2011

Local Jekyll and global Hyde: The dual identity of face identification

Sebastien R. Miellet
University of Wollongong, smiellet@uow.edu.au

Roberto Caldara
University of Fribourg

Philippe Schyns
University of Glasgow

Publication Details

Miellet, S. R., Caldara, R. & Schyns, P. (2011). Local Jekyll and global Hyde: The dual identity of face identification. Psychological Science, 22 (12), 1518-1526.

Local Jekyll and global Hyde: The dual identity of face identification

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 "iHybrid." 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

dual, hyde:, global, jekyll, identification, local, face, identity

Disciplines

Education | Social and Behavioral Sciences

Publication Details

Miellet, S. R., Caldara, R. & Schyns, P. (2011). Local Jekyll and global Hyde: The dual identity of face identification. Psychological Science, 22 (12), 1518-1526.

Local Jekyll and global Hyde:

The dual identity of face identification

Sébastien Miellet[1]
Roberto Caldara[2]
Philippe G. Schyns[1]

[1] Institute of Neuroscience and Psychology
University of Glasgow
[2] Department of Psychology
University of Fribourg
Switzerland

Word count: Manuscript = 3492 words
Abstract = 144 words

Corresponding author:

Sébastien Miellet or Philippe G. Schyns
Institute of Neuroscience and Psychology
58, Hillhead Street
Glasgow, G128QB
sebastien.miellet@glasgow.ac.uk

Running head: *i*Hybrids reveal local and global information use in face identification
Key words: Face identification, eye movements, gaze-contingent, *i*Hybrid

ABSTRACT

The cornerstone of the face processing research agenda is an understanding of the processes underlying the identification of faces. Here, we address this issue by examining whether local or global information supports face identification. We deploy a new methodology (*i*Hybrid) that 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 show that the visual system identifies faces on the basis of either local or global information depending on the first fixation location, even identifying the *same* identity using local information on one trial and global information on another trial. A validation in natural viewing conditions confirmed our results. These results clearly demonstrate that the face system is not rooted in a single, or even preferred, information gathering strategy for face identification.

An important issue in face processing concerns the information subtending facial identification. By information, we mean the variance across faces that the visual system uses for their identification. 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 can be turned-up, hawk-shaped, Greek, Roman, Nubian or Snub, and our visual cognition categorizes the color of the eyes, their almond shapes, the pouty mouths, the square jaws and the weak chins. But the question of interest is whether this undeniable local variance of faces generally contributes to their identification.

In ecologically valid conditions, faces are recognized over a wide variety of poses, conditions of illumination, obstruction of parts and over a considerable range of viewing distances. Facial features combine into higher-order global facial patterns that could be invariant over a large range of the challenging viewing conditions. In contrast, the local shape of the nose might become ambiguous outside the range of "conversational distances" and the color of the iris might simply vanish due to physical limitations of retinal sampling. Thus, a robust face identification mechanism could proceed from different sorts of information from the same face.

Face researchers typically discuss face identity mechanisms in terms of different formats of representation--i.e. specific face measurements over which variance is measured, from single pixel luminance, clusters of pixels forming face parts, higher-order relationships between face parts, entire faces at low, or at a combination of spatial frequency bands, local Gabor jets, three-dimensional variant structures with or without chromatic and textural cues, and so forth. Here, we embraced a different perspective on face identity, focusing instead on the distal stimulus, to ascertain whether the visual system uses information at local or global scales. Outcomes should inform and constrain issues of representation formats and processes, to which we return in the discussion.

Eye fixations describe how the visual system locally samples foveated information at a high resolution to categorize stimuli. The seminal work of Yarbus (1965) revealed that fixations follow a systematic triangular sequence sampling the local eyes and mouth over the course of face identification in Westerners (e.g., Althoff & Cohen, 1999; Groner, Walder & Groner, 1984; Henderson, Williams & Falk, 2005), but recent studies showed the deployment of central fixations in Westerners (Hsiao & Cottrell, 2008) and Easterners (Blais, Jack, Scheepers, Fiset & Caldara, 2008; Kelly, Miellet & Caldara, 2010; Kita, Gunji, Sakihara, Inagaki, Kaga, Nakagawa & Hosokawa, 2010; Rodger, Kelly, Blais & Caldara, 2010). Because retinal cell density and visual resolution decrease steeply outside the fovea, the centre of the face is likely to be the most advantageous face location to sample global information. In this context, the basic question 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 the underlying information (Caldara, Zhou & Miellet, 2010; Gosselin & Schyns, 2001; Schyns, Petro & Smith, 2007).

