Distinct information critically distinguishes judgments of face familiarity and identity

Marie L. Smith[1], Blanka Volna[1], Louise Ewing[1,2, 3]

[1] Department of Psychological Sciences, Birkbeck College, University of London

[2] ARC Centre of Excellence in Cognition and its Disorders, School of Psychology, University of Western Australia

[3] School of Psychology, University of East Anglia

Corresponding author:

Marie L. Smith

Senior Lecturer

Birkbeck College

Malet Street

London

WC1E 7HX

marie.smith@bbk.ac.uk

Keywords: Information use, face recognition, familiarity, spatial frequency

Abstract

Accurately determining the familiarity of another and correctly establishing their identity are vital social skills. A considerable body of work has explored their perceptual and neural underpinnings and debate remains regarding whether they are dissociable, i.e., separable parts of a dual process, or different aspects of a common retrieval process. Less is known about the specific visual information that guides familiarity judgments and how this compares to the information used to identify a face by name. Here we sought to establish the critical information underlying participants' judgments of facial familiarity and identification. We created a new standardized stimulus set comprising 6 personally familiar and 12 unfamiliar faces and applied the Bubbles reverse-correlation methodology to establish the information driving correct performance in each task. Results revealed that markedly different information underlies familiarity and identity judgments. When categorizing familiarity, participants relied more upon lower spatial-frequency, broad facial cues (eye and face shape) than when categorizing identity, which relied on fine details in the internal features (eyes and mouth). These results provide novel evidence of qualitatively distinct information use in familiarity and identification judgments and emphasize the importance of considering the task set for participants and their processing strategy when investigating face recognition.

Introduction

Rapidly and accurately recognizing an individual face is a vital function for successful social interaction and more generally for efficiently determining friend from foe. When presented with a face, an individual may recognize it as known to them simply as a result of it being familiar or not (a familiarity decision) or they may establish the particular identity of the individual (an identification decision). Hierarchical models of face processing suggest that the familiarity decisions come first, with person-specific identification occurring later, because these representations take longer to access (Bruce & Young, 1986).

Although it is generally accepted that familiar face recognition is a fast process, researchers have only recently shown that familiarity decisions can also be made rapidly (at around 380ms, Ramon, Caharel & Rossion, 2011; Barragan-Barragan-Jason, Lachat & Barbea, 2012; Jason, Besson, Ceccalidi & Barbea, 2013) when participants are constrained to respond quickly and stimuli are presented for a short period of time. Without such constraints, when participants are free to view the faces as they choose, reaction times are slower and considerably more variable (up to 900ms, see Ramon, Caharel & Rossion, 2011 for a review). This variability may be due in part to methodological differences in the type of response required (e.g. go-no go, matching, key-press), or the nature of the familiar faces (e.g. personally familiar vs. famous faces), but as suggested by Barragan-Jason and colleagues (Barragan-Jason, Cauchoix & Barbeau, 2015) it may also reflect differences in the specific strategies employed by the participants to form their decision: familiarity or identification.

Neural evidence regarding the processing of facial familiarity also remains mixed. Some researchers observe effects of familiarity as early as 170ms following stimulus onset (e.g. Jemel, Schuller & Goffaux, 2009; Caharel, Ramon & Rossion, 2013; Barragan-Jason, Cauchoix & Barbeau, 2015;) while others report significant differences begin consistently later (e.g. Bentin & Deouell, 2000; Eimer & Gosling, 2011; Zheng, Mondloch & Segalowitz, 2012). As with the reaction time experiments, however, these differences may partly reflect methodological differences between studies. For example, given that most researchers do not specifically set out to constrain participants towards a particular recognition strategy (e.g. a response speed constraint is thought to encourage a mostly familiarity based strategy, Barragan-Jason et al. 2015) it is impossible to rule out potential differences as a result of this demand characteristic. To the extent that these two categorizations (familiarity, identification) are in fact dissociable neural processes i.e., the separate constituents of a dual recognition process (e.g. Curran & Hancock, 2007; MacKenzie & Donaldson, 2007; Burns, Tree & Weidemann, 2015) this is a potentially important distinction (though see also Yovel & Pallor, 2004 for contrary evidence suggesting that recollection and familiarity constitute different aspects of the same retrieval process).

Another key issue to elucidate is the specific visual information that participants draw upon when recognizing faces. It has long been established that some regions of the face are more informative than others and the internal features in particular (eyes, nose and mouth) are routinely found to be of primary importance (e.g. Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & Ellis, 1985). This pattern of internal feature use has been confirmed in more recent eye tracking studies of face familiarity judgments for both familiar and unfamiliar face exemplars (Althoff and Cohen, 1999; Stacey et al, 2005; Barton et al, 2006; Heisz and Shore, 2008; van Belle, Ramon, Lefevre & Rossion, 2010; Bonifacci, Desideri & Ottaviani, 2015). Perhaps surprisingly, however, many studies show very little difference in fixation patterns for familiar and unfamiliar faces (e.g. Althoff and Cohen, 1999; Stacey et al, 2005; Bonifacci, Desideri & Ottaviani, 2015). Where differences are observed they have been related more to the number of fixations made (6 vs. 9 for familiar vs. unfamiliar faces respectively, Barton et al, 2006) or in the very precise internal location of later fixations (after 4/5 fixations, van Belle et al, 2010).

Notably, when response time is manipulated to permit only one fixation to the face, a central nasion fixation point (in the centre of the face, just below the eyes) proves to be optimal for extracting task relevant information for identification, emotion and gender categorization tasks (Peterson & Eckstein, 2012). Recent evidence, further suggests that old/new face decisions (for newly learned faces) can actually be made accurately in as little as one or two fixations to this area at the top of the nose between the eyes (Hsiao & Cottrell, 2008). Here the authors artificially constrained the viewing patterns of their participants by removing the face stimuli after a certain number of fixations and found that additional fixations to the center of the eyes did not result in improved performance in this task. This finding suggests that the authors' constraints may have forced participants into using a more familiarity-based strategy, reliant on the whole face, unlike in other studies where unlimited time to explore the faces could encourage and/or facilitate decisions based on more detailed local identity specific information. This interpretation is yet to be confirmed empirically, however, because to date no studies have explicitly explored differences in the use of visual information in familiarity and identification tasks, or even tried indirectly to force a particular strategy by manipulating response time whilst concurrently tracking eye movements.

