Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
22(12) 1518–1526

DOI: 10.1177/0956797611424290
http://pss.sagepub.com
SAGE

Local Jekyll and Global Hyde: The Dual Identity of Face Identification

Sébastien Miellet[1,2], Roberto Caldara[2], and Philippe G. Schyns[1]
[1]Institute of Neuroscience and Psychology, University of Glasgow, and
[2]Department of Psychology, University of Fribourg

Abstract

The main concern in face-processing research is to understand the processes underlying the identification of faces. In the study reported here, we addressed this issue by examining whether local or global information supports face identification. We developed a new methodology called "*i*Hybrid." This technique combines two famous identities in a gaze-contingent paradigm, which simultaneously provides local, foveated information from one face and global, complementary information from a second face. Behavioral face-identification performance and eye-tracking data showed that the visual system identified faces on the basis of either local or global information depending on the location of the observer's first fixation. In some cases, a given observer even identified the same face using local information on one trial and global information on another trial. A validation in natural viewing conditions confirmed our findings. These results clearly demonstrate that face identification is not rooted in a single, or even preferred, information-gathering strategy.

Keywords

face identification, eye movements, gaze contingent, *i*Hybrid

Received 11/22/10; Revision accepted 4/20/11

An important issue in face processing concerns what information is crucial for facial identification. Specifically, what individual variation among faces does the visual system use to identify them? The main issue concerns whether this diagnostic variance is *local* to facial features or *global* to the patterns comprising facial features. To illustrate, it is undeniable that faces vary locally: Noses, for example, can be turned up, hawk shaped, Greek, Roman, Nubian, or snub, and human visual cognition categorizes distinctive facial features, such as almond-shaped eyes, pouty mouths, square jaws, and weak chins. But the question of whether this undeniable local variance of faces generally contributes to their identification is still highly controversial.

In ecologically valid conditions, faces are recognized over a wide variety of poses and conditions of illumination, even when they are partially obstructed and the range of viewing distances is considerable. Individual facial features combine into higher-order global facial patterns that can be invariant over a large range of challenging viewing conditions. In contrast, the local shape of a nose might become ambiguous outside the range of conversational distances, and the color of an iris might simply vanish as a result of the physical limitations of retinal sampling. Thus, a robust face-identification mechanism would have to use different sorts of information from the same face.

Face researchers typically discuss face-identification mechanisms in terms of different formats of representation—that is, specific factors over which facial variance is measured, including single-pixel luminance, clusters of pixels forming face parts, higher-order relationships between face parts, entire faces at low-frequency bands or a combination of spatial-frequency bands, local Gabor jets, and three-dimensional variant structures with or without chromatic and textural cues. In the study reported here, we embraced a different perspective on face identity, focusing instead on distal stimuli to ascertain whether the visual system uses information at local or global scales.

During eye fixations, the visual system locally samples foveated information at a high resolution to categorize stimuli. The seminal work of Yarbus (1967) revealed that when Westerners fixate on faces, their eyes follow a systematic triangular sequence that locally samples the two eyes and mouth over the course of face identification (e.g., Althoff & Cohen, 1999;

Corresponding Authors:
Sébastien Miellet, Department of Psychology, Faucigny 2, Fribourg, 1700, Switzerland
E-mail: sebastien.miellet@unifr.ch

Philippe G. Schyns, Institute of Neuroscience and Psychology, 58 Hillhead St., Glasgow, G12 8QB, United Kingdom
E-mail: philippe.schyns@glasgow.ac.uk

Groner, Walder, & Groner, 1984; Henderson, Williams, & Falk, 2005). However, recent studies have shown that central fixations are deployed by both Westerners (Hsiao & Cottrell, 2008) and Easterners (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Kelly, Miellet, & Caldara, 2010; Kita et al., 2010; Rodger, Kelly, Blais, & Caldara, 2010). Because retinal cell density and visual resolution decrease steeply outside the fovea, the center of the face is likely to be the most advantageous location at which to sample global information. However, the basic question still remains to be clarified of whether face identification proceeds from extraction of local or global stimulus information because fixation on a face region does not necessarily imply usage of its underlying information (Caldara, Zhou, & Miellet, 2010; Gosselin & Schyns, 2001; Schyns, Petro, & Smith, 2007).

To address this question, we developed the "*i*Hybrid" technique, a novel methodology that combines a gaze-contingent window with a hybrid stimulus (see Fig. 1). A hybrid stimulus (Oliva & Schyns, 1997; Oliva, Torralba, & Schyns, 2006; Schyns & Oliva, 1994, 1999) comprises two different stimuli, each represented in different spatial-frequency bands spanning the full spectrum. The "*i*" (or "eye") refers to a gaze-contingent window that forces foveated information to be full spectrum and local, and leaves the global information outside the gaze-contingent window. Figure 1[1] illustrates the process of creating an *i*Hybrid stimulus using a fixation on the left eye. (A film clip illustrating the dynamics of fixations over one trial, with the dot representing the fixation location, is available in the Supplemental Material available online and on the third author's Web site: http://www.psy.gla.ac.uk/~philippe/iHybrid_example.mov.) Analysis of fixations on the *i*Hybrid stimuli teases apart information accrued locally, over multiple fixations, from information acquired globally, possibly in a single fixation, but outside the local foveated window. We used famous faces in the study reported here because they do not require extensive identity learning during the experiment and have been categorized in a variety of viewing conditions; both of these factors ensured that participants had considerable expertise with the stimuli.