To address the question, we developed the "*i*Hybrid," a novel methodology that combines a gaze-contingent window with the hybrid stimulus (see Figure 1). A hybrid stimulus (Oliva et al., 2006; Oliva & Schyns, 1997; Schyns & Oliva, 1994; Schyns & Oliva, 1999) comprises two different stimuli, each represented in different spatial frequency (SF) bands spanning the full SF spectrum. To this, we added the "*i*", a gaze-contingent window that forces foveated information to be full spectrum and local, leaving the global information outside the gaze-contingent window. Figure 1 explains the computation of the iHybrid and illustrates it with a fixation on the left eye. See Supporting Movie 1 for an illustration of the dynamics of fixations over one trial (the dot represents the fixation location):

http://www.psy.gla.ac.uk/~philippe/iHybrid_example.mov. The *i*Hybrid 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 because they do not require extensive identity learning in the experimental room, have been categorized in a variety of viewing conditions, leading to considerable expertise with the stimuli.

= = = = = = = = = = = =

FIGURE 1

= = = = = = = = = = = =

Methods

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

Stimuli. We used 18 famous male actor face pictures (from Butler et al., 2010, database). Original images were 260x260, 16-bits gray-level pixels and normalized for luminance, image positions of the eyes and the mouth, and external features. We created *i*Hybrids by combining two identities resulting in two possible combinations (e.g. Brad Pitt-foveal and William Macy-extrafoveal and the opposite). Across 18 identities, this produced a total of 18 *i*Hybrids identity pairings, counterbalancing which identity provides the local versus the global face information.

Procedure. The main experiment comprised two parts (identity information estimation and *i*Hybrid) described in turn.

Pre-test: Identity Information Estimation. For each participant we used QUEST (Watson & Pelli, 1983) to estimate the individual threshold of phase coherent face information (versus phase noise) required to identify each face (see Supporting Fig. 1). Phase coherence is an index of the identity information each picture represents. This measure enabled a more rigorous balancing of the identities composing the *i*Hybrids. Faces with a lower phase coherence threshold might be systematically identified when presenting the *i*Hybrids (their identity could systematically “stand out”). Instead, we wanted to ensure (a) that each observer could identify each face at 75% threshold with a given level of phase coherence, (b) that we balanced *i*Hybrids across identities 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 the *i*Hybrids comprised the original, not the noisy face pictures.

We instructed participants to verbally name the famous person briefly presented on a computer screen. Each trial comprised a central fixation cross (500ms), followed by the phase manipulated stimulus (100ms), a 100% noise field for 500ms and an unconstrained delay for verbal response. 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 et al., 2002). QUEST determined the percentage of phase coherence required for 75% threshold identification accuracy of each original identity. We maintained the viewing distance at 70 cm with a chinrest and presented the stimuli (face size: 15.6x19.5 deg of visual angle) on a grey background displayed on a Dell P1130 19” CRT monitor with a refresh rate of 170 Hz. We applied the same viewing parameters to the *i*Hybrid procedure.

*i*Hybrid Procedure. The pre-test revealed that an average of 45% of phase coherent face information was required on average to reach the 75% correct identification threshold (between 23% to 60%, congruent with the heterogeneity of famous faces). We constructed the *i*Hybrids with the constraint of a maximum relative phase coherence difference of 24% between any two identities. The face pairings were the same for all participants.

*i*Hybrid computation is time-intensive and so was pre-computed off-line for all identity pairs. First, we computed all possible fixation locations on the face (constrained to a lattice of 5*5 pixels on the original 260x260 pixel picture). Second, for each pair of identities (e.g. Brad Pitt and William Macy in Figure 1) we decomposed each face picture into 4 non-overlapping SF bands of 1 octave each (<3, 3-6, 6-12 and >12 cycles per deg of visual angle) using the Matlab Pyramid Toolbox (Simoncelli, 1997). For each one of the pre-computed fixation locations, we applied the spatial frequency filtering illustrated in Figure 1 for a fixation on the left eye. For the local information, a Gaussian window (of std = 25 pixels, ~ 1 deg of visual angle) extracted the information represented in the four SF bands at the point of fixation (the left eye of Brad Pitt in Figure 1, see Information: Local). For the global information, we applied the same technique, but extracting the information outside the Gaussian (see Information: Global). The iHybrid resulted from the sum of the local and the global information. For each identity pair, we generated two *i*Hybrids by swapping the face from which we sampled the local and global information.