While the value of eye movement recordings in establishing the pattern of fixations made when viewing a face is clear, it is important to acknowledge that participants' information use cannot easily be inferred from this variable. Most critically, it is worth noting that eye movement recordings cannot definitively distinguish between face regions that are fixated for whatever reason (e.g. as a result of a generic response to seeing a face) and those regions that are actually important and used for a particular task (e.g., identification). Reverse correlation approaches, introduced by Haig (1995) in the context of face perception, provide an elegant alternative. Examples include the Bubbles approach, where randomly located information samples from a stimulus are presented to participants and the importance of information from different locations is established via their subsequent responses (Gosselin & Schyns, 2001) and added noise techniques, where image noise is added to a stimulus and the associated responses again reveal the important parts of the stimulus for the categorization tasking place (Sekuler, Gaspar, Gold & Bennett, 2004). In reverse correlation experiments participants are presented with visually degraded stimulus information and asked to make categorization decisions. Significant associations between the availability of specific subsets of visual information and correct categorization performance are then used to precisely assign the credit for categorization performance to specific visual information. An important additional advantage of the Bubbles reverse correlation approach is that information can be independently sampled from different spatial frequency bands when preparing the stimuli. In this way it is possible to go beyond differences in the use of specific visual features (e.g. eyes vs. mouth) to explore more subtle differences in the use of visual information across spatial frequencies, e.g., from fine details to coarser shape information, and the importance of the overlap of information across SF bands (e.g. see Liu, Collin, Rainville & Chaudhuri, 2000).

The few studies that have employed reverse correlation to explore face recognition have done so with an explicit face identification task (Haig, 1985; Gosselin & Schyns, 2001; Caldara et al, 2005; Butler et al, 2010) and again confirmed the importance of the eyes, nose and mouth region for correct identification performance. They have also confirmed the importance of the mid-band of spatial frequencies (8-25 cycles per face), which may provide the optimal information for judgments of face identity (see Costen, Parker, & Craw, 1994, 1996; Nasanen, 1999; Ruiz-Soler & Beltran, 2006). Crucially, however, no reverse correlation study (to the best of our knowledge) has explored the visual information driving familiarity decisions (i.e., without an explicit naming component). Here, we set out to do just that and establish the critical information underlying participants' judgments of facial familiarity and then to compare this directly to the same participants' performance on a more traditional identification task (i.e., explicitly naming the same familiar individuals).

To this end we created a new standardized stimulus set comprising 6 personally familiar faces and 12 unfamiliar faces from members of the teaching and research staff at Birkbeck College respectively. Many studies of face recognition employ newly learned or famous faces, rather than personally known individuals, presumably for practical reasons. There are, however, clear differences in the underlying processing of faces that are newly learned in a laboratory setting (often from the same repeated static image that is also used later at test) or famous (but unknown) faces and personally familiar faces. This latter category is more likely, for example, to have been viewed extensively and individuated from others across a broad range of real life, highly variable viewing conditions and personally relevant interactions (Tong & Nakayama, 1999; Carbon, 2008; Heisz & Shore, 2008). Final year students at Birkbeck College categorized subsampled versions of these faces by familiarity in one session, and by individual name in a second session. From their responses we then characterized the spatial frequency specific diagnostic visual information that is crucial for the correct categorization of these faces by familiarity and identity. In this way we could address for the first time if observers use qualitatively different visual information to perform these two categorization tasks.

We expect to replicate existing reverse correlation and eye-tracking results regarding the importance of the internal features i.e. the eyes (predominantly the left eye), nose and mouth for face identification. Predictions regarding the visual information driving face familiarity decisions are more difficult to make. There is reason to believe that the same visual features will prove to be important (e.g. van Belle et al, 2010) though the results of Hsiao & Cottrell (2008) point to a greater importance of global face information when establishing only familiarity. From an evolutionary perspective, it makes sense that observers would be able to establish a generic sense of familiarity from a reasonable distance. Thus we would expect to find that facial familiarity decisions are driven more by larger features represented in lower spatial frequency bands i.e. information that would be detectable at a greater distances (Loftus & Harley, 2005), than the higher spatial frequency detail that should deliver the more fine grained information used for explicit identification. Clear differences in the use of visual information in these two categorization tasks would point more towards familiarity and identification representing distinct neural processes. A single process model might, by contrast, be more likely to require that the same information is used for both familiarity and identification categorizations, but with a more refined use of a subset of the information in the identification task.

Methods

Participants

Sixteen healthy volunteers (4 male, age range 21-44, mean age 31) participated in both conditions of the experiment, which was approved by the School of Psychological Science Ethics Committee at Birkbeck College, University of London, in exchange for a small payment. All participants had normal or corrected to normal vision, were in their final year of study at Birkbeck College and gave informed consent. All participants completed both a familiarity judgment task (response options: familiar, unfamiliar, don't know), and a face identification task (response options: named buttons for each of the 6 identities, unfamiliar, don't know) and a face ratings task. Face familiarity and face identification tasks took place in testing sessions held on separate days (with the familiarity task always preceding the identification task to limit any contamination of true familiarity categorizations by exposure to the identification task).

Stimuli

Stimuli comprised greyscale images of neutral expression faces posed by members of the teaching, research and administration staff at Birkbeck College, University of London. All photographs were taken in the same location, under standardized lighting conditions at the same distance from the poser. Stimulus images were further standardized via translation, and rotation to align the center of the eyes. In line with similar studies in this area we removed hairstyle information from the faces via the use of an oval facemask. Importantly the remaining external features e.g. face shape and external contour information remained intact. Past research using the bubbles approach in a famous face identification task found hairstyle information was not used (Butler et al, 2010), and previous eye-tracking studies with hairstyle included do not find many fixations to this region (Althoff and Cohen, 1999; Blais et al., 2008). Posers were also asked to remove any other distinctive non-face information (e.g. piercings or spectacles). 6 familiar faces were selected from a set of 12 faces of active teaching staff (based on an informal survey of 5 individuals in the final year of study). 12 unfamiliar faces were selected from a set of 20 faces to be similar in age, shape and presence/absence of facial hair. All stimuli were normalized via the Shine toolbox to ensure equivalent contrast (Willenbockel, Sadr, Fiset, Horne, Gosselin, Tanaka, 2010). See Figure 1A for an example stimulus from the face set.