Method

Participants

Twelve Western-Caucasian young adults (4 males, 8 females; mean age = 23.7 years) from the University of Glasgow participated in this study. All participants had normal or corrected-to-normal vision, gave written informed consent, and were paid for their participation in a protocol approved by the ethics committee at the University of Glasgow.

Stimuli

We used 18 face pictures of famous male actors (selected from images used in a study by Butler, Blais, Gosselin, Bub, & Fiset, 2010). Original images consisted of 260 × 260, 16-bit

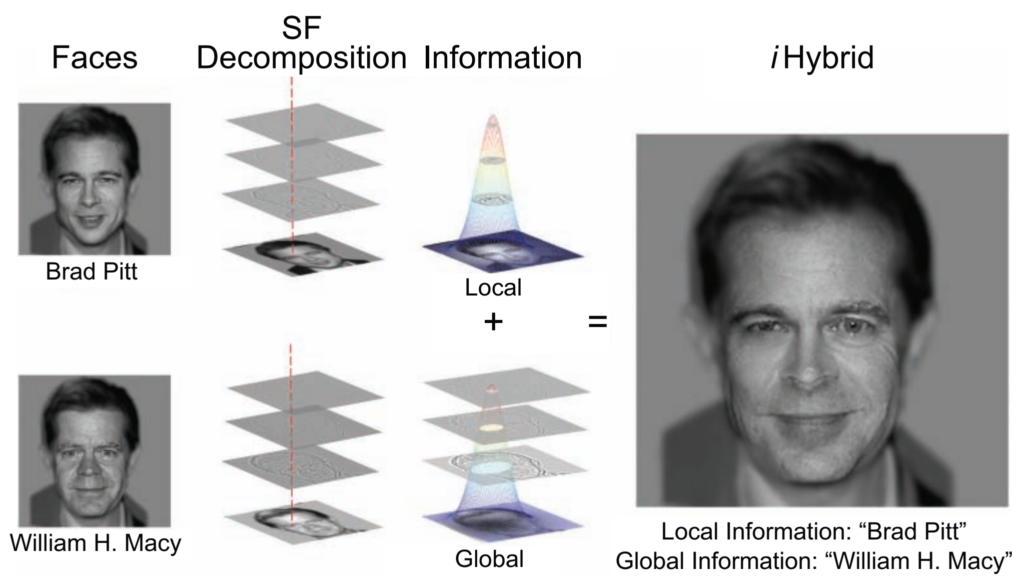

Fig. 1. Procedure used to create *i*Hybrid faces. The spatial frequencies (SFs) of two original face images (illustrated here with Brad Pitt and William H. Macy) were decomposed separately into four nonoverlapping SF bands of 1 octave each (<3, 3–6, 6–12, >12 cycles per degree of visual angle). A Gaussian window (*SD* = 25 pixels, ~1° of visual angle) was then centered on every potential fixation location on each face; this procedure formed a lattice of 5- × 5-pixel cells covering the original 260 × 260 image. When an observer fixated on the stimulus, the local information across the four SF bands for one identity was extracted through the Gaussian window at that location, and the complementary global SF information was extracted from the other identity. The sum of the complementary, fixation-dependent identities formed the *i*Hybrid stimulus. In the example illustrated here, the dashed red line indicates a fixation location at the left eye; local SF information was extracted from this location in the image of Brad Pitt, and the complementary SF information was taken from the image of William H. Macy. An observer who identifies the resulting face as Brad Pitt is using local information, and an observer who identifies this face as William H. Macy is using global information.

gray-level pixels and were normalized for luminance, image positions of the eyes and the mouth, and external features. We created 18 *i*Hybrid stimuli by pairing identities and combining the two identities in each pair in two ways (e.g., foveal information from Brad Pitt combined with extrafoveal information from William H. Macy, and the opposite combination); thus the identity providing the local face information and the identity providing the global face information were counterbalanced within each *i*Hybrid pair.

Procedure

We recorded eye movements using a desktop-mounted EyeLink 2000 (SR Research, Mississauga, Ontario, Canada; sampling rate = 1000 Hz). We maintained the viewing distance at 70 cm with a chin and forehead rest and presented the stimuli (face size: 15.6° × 19.5° of visual angle) on a gray background displayed on a Dell P1130 19-in. CRT monitor with a refresh rate of 170 Hz. Viewing was binocular, and we tracked the participant's dominant eye. We implemented the experiment in MATLAB (The MathWorks, Natick, MA) using the Psychophysics and EyeLink Toolbox extensions (Brainard, 1997; Cornelissen, Peters, & Palmer, 2002).