The experiment started with a standard 9-point calibration. Then each experimental trial started with a check that the eye tracker was 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—followed by one final central fixation cross for drift correction. A drift superior to 0.5° would automatically launch a standard 9-point calibration to recalibrate the eye tracker. To avoid first fixation locations to be constrained by a fixation cross presented before each trial, the screen went blank for a random time (between 0.5 and 1s.) after the calibration check. This procedure ensured a random gaze position when the *i*Hybrid was displayed on the screen. We then randomly selected one of the *i*Hybrid and started the *i*Hybrid procedure. We monitored the observer's fixation location and used this information to index online the pre-loaded *i*Hybrids. Each *i*Hybrid trial lasted for 1 s., followed by a 100% noise field presented for 500 ms. Observer identified each iHybrid 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 local and global information).

Apparatus. We recorded eye movements with the SR Research Desktop-Mount EyeLink 2K (with a chin/forehead rest, sampling rate 1000Hz). Viewing was binocular, we tracked the participant's dominant eye, calibrated the system with a standard 9-point fixation procedure and validated and repeated calibration when necessary. Specifically, at the beginning of each trial, participants were instructed to fixate a dot at the centre of the screen to perform a drift correction. Calibration would restart if the drift correction was more than 0.5°. We implemented the experiment in Matlab using the Psychophysics and EyeLink Toolbox extensions (Brainard, 1997; Cornelissen, Peters & Palmer, 2002). Updating of the display contingent to gaze position required 11 ms on average (between 8 and 14 ms, about 90 Hz refresh rate on the screen, eliminating any flickering).

Results

Participants correctly identified one of the two identities composing the *i*Hybrid with 95% accuracy. We measured proportion of local versus global identification responses as a proxy for use of local versus global face information and observed a non-significant trend for reporting the global face (51%) over the local face (44%), $F(1,11)=3.74$, $p=0.08$, $p\eta 2=.25$.

Further examination of individual observers did not reveal any systematic individual differences in reporting the global vs. the local face. Faces with a lower phase coherence threshold (i.e. faces identified with less information) did not bias response. Examining whether phase coherence conditioned on local versus global identification responses lead 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 the fixation map using all trials of the Experiment—i.e. without splitting them according to a local vs. a global response, see Figure 2, Panel *a*. The overall pattern of fixations is commensurate with the famous triangular pattern joining the two eyes and the mouth in face identification experiments (Forbus, 1967). To understand the fixation biases associated with local and global information use, we decomposed the overall fixation map according to local vs. global behavioural responses (see Figure 2, panel *b*).

Number of fixations, mean fixation durations, path length and mean saccade length as a function of local vs. global behavioral responses did not reveal any general significant effect of local vs. global information use (see Table 1, respectively $F(1,11)=2.09$, $F(1,11)=0.02$, $F(1,11)=4.34$, $F(1,11)=1.75$, $ps>.05$).

= = = = = = = = = = = = =

TABLE 1

= = = = = = = = = = = = =

However, fixation locations as a function of local vs. global behavioral responses revealed significantly different patterns. To establish significance, we used *i*Map (Caldara &

Miellet, in press). Specifically, we applied a one-tail *Pixel test* (Chauvin, Worsley, Schyns, Arguin & Gosselin, 2005) for the group local and global fixation maps ($Z_{crit}>4.86$; $p<.05$, corrected) and a two-tail *Pixel test* on the difference between the local and global maps (Z_{crit} $|5.01|$; $p<.05$, corrected, see Figure 2, Panel *b*).