For every experimental trial, subsampled versions of these faces were created by randomly sampling visual information from the original images using circularly symmetric Gaussian apertures, ‘bubbles’ (see Gosselin & Schyns 2001).

Only information located behind these apertures was visible to the participant, and could inform their decision, the rest was hidden from view (see Figure 1 for an example stimulus). The random position of the apertures ensures that on every trial a different combination of visual information was presented to the participant. Across trials, this random sampling approximates a uniform sampling of the input information space and allows a non-biased exploration of the importance of all of the available visual information for the categorization task. The number of apertures (bubbles) was adjusted on a trial-per-trial basis to maintain 75% correct performance for each observer in the categorization task (minimum 40, maximum 250). More bubbles means that more information has been shown to the participant (to counter poorer performance).

As we had clear predictions regarding the importance of visual information from different spatial frequency bands for familiarity and identity categorisation judgments, we sampled visual information across a number of spatial frequency bands in addition to across different locations across the face. Thus on each trial the randomly selected face image was first decomposed into five non-overlapping spatial frequency (SF) bands of one octave each (74-37, 37-18.5, 18.5-9.25, 9.25-4.63, 4.63-2.31 cycles-per-face) using the Pyramid toolbox for MATLAB (Simoncelli, 1999) before each SF band was independently sampled with randomly positioned bubbles. The size of the bubbles was adjusted at each scale to reveal 6 cycles per aperture and the number of bubbles per scale was adjusted to ensure equivalent information sampling across each scale (i.e. larger but few bubbles sampled information from the coarser low spatial frequency scales, smaller but more bubbles sampled information from the fine detail in the high frequency scales). The sampled information across each band was then recombined to produce one experimental stimulus comprised of a mixture of high and low SF information in randomly determined locations (for full details on the methods and an illustration of the stimulus generation process see Gosselin & Schyns, 2001 or Butler et al, 2010). Thus on each trial participants saw a mixture of low, mid and high spatial frequency information from randomly selected locations across the face, and had to base their categorization decisions on this information.

Stimuli were projected on a light gray background to the center of a screen at a distance of 70cm from the participant so that the visual angle was 5.7° to match Gosselin & Schyns, 2001 and Butler et al, 2010).

Experimental Procedure

Participants completed 864 trials in each of the familiarity and identification tasks (always in this order) across 12 blocks in 2 separate 45 minute testing sessions. A randomly selected familiar or unfamiliar face was presented on each trial (with familiar faces presented in half the experimental trials (72 repetitions per face) and unfamiliar faces in the remaining trials (36 repetitions per face)). To maintain concentration and motivation, short breaks were provided every 3-4 minutes (72 trials) consisting of generic motivational screens (e.g. "keep up the good work", odd numbered blocks) or interactive "puzzle-bubble games"[1] (even numbered blocks).

For both tasks the trial began with a 500ms fixation cross, which was immediately followed by a randomly selected face picture whose information was revealed by the bubbles apertures positioned across the locations of the face

---

[1] Participants were challenged to verbally label 'bubbled' images of films, TV shows or geographical locations (category = participant's choice) before and after the progressive addition of 'clues' that revealed additional features.

in the five spatial frequency bands. The sub-sampled face remained onscreen for 500ms and was replaced by a uniform grey screen until a response was given. Participants were instructed to respond as quickly as possible without making mistakes by pressing the appropriate labeled button of the computer keyboard. In the familiarity task participants were asked to respond to the general familiarity of the faces shown by pressing keys labelled as ‘familiar’, ‘unfamiliar’ or ‘don’t know’ (if they couldn’t decide from the information given). For the identification task, participants categorized the faces by name (appropriate labeled key for each of the six known identities), ‘unknown face’ or ‘don’t know’ (if they couldn’t decide from the information given). Participants were considered to be correct in the identification task only if they accurately identified the individual shown, selecting the wrong named key was considered an incorrect response. Don’t know responses were treated as incorrect responses. A short training session at the start of both experimental testing sessions ensured that participants could categorize the familiar face stimuli and four randomly selected unfamiliar face stimuli, presented without bubbles, correctly by familiarity / named identity as a function of the task being completed and ensured that participants were experienced in the response keys to be used.

In the first testing session, prior to the bubbles tasks, all participants were presented with non-standardised full-colour images of the familiar faces that would later be used in the experiment. The images were taken from publicly available sources e.g. staff pages on the Birkbeck College website, and participants were asked if the individuals were familiar to them and to write down their reason for knowing the individual (typically the name of the class that individual taught). Participants then viewed the standardized, oval-masked, gray scale images of the familiar faces alongside the twelve unfamiliar faces and rated them all for familiarity on a scale of 1 to 7.

Results

Performance Metrics

As expected, participants rated the familiar face stimulus set as much more familiar than the unfamiliar face stimulus set (Mean rating=5.82 vs. 1.24, t(15) = 23.4, p<0.001) confirming the familiarity of the known faces and unfamiliarity of the others.