The main experiment comprised two parts (identity-information estimation and *i*Hybrid identification).

Pretest: identity-information estimation. In order to create the *i*Hybrid stimuli, we used the QUEST procedure (Watson & Pelli, 1983) to estimate the threshold of phase-coherent face information each participant needed to identify each face (see Fig. S1 in the Supplemental Material). Phase coherence is an index of the identity information each picture contains. This measure enabled a more rigorous balancing of the identities composing the *i*Hybrid pairs. In an *i*Hybrid stimulus, the face with the lower phase-coherence threshold might be systematically identified more often that the face with the higher phase-coherence threshold. Instead, we wanted to ensure (a) that each observer could identify each face 75% of the time with a given level of phase coherence, (b) that we balanced each of the two faces in an *i*Hybrid pair for their required levels of phase coherence, and (c) that an eventual imbalance in phase-coherence threshold across two identities would not systematically bias identification responses. It is important to note, however, that we used the original face pictures, not the noisy ones used in the pretest, to create the *i*Hybrid stimuli.

We instructed participants to verbally name the famous person briefly presented on a computer screen. Each trial started with a central fixation cross (500 ms), followed by the phase-manipulated stimulus (100 ms), a 100% noise field (500 ms), and a blank screen during which observers had to respond (unlimited duration). We parametrically modified the phase spectrum of the original faces (by 5% increments) to generate stimuli ranging between pure noise (0% phase coherence) and the undistorted signal (100% phase coherence; Dakin, Hess, Ledgeway, & Achtman, 2002). Using QUEST, we determined the percentage of phase coherence required for 75% identification accuracy of each original identity.

iHybrid procedure. The pretest revealed that 45% of phase-coherent face information was required on average to reach the threshold of 75% correct identification (required phase coherence ranged from 23% to 60%, as expected given the heterogeneity of famous faces). We constructed the *i*Hybrid stimuli by pairing two identities that differed by at most 24% in the phase coherence required for 75% correct identification. The face pairings were the same for all participants.

Because *i*Hybrid computation is time-intensive, we created all stimuli before subjects came to the lab (see Fig. 1 for an illustration of this computation). First, we computed all possible fixation locations on each of the faces in a pair (constrained to a lattice of 5 × 5 pixels on the original 260- × 260-pixel picture). Second, we decomposed each identity into four nonoverlapping spatial-frequency bands of 1 octave each (<3, 3–6, 6–12, and >12 cycles per degree of visual angle) using the MATLAB Pyramid Toolbox (Simoncelli, 1997). For each one of the precomputed fixation locations, we applied spatial-frequency filtering (illustrated in Fig. 1 for a fixation on the left eye). The local information, represented in the four spatial-frequency bands at the point of fixation (the left eye of Brad Pitt in Fig. 1), was extracted through a Gaussian window (*SD* = 25 pixels, ~1° of visual angle). To extract global information, we applied the same technique to the information outside the Gaussian window (most of the face of William H. Macy in Fig. 1). The *i*Hybrid stimulus resulted from the sum of the local and the global information. For each identity pair, we generated two *i*Hybrid stimuli by swapping the face from which we sampled the local information and the face from which we sampled the global information.

The experiment started with a standard nine-point calibration. Then, each experimental trial started with a check to ensure the eye tracker remained properly calibrated: We first presented a central fixation cross, followed by four fixation crosses—one in the middle of each quadrant of the computer screen—and one final central fixation cross for drift correction. A drift greater than 0.5° would automatically launch another standard nine-point calibration to recalibrate the eye tracker. To avoid constraining first fixations to the location of the central fixation cross presented before each trial, the screen went blank for a random time (between 0.5 and 1 s) after the calibration check. This procedure ensured a random gaze position when the *i*Hybrid stimulus was displayed on the screen. We then randomly selected one of the *i*Hybrid stimuli and started the *i*Hybrid procedure. We monitored the observer's fixation location and used this information to index the precomputed *i*Hybrid stimuli on-line. Indexing, which was done continually during the course of each trial, consisted of retrieving the appropriate *i*Hybrid stimulus and updating the display contingent to the participant's gaze position; this process required 11 ms on average (between 8 and 14 ms, about 90 Hz refresh rate on the screen, eliminating any flickering). Each *i*Hybrid trial lasted for 1 s, followed by a 100% noise field presented for 500 ms. Observers identified each *i*Hybrid stimulus by naming aloud the famous actor they perceived. Note that each identity pair was presented twice in the experiment (to enable each identity to serve as the source of both local and global information).

Results

i*Hybrid procedure*

Participants identified one of the two identities composing the *i*Hybrid stimulus with 95% accuracy. We determined the proportion of local-identification responses and of global-identification responses. These proportions served as proxies for the use of local face information and global face information, respectively. We observed a nonsignificant trend for reporting the global face (51% of the time) over reporting the local face (44% of the time), $F(1, 11) = 3.74$, $p = .08$, $\eta_p^2 = .25$. Further examination of individual observers did not reveal any systematic individual differences in reporting the global versus the local face. Faces with a lower phase-coherence threshold (i.e., faces that could be identified with less information) were not selected more often than faces with a higher phase-coherence threshold. Phase coherence conditioned separately on local identification responses and global identification responses led to similar averages (identification of the local face—phase threshold of the local face: 44.12%, phase threshold of the global face: 45.75%; identification of the global face—phase threshold of the local face: 45.75%, phase threshold of the global face: 45.04%).

