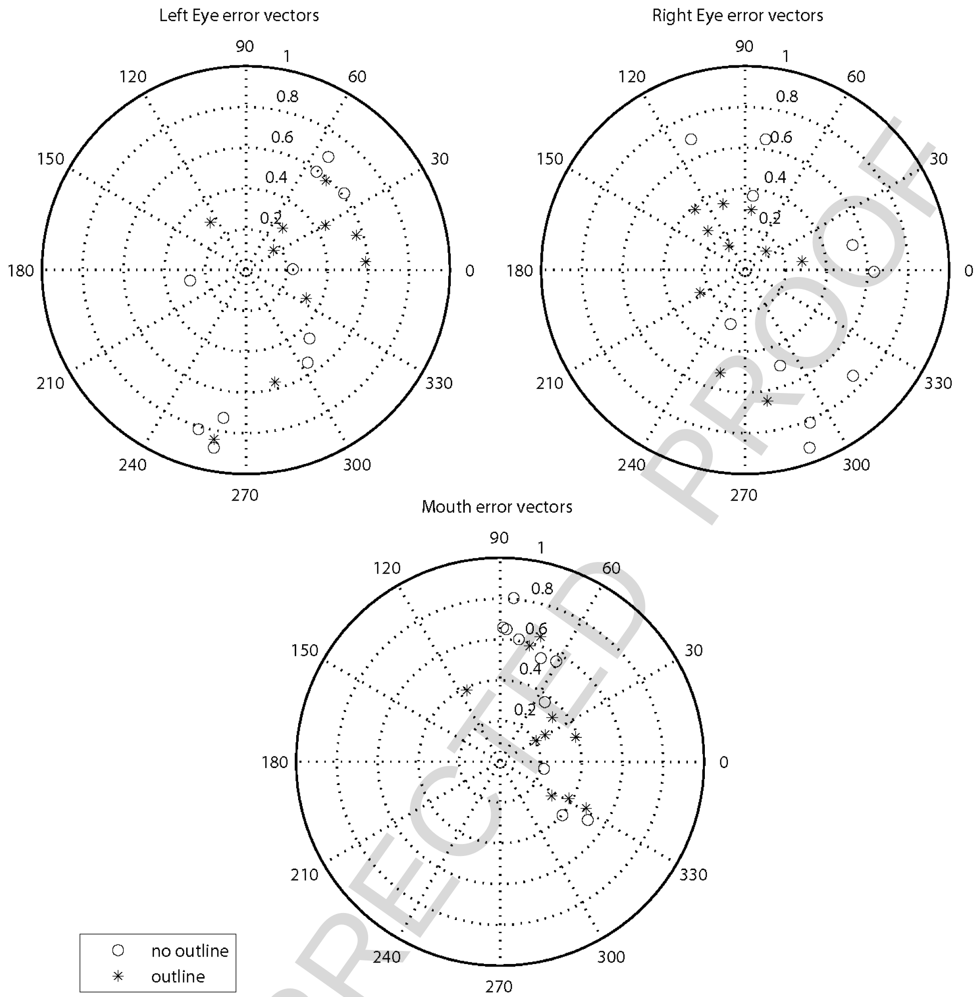

Figure 5. A summary of the angular distribution of error vectors for all three features. For each subject, the circular mean of error vectors (normalized to have unit magnitude but the original orientation) was taken across all faces in each condition. Here we display these mean vectors for all conditions and facial features. The distributions of angles in the outline present and outline absent conditions do not significantly differ.

Jammalamadaka, 2001). We find no effect of outline for any of our three features ($p > 0.5$ in each case). It appears that while the magnitude of errors decreases with the addition of the outline, the direction in which subjects make errors does not appreciably change. Thus, the outline is not simply serving to reduce 'expansion' of the internal feature assembly, but has a more generalized effect on error reduction.

VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 11 (578-641)

4.2. Experiment 2

In this task, we wish to know whether or not subjects can reliably pick out the original celebrity face from a full array of 'portraits' created by participants in Experiment 1. That is, were subjects in Experiment 1 able to create portraits that can reliably fool a new group of subjects? Furthermore, does recognition performance depend on the availability of external features? Given 11 images on each page, chance performance on this task is approximately 1/11, or ∼9%. If production ability is as fine-tuned as recognition ability, we expect that performance in this task will not differ significantly from this level. Also, if external features do not contribute to the recognition of face configuration, we expect that accuracy will not be affected by the presence or absence of the outline.

Furthermore, we can estimate the 'tolerance' of recognition by asking how big the errors in feature placement are for distracters erroneously selected by subjects in Experiment 2. Determining the magnitude of the error vectors for the 'false alarms' selected in both conditions provides a rough upper bound for acceptable configural distortions in face recognition.

4.2.1. Percent correct. Average performance on the recognition task was approximately 60% when the external outline was included and approximately 20% when it was absent. The 95% confidence interval of the 'outline-included' mean is [47.2%, 72.3%] and the interval for the 'outline-excluded' mean is [9.6%, 29.4%]. Neither of these intervals covers the chance level of performance, so we can conclude that subjects are able to perform the recognition task in both conditions. However, we can also see from these two intervals that subjects are significantly more accurate in the 'outline-included' condition. This leads us to conclude that external features provide important information for processing configural information in faces.

4.2.2. Average magnitude of error across targets. For each subject, we also determined the average magnitude of the error vectors across all face selected as a target. Correct selections are coded as zeroes. The result is an estimate for each feature of how much error was tolerated by subjects in both conditions of Experiment 2. A graph of these results is displayed in Fig. 6.

The low levels of error here indicate that subjects were correct a lot of the time, and also that any distracter images selected as targets did not differ greatly from veridicality. We conclude from this analysis that the 'tolerance' for error in recognition is less than that associated with production. To some extent, this is to be expected since we are essentially sampling from the production errors to obtain the recognition errors. As a result, we do not present a statistical comparison of error across the production and recognition tasks, but simply offer the results displayed in Fig. 6 as a rough upper bound of the fidelity of configural encoding for face recognition.

VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 12 (641-670)

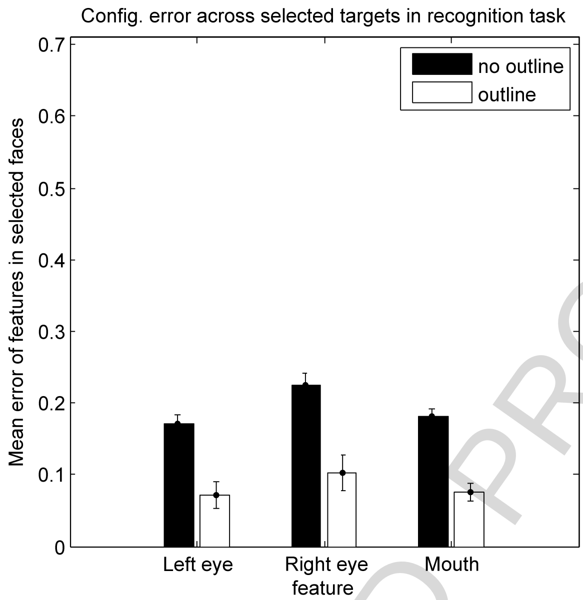

Figure 6. The mean magnitude of errors in faces selected as targets by subjects in Experiment 2. The data suggests that subjects are only 'tricked' by distracters that are extremely similar to the target, indicating that configural information can be used for recognition.

5. DISCUSSION

The results of Experiments 1 and 2 provide some interesting insights into the overall strength and nature of configural encoding in face recognition, as well as a look at the complex relationship between face production and perception.