For the Bubbles tasks, the staircase algorithm ensured that there was no significant difference in overall categorization performance across the two tasks (Mean percentage correct 74% vs. 74%, t(15)=0.475, p = 0.64). There was also no difference in performance as a function of the familiarity of the stimuli in either task (2 way repeated measures ANOVA for performance correct in the two tasks (familiarity and identity) and for the two classes of stimuli (familiar and unfamiliar) all F<1.1, p>0.3, see Table 1). However, slightly more information was required to achieve the same level of performance in the familiarity task compared to the identity task (median number: 118 vs. 89 bubbles, t(15) = 3.08, p=0.008)[2], which signals that participants found this categorisation judgment

---

[2] As the task performed is conflated with task order (the familiarity task always came first to avoid setting participants up in an identification strategy throughout), we cannot conclude whether or not the requirement for less information in the identity task is merely an effect of habituation to the

slightly more difficult (Royer et al, 2015). For the familiarity task, there was a significant difference in response times on correct trials as a function of the familiarity of the stimuli (t(15) = 3.58, p=0.003) with participants significantly faster to respond to familiar than unfamiliar faces (mean reaction times: 714 vs. 792ms). Comparative interpretation of reaction times in the identification task is challenging given the relatively more complex response demands (8 response keys, 6 representing the known identities cf., 2 response keys, representing familiar vs. unfamiliar). Nevertheless a main effect of familiarity was also observed here (t(15) = 2.96, p=0.01) with participants faster to indicate an unknown face than the familiar identities (1164 vs. 1001ms). It is also worth noting that participants' classifications of a face as unfamiliar were faster when performing the familiarity task (792ms) than when performing the identification task (1001ms, t(15) = 3.75, p =0.002), despite equivalent levels of performance accuracy in what is conceptually the same categorization decision (rejecting that a face is known to the participant).

Finally, there was no difference in use of the don't know response button across the two tasks (t(15) = 1.15, p=0.26) with on average 6.7% vs. 8.7% don't know responses in the familiarity and identification tasks respectively.

Bubbles Analysis

---

procedure, or truly reflects a difference in terms of information need. Considering the last 25% of trials only, by which time participants should have habituated to the slightly unusual Bubbles paradigm, there remains a significant difference in information requirement (t(15) = 2.25, p=0.04, 114 vs. 89 bubbles).

After the experiment, independently for each task, every trial was sorted as a function of whether or not the information presented to the participant resulted in a correct response in the categorization task. Observers will tend to be correct if the information necessary to perform the task has been provided to them and inversely they tend to be incorrect if this information is missing. To determine the specific information driving correct categorizations for familiarity decisions we summed together all of the bubbles leading to correct categorizations (e.g. in the familiarity task familiar as familiar, unfamiliar as unfamiliar) and divided that by the sum of all bubbles presented to generate probability maps. In these probability maps the pixel value at each location indicates the probability that presenting the visual information at that location would result in a correct familiarity categorization. Following Chauvin et al, 2005, we transformed these probabilities into z-scores using the non-informative region outside of the face image as a baseline and established those regions statistically associated with correct categorization performance by applying a $p<.05$ threshold and $p<.05$ cluster extent criterion (see Chauvin et. al., 2005 for full details of the specially developed statistical tests). To visualize the resulting diagnostic information we selected a representative face and revealed only the information in each spatial frequency band found to be significantly associated with correct categorization performance.

Figure 1A shows the diagnostic information used to correctly complete each task alongside the original face image. The distribution of this information across the spatial frequency bands is indicated by the bar charts in Figure 1B and represents the total number of significant pixels in each spatial frequency band (normalised by the size of the Gaussian bubble used to sample that band, see

Smith & Schyns, 2009). It can clearly be seen that the mid-band of spatial frequencies is of most importance to accuracy in both tasks. However, this is supported by higher spatial frequency details for the identity task, and relatively more low spatial frequency information in the familiarity task.

Turning to the visual features used and their split across spatial frequency bands (see Figure 1C), in the identity task participants focus on the left eye and mouth information in high spatial frequencies and continue to use primarily use the internal face area (eyes, nose and mouth) in the lower spatial frequencies. In the familiarity task, as in the identity task, participants also make some use of the higher spatial frequency detail around the left eye and mouth, though to a much lesser extent. Their strategy for information use in the mid band is markedly different to the identity task. Unlike the identity task when the central, classic t-shape portion of the face is used (Yarbus, 1967), here participants used the full extent of the eye shape along with external facial shape cues. This broad pattern of information use continues in the lowest two spatial frequency bands, with a global whole face region significantly linked to good categorization performance.

To compare information use across the two tasks, we computed the difference of the un-thresholded z-scored maps (to ensure that no features were missed by not quite reaching threshold). These differences were re-normalised to the baseline region and we applied the $p<0.05$ threshold corrected criterion on the resulting differences. Figure 1D indicates visual information that is significantly more used in the familiarity task (top row) and the identity task (bottom row) and confirms this general pattern of results (although we note that there is no significant difference in the use of the left eye in the second spatial frequency band).

Discussion

The main objective in the current study was to examine how task demands during categorisation judgments of known and unknown faces (i.e., asking participants to evaluate familiarity vs. identification) shapes their use of visual information. To that end, we investigated whether there are any differences in the specific visual cues and/or spatial frequency bands that critically inform these judgments. Overcoming the potential limitations associated with using newly learned or famous faces, we created a stimulus set of faces that were personally familiar to all our participants. This design choice allowed us to probe participants' robust memory representations of each individual: established and built up over real-life interactions.

Our results are broadly consistent with those of past studies (where these exist) of familiarity and identification judgments. For example, for the familiarity task performance metrics showed that participants made faster familiarity decisions for familiar faces (vs unfamiliar faces, e.g., van Belle et al, 2010) and for the identity task, classification images confirmed the importance of the mid-band of spatial frequencies for correct face recognition performance (Costen, Parker, & Craw, 1994, 1996; Nasanen, 1999; Ruiz-Soler & Beltran, 2006). The current findings also critically extend our understanding of how these judgments are made. For the first time we empirically confirm that marked differences exist in the specific visual information forming these categorizations, both in terms of the visual cues (features) and their representation across the spatial frequency bands tested.