Turning to fixation data, we computed a fixation map using all correct trials of the experiment—that is, without splitting them according to whether the response was based on local or global information (see Fig. 2a). The overall pattern of fixations revealed the standard T-shaped pattern of fixations joining the two eyes and the mouth in face-identification experiments (Yarbus, 1967). To understand the fixation biases associated with local- and global-information use, we decomposed the overall fixation map into two separate maps showing local and global behavioral responses, respectively (see Fig. 2b). We then created a map showing the difference between local and global fixations; this difference map revealed that observers fixated primarily on the left eye and the mouth when identifying the local face. When identifying the global face, observers fixated primarily on the center of the face.

Analyses of mean number of fixations, $F(1, 11) = 2.09$; mean fixation duration, $F(1, 11) = 0.02$; mean scan-path length, $F(1, 11) = 4.34$; and mean saccade length, $F(1, 11) = 1.75$ did not reveal any general significant effect of local- versus global-information use on these measures (all *p*s > .05; see Table 1).

However, analyzing fixation locations as a function of behavioral response (local vs. global) revealed significantly different patterns. To establish significance, we used the *i*Map toolbox (Caldara & Miellet, 2011). Specifically, we applied a one-tailed pixel test (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005) to the local and global fixation maps ($Z_{critical} > 4.86$, $p < .05$, corrected) and a two-tailed pixel test to the map showing the difference between local and global responses ($Z_{critical} > |5.01|$, $p < .05$, corrected; see Fig. 2b).

As Figure 2b shows, observers primarily fixated on the left eye and the mouth when identifying the local face of the *i*Hybrid stimulus, and primarily fixated on the center of the face when identifying the global face of the *i*Hybrid stimulus. This difference in loci of fixations was confirmed by the *z*-scored fixation durations in the significant areas (local face—eyes-mouth: 9.83, center: 3.15, global face—eyes-mouth: 4.24, center: 8.22), as well as the effect sizes of the difference in fixation durations between the local and global information-sampling strategies (Cohen's *d*—eyes-mouth: 1.4 ms, center: 1.75). We conducted 2 (strategy: local vs. global) × 2 (feature location: eyes-mouth vs. center) analyses of variance on average fixation duration, scan-path length, total fixation duration, and number of fixations per trial, and we found interactive effects on mean fixation durations, $F(1, 11) = 9.03$, $p < .02$, $\eta_p^2 = .45$; scan-path length, $F(1, 11) = 16.53$, $p < .002$, $\eta_p^2 = .60$; total fixation duration, $F(1, 11) = 34.57$, $p < .0001$, $\eta_p^2 = .76$; and number of fixations, $F(1, 11) = 14.26$, $p < .004$, $\eta_p^2 = .56$. The pattern of results confirmed that mean fixation durations, scan-path lengths, and fixation durations were longer for the eyes and mouth area than for the center face area when the local strategy was used and were longer for the center face area than for the eyes and mouth area when the global strategy was used; in addition, the number of fixations on the eyes and mouth area was larger than the number of fixations on the center face area when the local strategy was used, and the number of fixations on the center face area was larger than the number of fixations on the eyes and mouth area when the global strategy was used (see Table 2). No other effect reached significance.

The same observer would sometimes identify the same identity (e.g., Brad Pitt) in one *i*Hybrid stimulus using a local strategy and in the counterbalanced *i*Hybrid stimulus using a global strategy. On average, 64% of the trials induced such complementary use of the two different strategies to identify the same face. Note that in these trials, the "attractor" face was not necessarily the face with the lowest phase-coherence threshold (this was true in only 55% of the cases).

To further understand the relationship between strategy selection (local vs. global) and location of fixation, we computed the probability of identifying the local face and the probability of identifying the global face conditioned on the location of the first fixation on the *i*Hybrid stimulus. As Figure 2c shows, the probability of using a local strategy was significantly higher than the probability of using a global strategy when the first fixation was on the left eye or the mouth, $t(11) = 2.25$, $p < .05$. In contrast, the probability of using a global strategy was significantly higher than the probability of using a local strategy when the first fixation landed in the