= = = = = = = = = = = = =

FIGURE 2

= = = = = = = = = = = = =

Figure 2, panel *b* reveals that observers fixated primarily the left eye and the mouth when identifying the local face of the *i*Hybrid, and primarily the centre of the face when identifying the global face of the *i*Hybrid. This difference in loci of fixations is confirmed by the Z-scored fixations durations in the significant areas (Local face: eyes-mouth = 9.83, centre = 3.15; Global face: eyes-mouth = 4.24, centre = 8.22) as well as the effect sizes between local and global strategies (Cohen's d: eyes-mouth = 1.4, centre = 1.75). We conducted 2 (strategy: local versus global) x 2 (feature location: eyes/mouth versus centre, determined as significant areas above) ANOVAs on average fixation duration, path length, total fixation duration and number of fixations and found interactions on mean fixation durations ($F(1,11)=9.03$, $p<0.02$, $p\eta 2=.45$), path length ($F(1,11)=16.53$, $p<0.002$, $p\eta 2=.60$), total fixation duration ($F(1,11)=34.57$, $p<0.0001$, $p\eta 2=.76$) and number of fixations ($F(1,11)=14.26$, $p<0.004$, $p\eta 2=.56$). The pattern of results confirms longer mean fixation durations, path lengths, fixation durations and larger number of fixations on the eyes and mouth area for the local strategy and on the centre face area for the global strategy (see Table 2). No other effect reached significance.

= = = = = = = = = = = = =

TABLE 2

= = = = = = = = = = = = =

The same observer would sometimes identify the same identity (e.g. Brad Pitt) in one *i*Hybrid using a local strategy as illustrated in Figure 2 *b* and in the counterbalanced *i*Hybrid still as Brad Pitt using this time 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 (only true in 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 versus the global face conditioned on location of the first fixation on the iHybrid. Figure 2, Panel *c* reveals that the probability of a local strategy was significantly higher than 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 a global strategy was significantly higher than a local strategy when the first fixation landed in the centre 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 centre of the face) determines respectively the choice of local or global information sampling strategy and the report of one vs. the other identity of the same iHybrid.

<u>Validation: Local vs. Global Face Identification.</u> To validate that the local and global fixated information are both sufficient for face identification in natural viewing conditions we performed the following validation with ten novel Western-Caucasian young adults with normal or corrected vision (4 males; mean age 28.3 years).

We used all original 36 identity pairings to create hybrid identities specifically with local and global information filters. We created a “local” and a “global” information filters from the significant local and global regions of the difference fixation map in Figure 2 *b* (reported for clarity in Figure 3 *a*). For each identity pair, the local (or the global) filter extracted local (or the global) information from one identity to which we added complementary information from the other identity by inverting the filter. Figure 3 *b* illustrates the process, applying the local filter to Brad Pitt and adding complementary William Macy information. Figure 3B also illustrates applying the global filter to William Macy and adding complementary Brad Pitt information. For each face pair, we thus reconstructed the four possible hybrid faces (local-Pitt, global-Pitt, local-Macy, global-Macy), for a total of 36 hybrid faces. The viewing parameters were identical to those used during the *i*Hybrid procedure and the stimuli were also presented for 1s. However, it is important to emphasize that observers now saw full spectrum static stimuli on the screen, not eye-contingent displays. Observers were instructed to identify the faces by naming them aloud.

= = = = = = = = = = = = =

FIGURE 3

= = = = = = = = = = = = =

We predicted that the reconstructed hybrids would be preferentially identified according to the local or the global information extracted by the filters, validating these cues for face identification in more naturalistic viewing conditions, but still in the context of hybrid identities. The data confirmed these predictions (see Figure 3, Panel *b*, $t(9)=18.20$, $p<0.001$). Furthermore, face identification performance demonstrated a very strong inter-observers agreement (Fleiss’ kappa=0.94). This validation with local and global information extracted using the fixation maps from the iHybrid procedure confirms that sufficient information was indeed extracted to locally or globally to identify the faces in the iHybrid procedure. It also confirms that the use of this information generalizes to more naturalistic (i.e. non-eye-contingent) viewing conditions.

Discussion

*i*Hybrids 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 whereas another observer could use a global strategy. We did not find an association between choice of strategy and available identity information (as measured with percentage of phase coherence) but, in contrast, demonstrated a strong link between the strategy selection and location of the initial fixation on the face. First fixations on the eyes and mouth lead to a local strategy while initial fixations in the centre of the face promoted a global strategy. Note that when collapsing fixations across the two strategies we recovered the well-known pattern of fixations over the two eyes and the mouth of Yarbus (1965). Furthermore, hybrid faces reconstructed according to the local vs. global sampling of information lead 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. We now examine the implications our findings for face identification.