For identification judgements, participants made most use of the classic internal features (eyes, mouth and nose; Yarbus, 1967). There was also a relatively greater contribution of the fine detail in high spatial frequencies than the coarse shape outlines of the low spatial frequencies. This pattern of information use when labelling the identity of personally familiar faces is very much in line with previous studies with newly learned (Gosselin & Schyns, 2001; Caldara et al, 2005) and famous faces (Butler et al, 2010), as well as eye tracking findings with personally familiar faces (van Belle, Ramon, Lefevre & Rossion, 2010). This close alignment with extant research findings weakens any suggestion that the bubbles paradigm altered the processing strategy of the participants in the current study. The bias that we observe to the left side eye is also well documented (e.g. Mertens, Siegmund & Grusser, 1993; Gosselin & Schyns, 2001; Butler & Harvey, 2006) and may stem from the dominance of the right hemisphere of the brain for processing faces (Burt & Perrett, 1997; Rossion et al, 2003), whereby a right hemisphere lesion is necessary and sufficient for acquired prosopagnosia (an inability to recognise familiar faces, Hecaen & Angelergues, 1962; Meadows, 1974). Despite evidence suggesting that the mouth may be more important for newly learned faces as opposed to known famous faces (Butler et al, 2010), we replicated the majority of findings confirming the use of this visual feature in face identification (Gosselin & Schyns, 2001; Caldara et al, 2005; van Belle et al, 2010).

For familiarity judgements, participants made much greater use of low spatial frequency information. In contrast to the identity task, participants made only minimal use of the higher spatial frequency feature-specific details (e.g., the fine detail in the eyes). Instead, categorization performance relied primarily on the extraction of broader face shape cues available in the mid and lower spatial frequency bands. Moreover these same participants that focused on the internal t-shape facial features in the identification task, here relied upon face shape cues from the cheeks, external face contour and eye shape information to judge familiarity. This distinct information processing strategy may suggest that participants consider the face more globally when judging familiarity (perhaps in a more holistic manner (e.g. Farah et al, 1995; Goffaux et Rossion, 2006) though the link between spatial frequencies and holistic processing is not a clear-cut one (e.g. see Rossion et al, 2013). It is clear, however, that when judging familiarity participants are extracting information that could be gleaned at greater distance from the individual rather than seeking out detailed local featural information (perhaps more relevant for confirming a specific identity) which would become available as the individual moves physically closer.

Our results for the familiarity task are broadly in line with the findings of an eye-tracking study by Hsiao & Cottrell (2008). These authors reported that participants could classify newly learned faces as old (vs. new) from as little as two fixations to the central upper region of the face when they were restricted in the number of fixations they could make (after a set number of fixations the stimulus was replaced by an average face mask). Fixation to a central location where there is a perceptual span large enough to cover the whole stimulus (e.g., on the nose or between the eyes) as opposed to the features themselves, is argued to represent the most informative “center of information” for such a task, which relies on a whole-face representation (Dailey & Cottrell, 1999).

Given that categorization task was necessarily conflated with task order (we always ran the familiarity task before the identification task), it is important to rule out the possibility that participants’ experience with the familiar faces during participants' first testing session was driving the differences in information use we observed during session two. We stress that any additional 'training' with each (already familiar) familiar face would have been minimal, limited to a maximum of 36 seconds of exposure to partial versions of each identity (revealed through randomly positioned bubbles on each trial)[3]. Nevertheless we ran a new group of 12 new participants from the same final year student population (4 male, mean age 33.4 [22-54]) through the same identification task, without the preceding familiarity task. Crucially, we observed no difference in the performance of this second group of participants in comparison to those individuals who completed the identification task after the familiarity session. They indicated strong familiarity with the stimuli (mean familiarity rating for the familiar vs. unfamiliar faces was 4.9 vs. 1.15, t(11) = 15.6, $p<0.001$), performed the bubbles task with equal accuracy (74.4% vs. 75% correct, t(26)=0.4, p=0.69) and required the same amount of information to do so (88 vs. 106 bubbles, t(26) = 0.92, p=0.37). Furthermore, like our original participants their information use was driven by high spatial frequency information (left sided eye, center of the mouth) with minimal contribution from the lower spatial frequency bands (see Supplementary Figure 1). Taken together, we propose that these results make it difficult to conclude that the differences in information use observed in the familiarity and identification tasks were driven simply by any additional training or the fixed task order.

In recognition memory research generally, and face perception particularly, there is an ongoing tension between single and dual process models

---

[3] In the familiarity task there were 72 trial repetitions of each familiar identity (36 per unfamiliar), in which different parts of each face were presented, for a short duration (500ms).

of recognition. In dual process models, it is proposed that *familiarity* (the general feeling of knowing a face, without any additional contextual details) is supported by a different processing route than *recollection* (where contextual detail is explicitly remembered). In contrast, single process models assume a single processing route that serves both familiarity and recollection appraisal (Yovel & Paller, 2004; Curran & Hancock, 2007; MacKenzie & Donaldson, 2007; Burns, Tree & Weidemann, 2014). **Our new evidence that participants use qualitatively different information when performing familiarity and identification categorizations on the surface appears more consistent with a dual process model. However, a single process model, whereby identification decisions naturally follow familiarity decisions due to the requirement to accumulate more detailed information to support the former, could also account for our results.**

In fact, it has recently been proposed that the processing of facial familiarity is a multi-stage process in which initial crude familiarity distinctions can take place rapidly (as early as 140ms following stimulus onset), but improved performance requires subsequent refinement at later stages (Barragan-Jason et al, 2015). This proposal is very much in line with the classical hierarchical models of face recognition (Bruce & Young, 1986). Barragan-Jason and colleagues (2015) report that participants in a speeded familiarity judgment task (with famous faces) can be split into two groups as a function of the presence or absence of early (170ms) neural differences to familiar vs. unfamiliar faces. Participants with earlier neural discrimination of familiar and unfamiliar faces produced a faster reaction time distribution but more false alarms than did participants whose neural activity did not discriminate until significantly later (285ms) but who showed more accurate discrimination. Behavioural evidence that participants naturally fall into fast vs. slow responders in familiarity tasks (Barragan-Jason et al, 2012, 2013) further supports the notion that familiarity decisions may well be a result of multiple processes taking place at different neural levels, with individual observers naturally making use of different processing routes to accomplish the same tasks. Van Belle et al (2010) also propose a multi stage familiarity decision process, in which the visual system initially extracts sufficient information within a few fixations to form basic old/new distinctions. Further fixations to detailed feature information are then required in order to confirm a match to a particular identity representation.