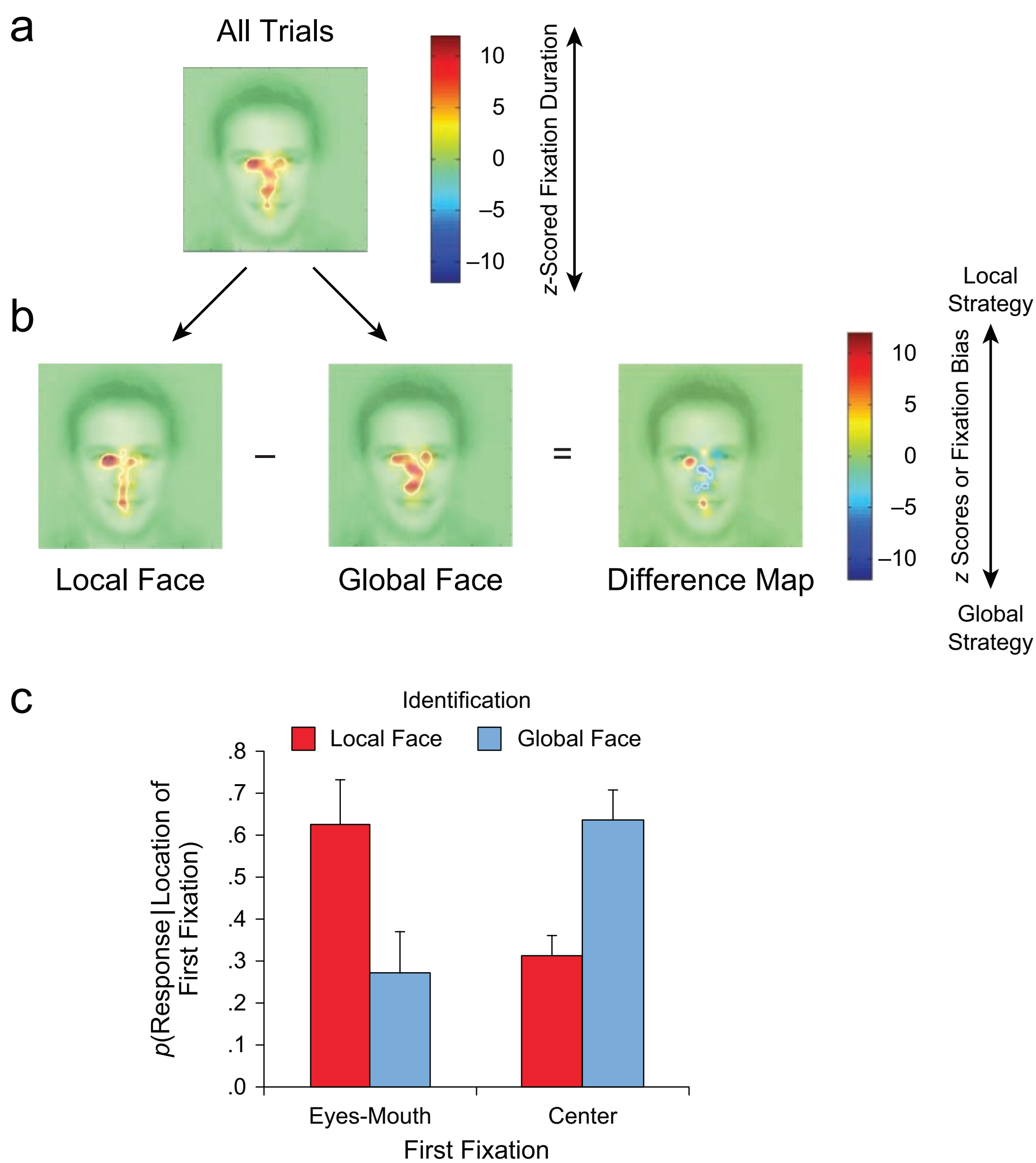

Fig. 2. Results from the experiment. A fixation map (a) was created by collapsing *z*-scored fixation durations across all correct trials for all *i*Hybrids in the experiment. This fixation map was then split into separate maps (b) showing fixation durations when responses were based on local face identifications and fixation durations when responses were based on global face identifications. Subtracting the fixation map for global identification responses from the fixation map for local identification responses resulted in a difference map. On the local and global maps shown here, white contours surround regions of significantly longer fixation durations than were observed for other areas; on the difference map, white contours indicate regions of significant differences between the local and global maps (i.e., fixation bias). The graph (c) shows the conditional probability of identification responses (local face or global face) as a function of the location of the first fixation (in the eyes-mouth region or in the center of the face). Error bars show standard errors of the mean.

Table 1. Mean Number of Fixations, Fixation Duration, Scan-Path Length, and Saccade Length as a Function of the Strategy Used to Identify the *i*Hybrid Stimulus

Strategy used	Number of fixations	Fixation duration (ms)	Scan-path length (°)	Saccade length (°)
Local	3.25	391	6.74	1.88
Global	2.99	386	5.78	1.73

Table 2. Mean Fixation Duration, Scan-Path Length, Total Fixation Duration, and Number of Fixations as a Function of Feature Location and the Strategy Used to Identify the *i*Hybrid Stimulus

	Local strategy		Global strategy	
Variable	Eyes-mouth	Center	Eyes-mouth	Center
Fixation duration (ms)	402	296	313	379
Scan-path length (°)	1.46	0.40	0.61	1.07
Total fixation duration (ms)	222	47	70	206
Number of fixations	0.58	0.18	0.24	0.56

center of the face, $t(11) = 3.04$, $p < .02$. Thus, the first information sampled (either via a fixation on the eye or the mouth or via a fixation on the center of the face) determined the choice of a local or a global information-sampling strategy, and consequently which identity of the *i*Hybrid stimulus was reported.