The face identification literature presents long-standing questions on the specialized processing of faces compared to other objects and scenes (e.g., Kanwisher, 2000). Some argue that faces are processed as relatively undifferentiated wholes (i.e. *holistically*), without a differentiated and explicit representation of the individual facial features (Farah, Tanaka & Drain, 1995; Tanaka and Farah 1993). Others argue that faces are processed on the basis of both *configural* and *featural* information, with separate representational systems for both (e.g. Cabeza & Kato 2000; Sergent 1984). Configural processing would primarily use as information the metric distances between individual face features (for a review, see Maurer, Le Grand & Mondloch, 2002; though 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 and/or color of the eyes, the shape of the chin, nose, mouth and so forth) or combination of features (Schyns, Bonnar & Gosselin, 2002; Sekuler, Gaspar, Gold & Bennett, 2004; Schyns et al, 2009; VanRijsbergen & Schyns, 2010). Notwithstanding the conceptual difficulty of teasing apart these positions and empirically testing them, our results suggest that when the visual system locally extracts 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 the centre of the face, then this global information is probably sufficient for a direct comparison with memory. Thus, it is not surprising that the face identification literature is roaring with debates on representation formats (featural vs. holistic or configural): Sampling of face information with fixations suggests 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 preference for one or the other (for unfamiliar faces), and thus account for the complex pattern of data in this literature (see Althoff & Cohen, 1999; Groner et al., 1984; Henderson et al., 2005 vs. Blais et al., 2008;

Caldara et al., 2010; Jack et al., 2009; Hsiao & Cottrell, 2008; Kelly et al., 2010; Kita et al., 2010). Accordingly, a cultural bias for a global strategy in Easterners could arise from an equally effective general visual information extraction style used for object identification (Kelly et al., 2010), which might originate from a cultural norm to direct the first fixation less often directly towards 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. This questions a mandatory route to recover face identity.

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 & 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*, e3022. 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. (in press). *i*Map: A Novel Method for Statistical Fixation Mapping of Eye Movement data. Behavior Research Methods.

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the spotlight reveals universal information use for face recognition. *PLoS ONE, 5* (3): e9708. 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*, 659-667.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behavioral Research Methods, Instruments and 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* (14), R476-7.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception & 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 (Ed.), *Theoretical and applied aspects of eye movements research* (pp. 523-533). Amsterdam: 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* (8), 759-763.

Kelly, D. J., Miellet, S., & Caldara, R. (2010). Culture shapes eye movements for visually homogeneous objects. *Frontiers in Psychology*, *1*:6. doi: 10.3389/fpsyg.2010.00006.

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. doi:10.1371/journal.pone.001105.

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*(11), 1491-1503.

Schyns, P., & 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., 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., 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. Doi:10.1371/journal.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]”, available from the author: 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.* doi: 10.3389/fpsyg.2010.00159.

Van Rijsbergen, N. & Schyns, P.G. (2010). Dynamics of trimming the content of face representations for categorization in the brain. *PLoS Computational Biology.* 5(11): e1000561.

Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33* (2), 113-20.

Yarbus, A.L. (1965). *Role of eye movements in the visual process*. Moscow: Nauka.

Acknowledgments

SM, RC and PS are supported by The Economic and Social Research Council and Medical Research Council (ESRC/RES-060-25-0010).

Table 1.

Number of fixations, mean fixation duration (in msec.), path length (in deg.) and mean saccade length (in deg) in function of the *i*Hybrid identification strategy (local versus global).

	Number of fixations	Mean fixation duration (msec.)	Path length (deg.)	Mean saccade length (deg.)
Local strategy	3.25	391	6.74	1.88
Global strategy	2.99	386	5.78	1.73

Table 2.

Mean fixation duration (in msec.), path length (in deg.), total fixation durations (in msec.) and number of fixations in function of the *i*Hybrid identification strategy (local versus global) and the feature location (eyes/mouth versus center).