These models of familiarity processing sit nicely with theoretical models that posit a coarse to fine hierarchy of information processing in the visual system. Here, the suggestion is that the coarse structure of an exemplar is first extracted from low spatial frequencies and then used as an index into the fine detail available at higher spatial frequencies (e.g. Marr, 1982; Sergent, 1986, Bar, 2007). Such coarse to fine processing has been observed in face selective regions of cortex during face processing tasks (Goffaux et al, 2011). When bandpassed low, mid and high spatial frequency faces were presented for different durations in an fMRI study, Goffaux and colleagues (2011) clearly identified later processing of high spatial frequency information in the fusiform face area and right occipital face regions (areas commonly thought to represent identity, Haxby, Hoffman, & Gobbini, 2000). More recently, Ramon and colleagues (2015) recorded the neural response to familiarity categorizations of personally familiar faces as they were gradually revealed in a coarse-to-fine display paradigm. While activity in ventral occipitotemporal face selective brain regions responded to increasingly detailed face information, it was the structures of the medial temporal lobe (perirhinal cortex, amygdala, hippocampus) that responded categorically when sufficient information to correctly establish familiarity was accrued. An interesting direction for future research could be to combine reverse correlation with electroencephalographic recordings (e.g. as in Smith et al, 2009) or fMRI (Smith et al, 2008) to shed more light on the precise information processing functionality of face selective brain regions and the time-course of visual information processing that are associated with face recognition processes in familiarity and identification tasks.

An alternative (albeit less interesting) interpretation of the current results is that the differences in information use we observe are driven by differences in the levels of chance error associated with 3 vs. 8 response keys in the familiarity and identity tasks respectively. This discrepancy in the number of response options is near-unavoidable when trying to ensure that participants are performing the identification task correctly (and distinctly from the familiarity judgment task), i.e., by fully individuating each stimulus. In the current study it is important to note that participants performed equally well in both tasks (correctly assigning familiar faces to that response key, and correctly identifying each individual by name) and at the level set by the staircase algorithm (75% correct). Participants also had the option of a ‘don’t know’ response in both tasks, a non-standard addition to the bubbles paradigm, and were encouraged to use it. As such, they were not forced to make chance errors when they felt that they could not actually perform the task, but rather use this key. Finally, to the extent that our findings: are in line with the predictions of long-standing and widely accepted theories of face perception; fit broadly with other similar studies of identification tasks using famous faces with a verbal response (Butler et al, 2010) and newly learned faces with multiple response keys (Caldara et al, 2005, Schyns & Gosselin, 2001); and have been replicated with a new set of participants (identification task alone, see Supplementary Figure 1), we contend that an explanation of this sort seems unlikely.

A current limitation of reverse correlation techniques is the requirement for uniformity in the location of visual information in stimulus image space (e.g. the eyes should be in roughly the same position in image space in each image so that repeated use of this information across exemplars is clear when one conducts the reverse correlation analysis). This unfortunately precludes the effectiveness of using different exemplars of the same individual and necessarily requires the repetition of stimuli across trials. Butler et al (2010) elegantly sidestepped this issue by using a large database of famous face, so that their familiar faces need not be repeated. Although they replicated many of the findings of similar studies with newly learned faces (e.g. Gosselin & Schyns, 2001; Caldara et al, 2005), their results indicated less importance of the mouth in identifying famous faces and one explanation of this may been the lack of stimulus repetition. An interesting extension of the current study would be to employ such a large stimulus set in a familiarity categorization task to explore any effect of stimulus repetition on the information used to establish familiarity from faces. Caution would be needed in interpreting the results, however, because such stimuli would also differ from those in the current study with regards to their *type* of familiarity to participants (famous vs. personally familiar).

The current study presents unique, novel evidence of variability in the specific visual information that is attended to and processed in order to make familiarity and identification judgments of the same face stimuli. From a methodological standpoint, these results highlight the importance of considering task demands in face research. Regions in face-sensitive cortex (fusiform gyrus, inferior occipital gyrus, and superior temporal sulcus) are known to respond as a function of the type of processing taking place (e.g. featural vs. configural), rather than as a function of changes in the stimulus per se (e.g., Cohen-Kadosh et al, 2009). In the light of this, researchers should carefully consider the task asked of participants and the explicit strategy they will employ to achieve it. In not doing so, they risk encouraging systematic processing biases that could cloud clear understanding of face recognition processes.

References

Alhoff, 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.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions*. Trends in Cognitive Science*, 11, 280-289.

Barragan-Jason, G., Besson, G., Ceccaldi, M., & Barbeau, E. J. (2013). Fast and famous: Looking for the fastest speed at which a face can be recognised. *Frontiers in Psychology*, 4, 100. http://dx.doi.org/10.3389/fpsyg.2013.00100

Barragan-Jason, G., Chauchoix, M., & Barbeau, E. J. (2015). The neural speed of familiar face recognition. *Neuropsychologia*, 75, 390-401.

Barragan-Jason, G.,Lachat, F., & Barbeau, E. J. (2012). How fast is famous face recognition? *Frontiers in Psychology,* 3, 454. http://dx.doi.org/10.3389/fpsyg.2012.00454

Barton, J. J., Radcliffe, N., Cherkasova, M.V., Edelman, J. & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion and morphing on scanning fixations. *Perception*. 35, 1089-1105.

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology*, 17, 35-55.

Bonifacci, P., Desideri, L., & Ottaviani, C. (2015). Familiarity of faces: Sense or feeling? An exploratory investigation with eye movements and skin conductance. *Journal of Psychophysiology*, 29, 20-25.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305-327. http://dx.doi.org/10.1111/j.2044-8295.1986.tb02199.x

Burns, E. J., Tree, J. J., & Weidemann, C., T. (2015). Recognition memory in developmental prosopagnosia: Electrophysiological evidence for abnormal routes to face recognition. *Frontiers in Human Neuroscience,* 8, 622.

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

Butler, S., & Harvey, M. (2006). Perceptual biases in chimeric face processing: Eye-movement patterns cannot explain it all. *Brain Research*, 1124, 96-99.

Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgments of facial attractiveness, age, gender, speech and expression. *Neuropsychologia,* 35, 685-693.

Caharel, S., Ramon, M., & Rossion, B. (2013). Face familiarity decisions take 200ms in the human brain: Electrophysiological evidence from a go/no-go speeded task. *Journal of Cognitive Neuroscience,* 26, 81-95.