Validation: local versus global face identification

To validate that both local and global fixated information are sufficient for face identification in natural viewing conditions, we performed the following validation with 10 Western-Caucasian young adults with normal or corrected-to-normal vision (4 males, 6 females; mean age = 28.3 years), none of whom participated in the main experiment.

Using the fixation maps obtained in the main experiment, we applied local- and global-information filters to reconstruct 36 hybrid identities. We created these filters from the significant local and global regions of the difference fixation map (shown in Fig. 2b and in Fig. 3a). We used the filters to extract local and global information from each identity within a pair; then we added the rest of the data (i.e., the complementary information) from the other identity. For each face pair and for the two fixation-strategy filters, we thus constructed four hybrid faces (Pitt-local/Macy-complementary, Macy-local/Pitt-complementary, Pitt-global/Macy-complementary, Macy-global/Pitt-complementary; see Fig. S2 in the Supplemental Material), for a total of 36 hybrid faces (Fig. 3a illustrates this process). The viewing parameters were identical to those used during the *i*Hybrid procedure, and the stimuli were again presented for 1 s. However, it is important to emphasize that in this validation posttest, observers saw full-spectrum static stimuli on the screen, not gaze-contingent displays. Observers were instructed to identify the faces by naming them aloud.

We predicted that the reconstructed hybrids would be preferentially identified according to the local or the global information extracted by the filters. If this occurred, it would validate these cues for face identification in more naturalistic viewing conditions, but still in the context of hybrid identities. The data confirmed these predictions (see Fig. 3b), $t(9) = 18.20$, $p < .001$. Furthermore, face-identification performance demonstrated a very strong interobserver agreement (Fleiss's $\kappa = .94$). This validation with local and global information extracted using the fixation maps from the *i*Hybrid procedure confirms that sufficient information was indeed extracted to locally or globally identify the faces in the *i*Hybrid procedure. It also confirms that the use of this information generalizes to more naturalistic (i.e., not eye contingent) viewing conditions.

Discussion

The use of *i*Hybrid stimuli revealed the existence of two distinct, equally frequent, and equally effective information-sampling strategies for face identification. The local strategy involves fixations on the eyes and the mouth, whereas the global strategy relies on central fixations of the face. All observers used both strategies, often to recover the very same identity. No strategy was systematically associated with specific identities: One observer could use a local strategy to identify Tom Cruise, for example, whereas another observer could use a global strategy to identify him. We did not find an association between choice of strategy and the amount of identity information available (as measured by percentage of phase coherence in the identity), but instead demonstrated a strong link between strategy selection and location of the initial fixation on the face. First fixations on the eyes and mouth led to a local strategy, whereas initial fixations in the center of the face promoted a global strategy. Note that when we collapsed fixations across the two strategies, we validated Yarbus's (1967) well-known pattern of fixations over the two eyes and the mouth. Furthermore, hybrid faces reconstructed according to the local versus global sampling of information led to the predicted identifications. We therefore conclude that the face system flexibly uses local or global stimulus information to identify faces depending on the constraint of the information sampled in the initial fixation. These findings have important implications for face identification.

The face-identification literature presents long-standing questions on the specialized processing of faces compared with the processing of other objects and scenes (e.g., Kanwisher, 2000). Some researchers argue that faces are processed as relatively undifferentiated wholes (i.e., *holistically*), without a differentiated and explicit representation of the individual facial features (e.g., Farah, Tanaka, & Drain, 1995; Tanaka & Farah, 1993). Other researchers argue that faces are processed on the basis of both *configural* and *featural*