From Experiment 1 we are first able to conclude that the ability to reconstruct the configuration of facial features is significantly aided by the presence of the external outline. It is important to note that the distribution of error vector angles does not change significantly with the inclusion of the external outline, but that the magnitude of the error vectors does decrease. This demonstrates that configural errors are not merely the result of an inability to determine the proper scale of the face from the internal features alone (though this could also be interesting). Were this so, we might surmise that aspects of feature configuration are accurately encoded, but without sufficient information to recover the proper scale from the size of individual features. What we find instead is that the absence of the external outline does not simply induce a uniform 'expansion' of the face that is reduced in the presence of a constraining contour. The addition of the outline reduces the magnitude of the errors without significantly affecting their directions. This suggests that the traditional definitions of face configuration might need to be revised to incorporate relationships between internal and external features. Rather than conceiving face geometry as a set of relationships between discrete facial features such as the eyes, nose, mouth and eyebrows, it may be much more useful to include a wider range of distances and distance ratios. We do not know the extent to which measurements linking internal and external facial features (or various combinations of them) carry diagnostic information for recognition. The results of Experiment 1 provide some evidence that they might, however. At the very least, these results indicate that just as there are global influences on the recognition of individual features (Tanaka and Farah, 1993) there also appear to be similar affects on second-order relationships between features.

Experiment 2 provides us with more evidence that external features play an important role in configural processing, as well as leaving us with an intriguing conundrum that is central to the relationship between vision and art. In our production task, we find that in the best case (when head outline is included), subjects make errors in feature placement that are approximately 20–25% of the true distance between the nose and the feature under consideration. Objectively, it is hard to say whether this is a large or small amount of error. When we consider the level of performance achieved by subjects in Experiment 2, however, we must conclude that these errors are too large for naïve observers to be tricked. Here then, is the paradox: Why is production so much harder than recognition? The errors made in our production task indicate that configural information may not be strongly encoded, but subjects' ability to detect the true face out of a large set of distracters indicates that recovering the true configuration is possible. What then, is the nature of face representation with regard to configuration? If it is encoded strongly, what renders it inaccessible to subjects who are performing the production task?

Given that observers in Experiment 2 are capable of comparing the veridical image to the distracters and selecting the correct target most of the time, it seems that a participant in Experiment 1 should be able to adopt a similar strategy. However, rather than comparing a large set of images to one another, the 'portrait artist' in Experiment 1 must instead compare the current image to some internal standard. Could the difference in performance between production and recognition be a consequence of this internal representation being weaker or noisier than a physical image? We suggest that this is not necessarily the case. The fact that the veridical image is displayed in the test array during the recognition task does not obviate the need for an internal standard. Comparing an image to another image may be easier than comparing it to an imagined one, but recognition as we have defined it here cannot be carried out without some ultimate appeal to an internal representation of the face. It is not enough to judge that two images differ; one also has to decide which one is veridical. Thus, both tasks rest on the strength of memory for each face.

We suggest that the primary difficulty in producing a convincing replica of each celebrity's face is not memory for the face or weak encoding of configuration, but an online hysteresis effect related to holistic face processing. First, consider the composite face effect (Young *et al.*, 1987) in which a novel face is made by combining the top half of one face with the bottom half of another. Identifying the constituent faces in such a stimulus is very difficult. It has been suggested that the reason for this difficulty is that the combined stimulus, though strange looking, is sufficient to engage global processes that overwhelm the purely local analysis of the top or bottom halves. The new composite does not look familiar when considered globally and local features are unable to make themselves heard, so to speak. A powerful validation of this hypothesis is the finding that slightly misaligning the top and bottom halves of composite faces makes identification much easier.

Now let us consider the act of 'portraiture' in our production task with this phenomenon in mind. Most subjects in Experiment 1 began by placing the features into a rather haphazard arrangement that respects 'first-order' configuration, but is not meant to capture all the 'second-order' relationships they know are important. As they begin altering the positions of individual features in a more fine-grained way, however, many subjects report a feeling of not knowing at all how to proceed. A common complaint is that while the face does not look 'right', it is not at all clear what is 'wrong' about it. It is at this point that we suggest that the subject is suffering from a version of the composite face effect. The erroneous locations of the facial features make up a *new* face that is not a good representation of the intended individual, but nonetheless satisfies global mechanisms. Just as global processing makes the top and bottom half of a composite face hard to recognize, so too might coarse global processing impair subjects' ability to fine-tune the placement of the eyes, nose and mouth in our production task or any form of portraiture.