Caldara, R., Schyns, P.G., Mayer, E., Smith M. L., Gosselin, F., & Rossion, B. (2005). Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *Journal of Cognitive Neuroscience*, 17, 1652-1666.

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

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. doi:10.1167/5.9.1

Cohen Kadosh, K., Henson, R. N. A., Cohen Kadosh, R., Johnson, M. H. & Dick, F. (2009). Task-dependent activation of face sensitive cortex: An fMRI adaptation study. *Journal of Neuroscience, 22,* 903-917.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high pass and low pass spatial filtering on face identification. *Perception & Psychophysics*, 58, 602-612.

Curran, T., & Hancock, J. (2007). The FN400 indexes familiarity-based recognition of faces. *Neuroimage,* 464-471.

Dailey, M. N., & Cottrell, G. W. (1999). Organization of face and object recognition in modular neural networks. *Neural Networks, 12,* 1053-1074.

Eimer, M., & Gosling, A. (2011). An event related potential study of explicit face recognition. *Neuropsychologia*, 49, 2736-45.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979) Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face perception. *Perception*, 8, 431-439.

Farah, M. J., Wilson, K.D., Drain, M., & Tanaka, J. N. (1995). What is 'special' about face perception? *Psychological Review, 105, 482-498.*

Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2011). From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cerebral Cortex,* 21, 467-76. doi: 10.1093/cercor/bhq112

Goffaux, V., & Rossion, B. (2006). Faces are 'spatial' – holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1023-1039.

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

Haig, N. D. (1985). How faces differ – a new comparative technique. *Perception,* 14(5), 601-615.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I., (2000). The distributed human neural system for face perception. *Trends in Cognitive Science,* 4, 223-233.

Hecaen, M. D. & Angelergues, M. D. (1962). Agnosia for faces (Prosopagnosia). *Archives in Neurology,* 7(2), 92-100.

Heisz, J.J. & Shore, D. I. (2008). More efficient scanning for familiar faces. *Journal of Vision*, 8, 1-10.

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

Jemel, B., Schuller, A., M., & Goffaux, V. (2010). Characterising the spatio-temporal dynamics of the neural events occurring prior to and up to overt recognition of famous faces. *Journal of Cognitive Neuroscience,* 22, 2289-2305.

Liu, C. H., Collin, C. A., Rainville, S. J. & Chaudhuri, A. (2000). The effects of spatial frequency overlap on face recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 26, 956-979.

Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identity someone close than far away? *Psychonomical Bulletin & Review,* 12, 43-65.

MacKenzie, G., & Donaldson, D. I. (2007). Dissociating recollection from familiarity: Electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *Neuroimage*. 36, 454-463.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco (CA): Freeman.

Meadows, J.C. (1974). The anatomical basis of prosopagnosia. *Journal of Neurology, Neurosurgery & Psychiatry,* 37(5), 489-501.

Mertens, I., Siegmund, H., & Grusser, O. J. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia,* 31, 989-998.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, 39, 3824-3833.

Peterson, M.F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings in the National Academy of Sciences,* 109:48, 3314-3323.

Ramon, M., Caharel, S., & Rossion, B. (2011). The speed of recognition of personally familiar faces. *Perception*, 40, 437-439.

Ramon, M., Viziolo, L., Liu-Shuang, J., & Rossion, B. (2015). Neural microgenesis of personally familiar face recognition. *Proceedings of the Natural Academy of Sciences,* 112(35), 4835-4844.

Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic face perception. *Visual Cognition,* 21, 139-253.

Rossion, B., Joyce, C. A., Cottrell, G.,W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word and object processing in the visual cortex. *Neuroimage, 20, 1609-1624.*

Royer, J., Blais, C., Gosselin, F., Duncan, J. & Fiset, D. (2015). When less is more: The impact of face processing ability on the recognition of visually degraded faces. *Journal of Experimental Psychology: Human Perception and Performance*.

Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research*, 70, 273–292.

Simoncelli, E. P. (1999). *Image and multi-scale pyramid tools.* Computer Software.

Sekuler, A. B., Gaspar, C., Gold, J.M. & Bennett, P. J. (2004). Inversion leads to quantitate, not qualitative, changes in face processing. *Current Biology*, 14, 391-396.

Sergent, J. (1986). Microgenesis of face perception. *Aspects of Face Perception,* 28, 17-33.

Smith, M.L., Fries, P., Gosselin, F., Goebel, R., & Schyns, P.G (2009). Inverse Mapping the Neuronal Substrates of Face Categorizations, *Cerebral Cortex*, 19, 2428-2438.

Smith, F. W., Muckli, L., Brennan, D., Pernet, C., Smith, M. L., Belin, P., Gosselin, F., Hadley, D., Cavanagh, J., & Schyns, P.G. (2008). Classification Images reveal the Information Sensitivity of brain voxels in fMRI. *NeuroImage*, 40, 1643-1654.

Smith, F.W., & Schyns, P. G. (2009). Smile through your fear and sadness. *Psychological Science,* 20, 1202-1208. doi: 10.1111/j.1467-9280.2009.02427.x

Stacey, P. C, Walker, S. & Underwood, J. (2005). Face processing and familiarity: evidence from eye movement data. *British Journal of Psychology*. 96, 407-422.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence for visual search. *Journal of Experimental Psychology: Human Perception and Performance,* 25, 1016-1035.

van Belle, G., Ramon, M., Lefevre, P. & Rossion, B. (2010). Fixation patterns during recognition of personally familiar and unfamiliar faces. *Frontiers in Psychology*. 1, 1-8.

Willenbockel, V. Sadr, J. Fiset, D. Horne, G. Gosselin, F., & Tanaka, J. (2010). The SHINE toolbox for controlling low-level image properties. *Behavior Research Methods*, 42, 671-684.

Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Plenum Press. (Translated from Russian by Basil Haigh. Original Russian edition published in Moscow in 1965.)

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

Yovel, G., & Pallor, K. A. (2004). The neural basis of the butcher on the bus phenomenon: when a face seems familiar but is not remembered. *Neuroimage* 21, 789-800.