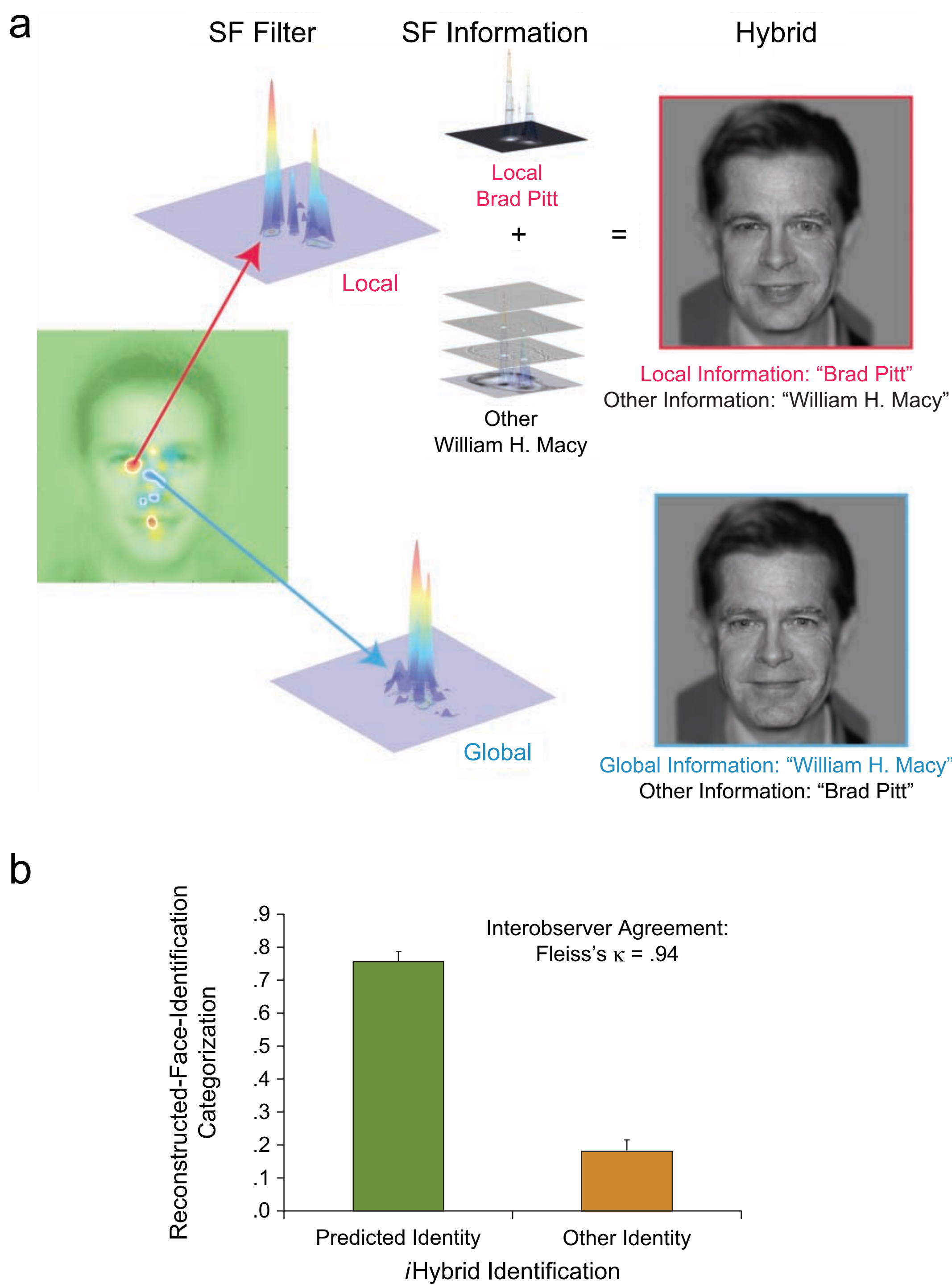

Fig. 3. Illustration of reconstructed faces (a) and results (b) from the posttest validation analysis. Using the difference maps from the *i*Hybrid analyses (a), we isolated the regions associated with local-identification responses (colored in red) and the regions associated with global-identification responses (colored in blue). In the example shown here, one reconstructed hybrid was created by applying a spatial-frequency (SF) filter to select the local information for Brad Pitt and completing the image with the complementary information from William H. Macy. The other reconstructed hybrid was created by using an SF filter to select the global information for Macy and filling in the image with the complementary information from Pitt. Thus, two reconstructed hybrids were created for each *i*Hybrid stimulus (i.e., four reconstructed hybrids were created for each face pair) by applying local and global SF filters separately to the original face images. The graph (b) shows the proportion of identifications that corresponded with the information selected by the SF filter (predicted identity) and the proportion of face identifications that corresponded with the person from whom the complementary information was derived (other identity). Error bars show standard errors of the mean.

information, with separate representational systems for each kind of information (e.g., Cabeza & Kato 2000; Sergent 1984). Configural processing would primarily use the metric distances between individual face features as information (for a review, see Maurer, Le Grand, & Mondloch, 2002; however, see Taschereau-Dumouchel, Rossion, Schyns, & Gosselin, 2010, for a statistical demonstration that metric distances are generally insufficient). Featural processing would rely on information from individual features (e.g., the shape or color of the eyes or the shape of the chin, nose, and mouth) or a combination of features (e.g., Schyns, Bonnar, & Gosselin, 2002; Schyns, Petro, & Smith, 2009; Sekuler, Gaspar, Gold, & Bennett, 2004; van Rijsbergen & Schyns, 2010).

Notwithstanding the conceptual difficulty of teasing apart these positions and empirically testing them, our results suggest that when the visual system extracts local features over multiple fixations, it does not acquire an undifferentiated whole—because this is incompatible with multiple fixations on different features. When the visual system extracts information globally from fixating on the center of the face, then this global information is probably sufficient for a direct comparison with the version of that face stored in memory. Thus, it is not surprising that the face-identification literature is full of debates on representation formats (featural vs. holistic or configural): Sampling of face information with fixations suggests the use of both local and global information to achieve identification of the same familiar face.