	Local Strategy		Global strategy	
	Eyes/mouth	center	Eyes/mouth	center
Mean fixation duration (msec.)	402	296	313	379
Path length (deg.)	1.46	0.40	0.61	1.07
Total fixation duration (msec.)	222	47	70	206
Number of fixations	0.58	0.18	0.24	0.56

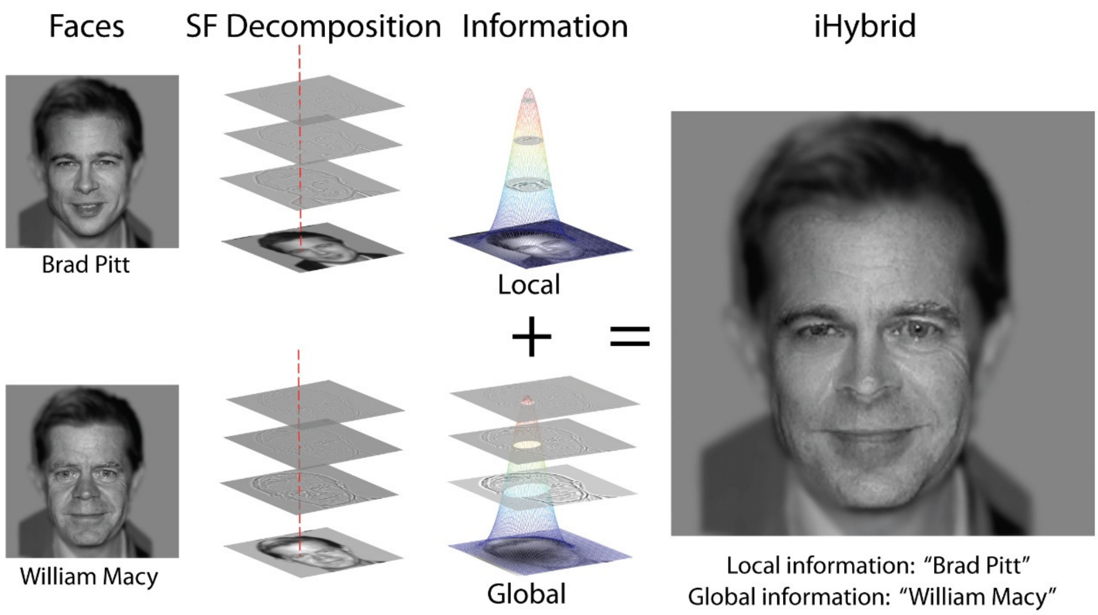
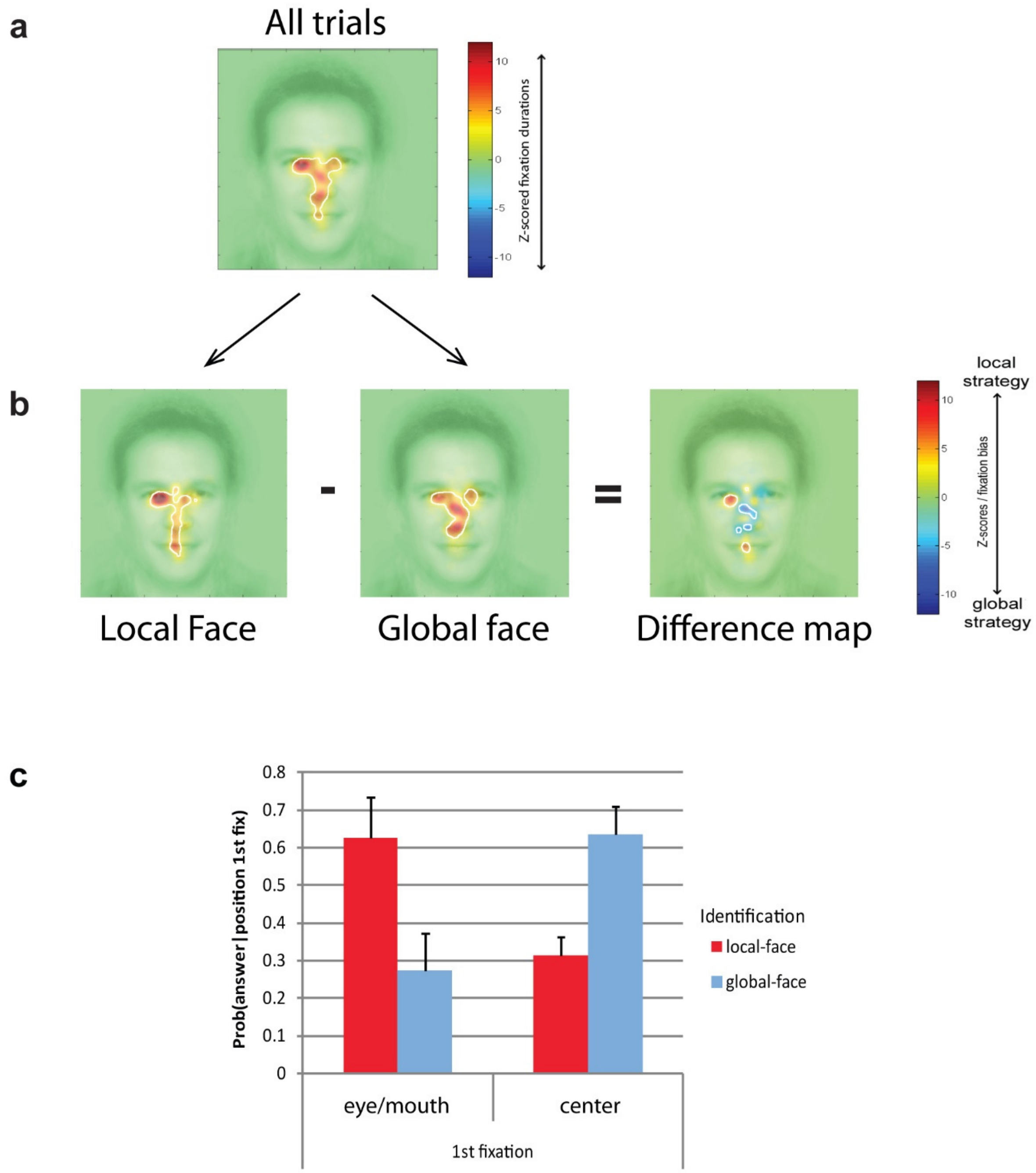