Zheng, X., Mondloch, C. J., & Segalowitz, S. J. (2012). The timing of individual face recognition in the brain. *Neuropsychologia*, 50, 1451-1461.

Figure Caption

Figure 1. A. The original stimulus along inside the visual information used by participants to correctly categorize faces by familiarity and identity. B. The degree to which information from each of the five spatial frequency bands samples is used in the categorization of familiarity and identity (1=high spatial frequency, 5 = low spatial frequency). C. The breakdown of information used in each task in each spatial frequency band. D. The visual information used significantly more in the familiarity task than the identity task (top row) and used more in the identity task than the familiarity task (bottom row).

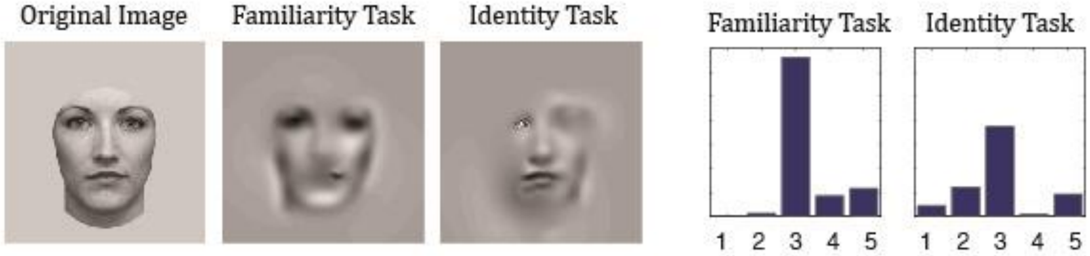
Supplementary Figure Caption

Supplementary Figure 1. The visual information used by participants in the second group to correctly categorize faces by identity (left), alongside the breakdown of information used in each task in each spatial frequency band.

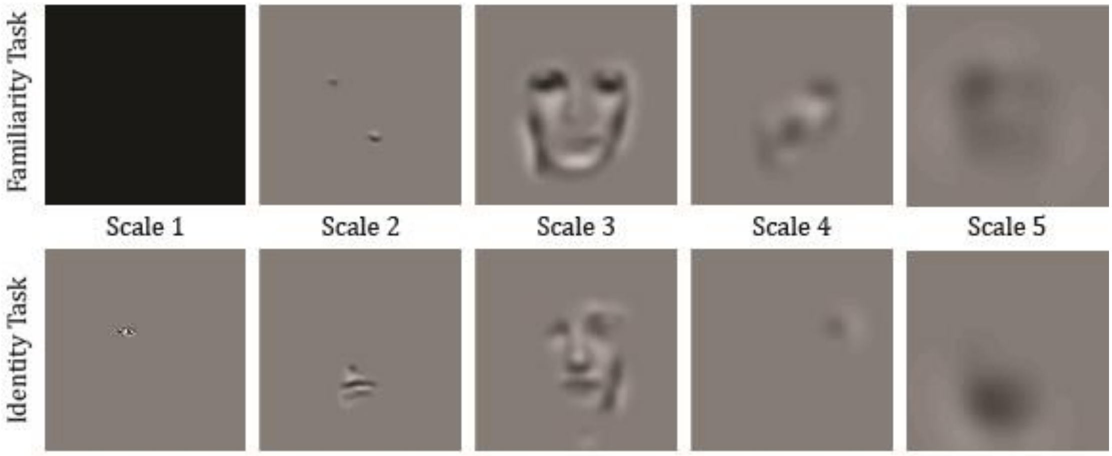
Figure 1

A. Information use

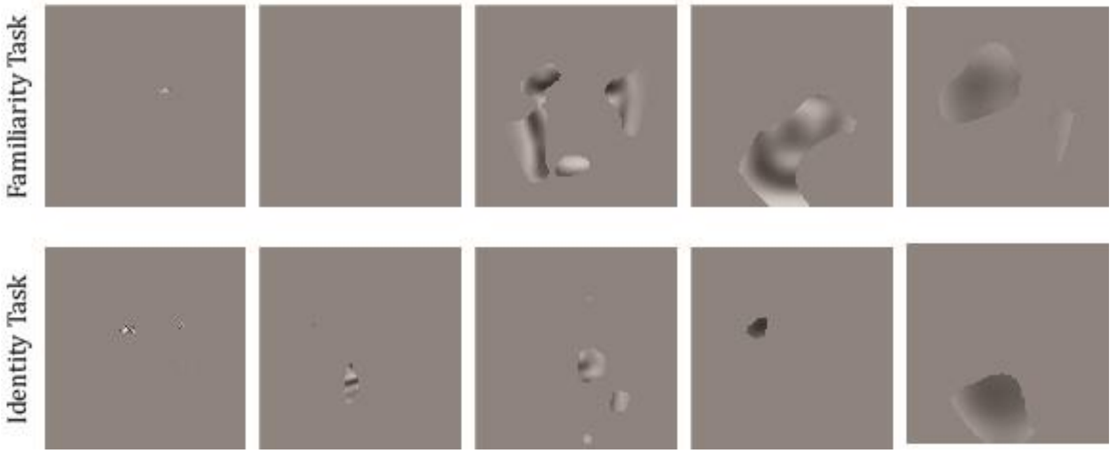
B. SF use



C. Information use - scale split



D. Significant differences in information use



	Familiarity Task					Identification Task					
	Familiar Faces	Unfamiliar Faces	Statistical Comparison	Overall		Familiar Faces	Unfamiliar Faces	Statistical Comparison	Overall		Overall Statistical Comparison
Performance Correct (%)	69.8	77.8	t(15)=1.05, p=0.31	73.8		72.2	76.2	t(15)=0.79, p=0.44	74.4		
Number of Bubbles	n/a	n/a	n/a	118		n/a	n/a	n/a	88		t(15)=3.08, p=0.008*
Reaction Times (ms)	714	792	t(15)=3.58, p=0.003*	753		1164	1001	t(15)=2.96, p=0.01*	1082		

Table 1. Performance metrics (performance accuracy, reaction times, number of bubbles required) for the familiarity and identification tasks.

Supplementary Figure 1