If both local and global strategies exist in the visual system (at least for identification of famous faces), culture could also determine a preferential strategy for unfamiliar faces and thus account for the complex pattern of data in this literature (see, e.g., Althoff & Cohen, 1999; Groner et al., 1984; and Henderson et al., 2005, for data indicating local strategies and see, e.g., Blais et al., 2008; Caldara et al., 2010; Hsiao & Cottrell, 2008; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Kelly et al., 2010; and Kita et al., 2010, for data indicating global strategies). Accordingly, a cultural bias for a global strategy in Easterners could arise from an equally effective style of general visual-information extraction used for object identification (Kelly et al., 2010), which might originate from a cultural norm to direct the first fixation less often directly toward the eyes.

In sum, we have shown that the eye movement strategies of face identification flexibly use local or global face information, even when the same observer identifies the same face. Our results challenge the notion of a mandatory route for recovering face identity.

Acknowledgments

We thank Frederic Gosselin and Daniel Fiset for providing the famous-faces data set.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

S. M., R. C., and P. G. S. are supported by The Economic and Social Research Council and Medical Research Council (ESRC/RES-060-25-0010).

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Note

1. Note that Figures 1 and 3 do not show the actual stimuli, which could not be published for licensing reasons. Rather, the figures show re-creations from alternative images of the actors whose faces were used in the studies. The faces used for Figures 1 and 3 are almost identical to the ones used in the studies, and the figures were created using exactly the same techniques as in the studies.

References

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 997–1010.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*, *3*(8), e3022. Retrieved from http://www.plosone.org/article/info:doi/10.1371/journal.pone.0003022

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics*, *72*, 1444–1449.

Cabeza, R., & Kato, T. (2000). Features are also important: Contributions of featural and configural processing to face recognition. *Psychological Science*, *11*, 429–433.

Caldara, R., & Miellet, S. (2011). iMap: A novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods, 43,* 864–878.

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the spotlight reveals universal information use for face recognition. *PLoS ONE*, *5*(3), e9708. Retrieved from http://www.plosone.org/article/info:doi/10.1371/journal.pone.0009708

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*(9), Article 1. Retrieved from http://www.journalofvision.org/content/5/9/1

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*, 613–617.

Dakin, S. C., Hess, R. F., Ledgeway, T., & Achtman, R. L. (2002). What causes nonmonotonic tuning of fMRI response to noisy images? *Current Biology*, *12*, R476–R477.

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.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A new technique to reveal the use of visual information in recognition tasks. *Vision Research*, *41*, 2261–2271.

Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movements research* (pp. 523–533). Amsterdam, The Netherlands: Elsevier.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*, 98–106.

Hsiao, J. H.-W., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*, 998–1006.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, *19*, 1543–1548.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*, 759–763.

Kelly, D. J., Miellet, S., & Caldara, R. (2010). Culture shapes eye movements for visually homogeneous objects. *Frontiers in Psychology*, *1*. Retrieved from http://www.frontiersin.org/perception_science/10.3389/fpsyg.2010.00006/full

Kita, Y., Gunji, A., Sakihara, K., Inagaki, M., Kaga, M., Nakagawa, E., & Hosokawa, T. (2010). Scanning strategies do not modulate face identification: Eye-tracking and near-infrared spectroscopy study. *PLoS ONE*, *5*(6), e11050. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2883577/

Maurer, D., Le Grand, R., & Mondloch, C. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs, or fine scale edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, *34*, 72–107.

Oliva, A., Torralba, A., & Schyns, P. G. (2006). Hybrid images. *ACM Transactions on Graphics*, *25*, 527–532.

Rodger, H., Kelly, D., Blais, C., & Caldara, R. (2010). Inverting faces does not abolish cultural diversity in eye movements. *Perception*, *39*, 1491–1503.

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.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*, 195–200.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Schyns, P. G., Petro, L., & Smith, M. L. (2007). Dynamics of visual information integration in the brain to categorize facial expressions. *Current Biology*, *17*, 1580–1585.

Schyns, P. G., Petro, L., & Smith, M. L. (2009). Transmission of facial expressions co-evolved with their efficient decoding in the brain: Behavioral and brain evidence. *PLoS ONE*, *4*(5), e5625. Retrieved from http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0005625

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. (1984). An investigation into component and configurational processes underlying face recognition. *British Journal of Psychology*, *75*, 221–242.

Simoncelli, E. P. (1997). Image and multi-scale pyramid tools [Computer software]. Retrieved from www.cns.nyu.edu/~eero/software.html

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology A*, *46*, 225–245.

Taschereau-Dumouchel, V., Rossion, B., Schyns, P. G., & Gosselin, F. (2010). Interattribute distances do not represent the identity of real-world faces. *Frontiers in Perception Science*, *1*. Retrieved from http://www.frontiersin.org/perception_science/10.3389/fpsyg.2010.00159/full

van Rijsbergen, N. J., & Schyns, P. G. (2010). Dynamics of trimming the content of face representations for categorization in the brain. *PLoS Computational Biology*, *5*(11), e1000561. Retrieved from http:// www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000561

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120.

Yarbus, A. L. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York, NY: Plenum Press.