Faces
SF Decomposition
Information
iHybrid
Brad Pitt
Local
+
=
William Macy
Global
Local information: "Brad Pitt"
Global information: "William Macy"

Figure 1.



a
All trials
Z-scored fixation durations
b
Local Face
Global face
Difference map
local strategy
global strategy
Z-scores / fixation bias
c
Prob(answer|position 1st fix)
Identification
local-face
global-face
eye/mouth
center
1st fixation

Figure 2.

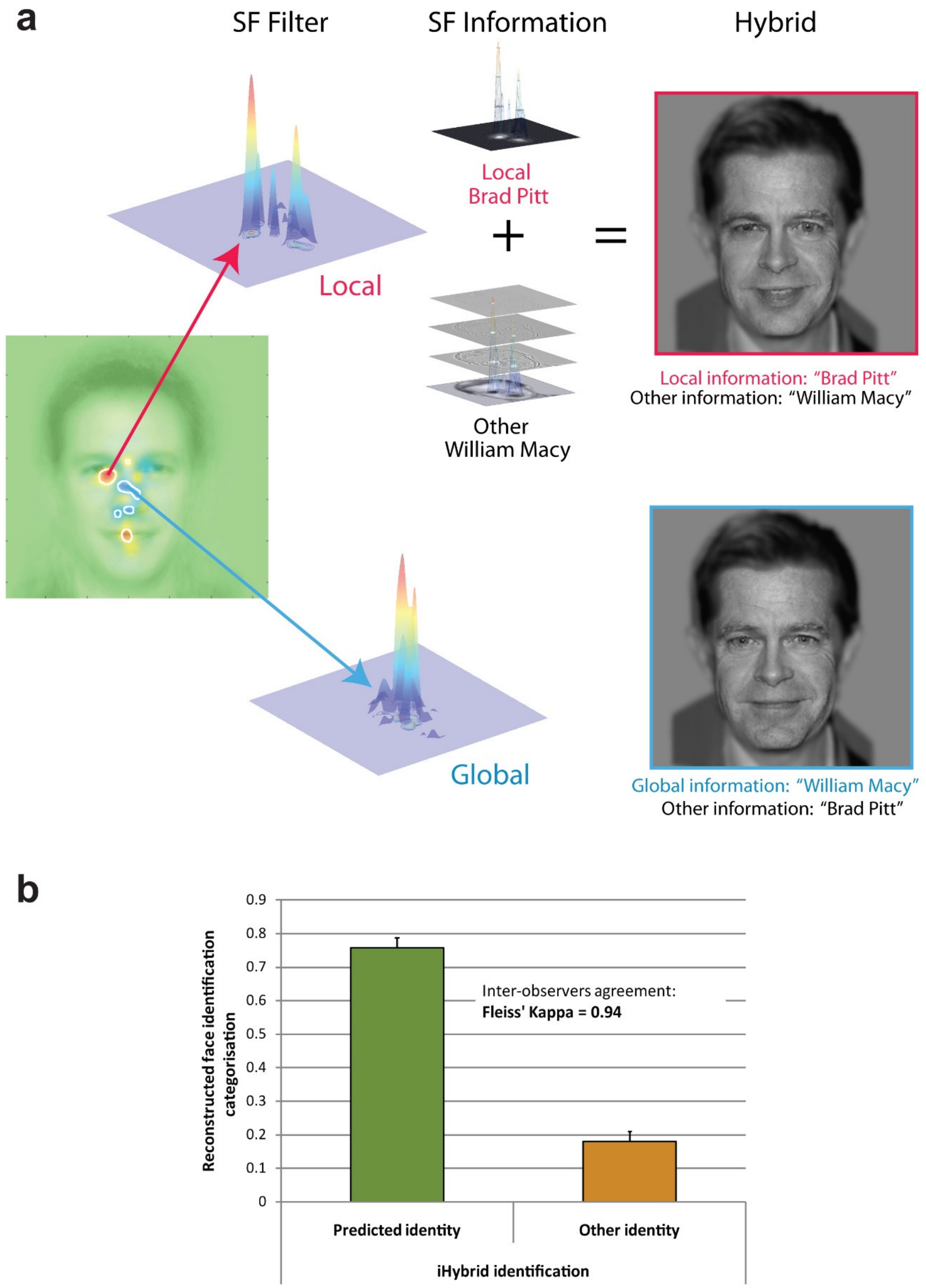
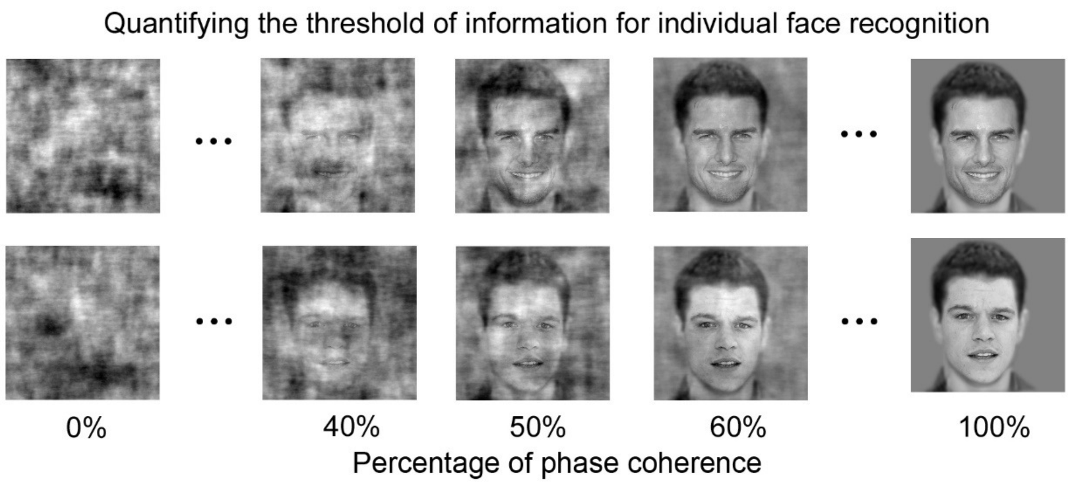

a
SF Filter
SF Information
Hybrid
Local
Brad Pitt
+
=
Local
Other
William Macy
Local information: “Brad Pitt”
Other information: “William Macy”
Global
Global information: “William Macy”
Other information: “Brad Pitt”
b
Reconstructed face identification categorisation
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Inter-observers agreement:
Fleiss' Kappa = 0.94
Predicted identity
Other identity
iHybrid identification

Figure 3.



Complementary Fig. 1