This analysis of our production task makes some interesting predictions for further experiments, and also suggests some possible techniques for improving portraiture skills. If pervasive global processing is to blame for the failure to accurately reproduce face configuration, the same manipulations that eliminate the composite face effect should reduce errors in our task. Some form of misalignment that 'breaks' the global organization of the face may greatly reduce the errors evident in our task. It may also be useful to examine how placement errors are affected by only allowing subjects to view a small portion of the image at a time. Masking off large portions of the image may be an alternative means of limiting the impact of global processes on portrait-creation. As an interesting side note, one instruction often given to budding artists is to attempt drawing copies of inverted images (Edwards, 1989). In many cases, the drawings created from the inverted image are substantially more accurate than those created from the upright image. While it is usually not discussed in these terms, the benefits of this strategy may result from the disruption of global templates that impede important local analyses. We also note that other complex objects with a conspicuous global configuration (such as houses) may not be processed in the same way as faces, leading to a different pattern of behavior in tasks like those presented here. Global features can profoundly affect local processing for face stimuli, but this is not necessarily the case for non-face objects (Tanaka and Farah, 1993). For example, correctly recognizing an individual face part (such as a nose) is easier when it appears in the context of the entire face. Recognition of an individual house part (such as a door) is not influenced as much by the presence or absence of the entire house. The lack of strong effects of global context for non-face objects could thus lead to much better performance in our production task.

An additional issue that may also be worth investigating is the difference between observers' ability to create a portrait from memory (as we have required our participants to do here) and their ability to make a good copy of a briefly presented veridical image using the same feature placement paradigm. Visual memories often deviate towards 'simpler' forms, so distinctive aspects of facial configuration may be difficult to recover from memory, but relatively easy to encode and reproduce given even a briefly presented stimulus. If portrait construction is more accurate in this scenario, that would provide good evidence that facial configuration features are available to observers in the short-term, but may be overwhelmed in the long-term by the tendency to remember pictures in terms of a 'better' or more typical stimulus.

Finally, we suggest that it would prove valuable to complement the results we have presented here with a computational analysis of the configuration of features in existing portraits. Where possible, it would be interesting to determine the extent to which artists actually do capture the correct distances and angles between discrete facial features. Though it seems portrait artists are not encouraged to explicitly measure and replicate face geometry in this manner, it may be the case that they do so implicitly. In particular, studying caricatures in this manner may prove particularly intriguing. Given that recognition is possible from a caricature even though the appearance of individual features can be exaggerated dramatically, it may be the case that some amount of diagnostic information is contained in the second-order relationships between features we have discussed here. Overall, further exploration into how facial likeness is captured by artists is likely to provide deep insights into the representations and mechanisms that support face recognition in even the poorest draftsman.

6. CONCLUSIONS

Through the use of a face-production task, we have demonstrated that adult observers are generally incapable of accurately synthesizing the facial configuration of familiar individuals. Subjects' ability to accurately recreate feature configurations is greatly aided by the addition of the external outline, however, suggesting that the encoding of second-order relationships between facial features may incorporate geometric relationships between internal and external points. Finally, we note that for the general population, recognition abilities far outpace production skills. We suggest that strong global biases create a hysteresis effect during portrait production, making synthesis of the target face difficult. Strategies that minimize the influence of global processing should reduce errors in portrait creation. The use of a productive task, combined with future analysis of portraits and caricatures, provides an opportunity to study the complexities of face recognition with a wide range of new tools, and demonstrates how fruitful it can be to have a dialog between art and vision.

Acknowledgements

The authors are grateful to Winston Chang, Erin Conwell and Richard Russell for their observations and comments regarding this work. B. J. B. is a National Defense Science and Engineering Graduate Fellow. P. S. is supported by an Alfred P. Sloan Fellowship and a Merck Foundation grant.

REFERENCES

Bartlett, J. C. and Searcy, J. (1993). Inversion and configuration of faces, *Cognitive Psychol.* **25**, 281–316.

Batschelet, E. (1981). *Circular Statistics in Biology*. Academic Press, New York, USA.

Blake, W. (1981). *Portrait Drawing*. Watson Guptill Publications, New York, USA.

Bredart, S. and Devue, C. (2006). The accuracy of memory for faces of personally known individuals, *Perception* **35**, 101–106.

Brunelli, R. and Poggio, T. (1993). Face recognition: features versus templates, *IEEE Trans. Pattern Anal. Machine Intell.* **15**, 1042–1052.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. University of California Press, Berkeley, CA, USA.

Collishaw, S. M. and Hole, G. J. (2002). Is there a linear or a nonlinear relationship between rotation and configural processing of faces? *Perception* **31**, 287–296.

Diamond, R. and Carey, S. (1986). Why faces are and are not special: an effect of expertise, *J. Exper. Psychol.: General* **115**, 107–117.

Edwards, B. (1989). *Drawing on the Right Side of the Brain*. Tarcher/Putnam, New York, USA.

Ge, L., Luo, J., Nishimura, M. and Lee, K. (2003). The lasting impression of Chairman Mao: hyperfidelity of familiar-face memory, *Perception* **32**, 601–614.

Graves, D. R. (1974). *Drawing Portraits*. Watson Guptill Publications, New York, USA.

Graves, D. R. (1984). *Drawing a Likeness*. Watson Guptill Publications, New York, USA.

Haig, N. D. (1984). The effect of feature displacement on face recognition, *Perception* **13**, 505–512.

Hogarth, B. (1965). *Drawing the Human Head*. Watson Guptill Publications, New York, USA.

Hosie, J. A., Ellis, H. D. and Haig, N. D. (1988). The effect of feature displacement on the perception of well-known faces, *Perception* **17**, 461–474.

Jammalamadaka, S. R. (2001). *Topics in Circular Statistics*. World Scientific, River Edge, NJ, USA.

Leder, H. and Bruce, V. (2000). When inverted faces are recognized: the role of configural information in face recognition, *Quart. J. Exper. Psychol. A* **53**, 513–536.

Le Grand, R., Mondloch, C. J., Maurer, D. and Brent, H. P. (2001). Early visual experience and face processing, *Nature* **410**, 890.

Lewis, M. B. and Johnston, R. A. (1997). The Thatcher illusion as a test of configural disruption, *Perception* **26**, 225–227.

Mondloch, C. J., Le Grand, R. and Maurer, D. (2002). Configural face processing develops more slowly than featural face processing, *Perception* **31**, 553–566.

O'Donnell, C. and Bruce, V. (2001). Familiarisation with faces selectively enhances sensitivity to changes made to the eyes, *Perception* **30**, 755–764.

VSP 2001/04/20 Prn:14/08/2007; 10:41 {RA} F:sv2227.tex; VTeX/R p. 17 (941-965)

Rhodes, G., Brake, S. and Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition* **47**, 25–57.

Sinha, P. and Poggio, T. (1996). I think I know that face. . . , *Nature* **384**, 404.

Tanaka, J. W. and Farah, M. J. (1993). Parts and wholes in face recognition, *Quart. J. Exper. Psychol. A* **46**, 225–245.

Wiskott, L., Fellous, J.-M., Kruger, N. and von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching, *IEEE Trans. Pattern Anal. Machine Intell.* **19**, 775–779.

Yin, R. K. (1969). Looking at upside-down faces, *J. Exper. Psychol.* **81**, 141–145.

Young, A. W., Hay, D. C., Mcweeny, K. H., Flude, B. M. and Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features, *Perception* **14**, 737–746.

Young, A. W., Hellawell, D. and Hay, D. C. (1987). Configurational information in face perception, *Perception* **16**, 747–759.