Sliwinska, MW, Searle, LR, Earl, M, O'Gorman, D, Pollicina, G, Burton, AM and Pitcher, D

Face learning via brief real-world social interactions includes changes in face-selective brain areas and hippocampus

Face learning via brief real-world social interactions induces changes in face-selective brain areas and hippocampus

Magdalena W. Sliwinska[1,2], Lydia R. Searle[1], Megan Earl[1], Daniel O'Gorman[1], Giusi Pollicina[1], A. Mike Burton[1,3] & David Pitcher[1]

1. Department of Psychology, University of York, U.K.
2. School of Psychology, Liverpool John Moores University, UK
3. Faculty of Society & Design, Bond University, Australia

Corresponding author: David Pitcher: david.pitcher@york.ac.uk
Department of Psychology, University of York, Heslington, York, YO105DD, U.K.

Abstract

Making new acquaintances requires learning to recognise previously unfamiliar faces. In the current study, we investigated this process by staging real-world social interactions between actors and the participants. Participants completed a face-matching behavioural task in which they matched photographs of the actors (whom they had yet to meet), or faces similar to the actors (henceforth called foils). Participants were then scanned using functional magnetic resonance imaging (fMRI) while viewing photographs of actors and foils. Immediately after exiting the scanner, participants met the actors for the first time and interacted with them for ten minutes. On subsequent days, participants completed a second behavioural experiment and then a second fMRI scan. Prior to each session, actors again interacted with the participants for ten minutes. Behavioural results showed that social interactions improved performance accuracy when matching actor photographs, but not foil photographs. The fMRI analysis revealed a difference in the neural response to actor photographs and foil photographs across all regions of interest only after social interactions had occurred. Our results demonstrate that short social interactions were sufficient to learn and discriminate previously unfamiliar individuals. Moreover, these learning effects were present in brain areas involved in face processing and memory.

Introduction

An extensive behavioural literature has demonstrated large differences between the recognition of familiar and unfamiliar faces. Studies demonstrate that participants perform significantly better in familiar, compared to unfamiliar face recognition tasks (Ellis et al., 1979; Yarmey & Barker, 1971). Familiar face recognition is also less affected by changes in pose, lighting conditions or picture characteristics than unfamiliar face recognition (Hancock et al., 2000; Hill & Bruce, 1996). Other factors also differentiate familiar and unfamiliar face recognition, including image degradation (Burton et al., 1999), removal of external features (Ellis et al., 1979), viewpoint (Bruce, 1982) and context (Dalton, 1993). Matching unfamiliar faces also correlates more strongly with object matching than it does with familiar face matching (Megreya & Burton, 2006). Taken together, these findings demonstrate that familiar and unfamiliar faces are processed, to some extent differently, and that the mental representation of familiar faces is more robust to a wide range of environmental factors (Burton et al., 2011).

Face processing is a complex cognitive operation that is processed across multiple brain areas (Haxby et al., 2000). These include the fusiform face area (FFA) (Kanwisher et al., 1997; McCarthy et al., 1997; Parvizi et al., 2012), the occipital face area (OFA) (Gauthier et al., 2000; Pitcher et al., 2009), the posterior superior temporal sulcus (pSTS) (Campbell et al., 1990; Pitcher & Ungerleider, 2021; Puce et al., 1997) and face-selective voxels in the amygdala (Morris et al., 1996; Pitcher et al., 2019). Many prior fMRI studies have investigated how familiar faces are processed in the brain (Gobbini & Haxby, 2007; Natu & O'Toole, 2011). The majority of these studies have focused on the role of the FFA. FFA activity was correlated with behavioural performance when participants correctly identified target faces (Grill-Spector et al., 2004) and the FFA is sensitive to identity changes of famous individuals (Rotshtein et al., 2005). Numerous studies have also reported greater activity in the FFA for familiar than unfamiliar faces (Elfgren et al., 2006; Gobbini et al., 2004; Pierce et al., 2004; Ramon et al., 2015; Sergent et al., 1992; Weibert & Andrews, 2015), although it is also important to note that some studies have reported no difference (Gorno-Tempini & Price, 2001; Leveroni et al., 2000). It is unclear why the existing literature is inconsistent regarding the role of the FFA in the recognition of familiar faces. One possible explanation is that the nature of the personal familiarity used across different fMRI studies can vary (e.g., personally familiar, visually familiar or famous faces). Natu and O'Toole (2011) have suggested that this variation may account for the differences in the results reported across fMRI studies of familiar and unfamiliar face recognition.

Neuroimaging studies have identified brain areas other than the FFA that also exhibit different neural responses to familiar and unfamiliar faces. For example, in one study participants viewed photographs of unfamiliar, famous, and emotional faces (Ishai et al., 2005). Results showed famous faces and emotional faces elicited a greater neural response across a network of face-selective areas including the FFA, OFA, pSTS, amygdala and hippocampus. A greater response to recently learned faces compared to unlearned faces in the hippocampus has also been reported in a study in which Caucasian participants learned unfamiliar South Korean faces (Ishai & Yago, 2006). Subsequent research by the same group further demonstrated that familiar and unfamiliar faces are associated with unique functional connectivity patterns between face areas in the occipitotemporal cortex and the orbitofrontal cortex (Fairhall & Ishai, 2007).

Prior electroencephalography (EEG) studies of familiar and unfamiliar face processing have also demonstrated some neural sensitivity to familiarity. For example, the N250r event-related potential (ERP) component, recorded over the inferior temporal regions, is modulated by repetitions of face stimuli which is greater for familiar (famous) than unfamiliar faces (Schweinberger et al., 2004; Schweinberger et al., 2002). A more recent EEG study also demonstrated a robust dissociation in the neural response to personally familiar, versus unfamiliar faces that ranged from 200 to 600 milliseconds after stimulus onset (Wiese et al., 2019). Studies have also shown that faces can become familiar over the course of an experiment resulting in differences between familiar and unfamiliar faces in ERP components (especially the N250). These studies have used computer based face learning (Tanaka et al., 2006; Tanaka & Pierce, 2009), videos (Kaufmann et al., 2009) and photorealistic caricatures (Limbach et al., 2018), demonstrating that previously unfamiliar faces can acquire some level or familiarity from a computer screen in a lab environment.

One of the problems with lab-based investigations of face-learning is that they typically lack the context surrounding the way we meet new people day to day, effectively 'controlling away' critical factors in learning (Burton, 2013a; Jenkins & Burton, 2011). Of course, it is extremely difficult to study naturalistic face learning without sacrificing the levels of control available in a lab-based studies. In this study, we investigated whether brief social interactions would be sufficient to bring about learning that could be measured by changes in both behaviour and neural response. To do so, we used a setting that allowed participants to meet previously unknown people, multiple times, interspersed with multiple imaging events. In doing so, we aimed to complement existing lab-based studies by providing evidence from more naturalistic learning. While this inevitably relinquishes some lab control, we hold that converging evidence from artificial and natural learning studies gives the best chance of providing an understanding of this complex and poorly understood phenomenon.

In addition to face-selective areas, another brain area implicated in the recognition of familiar faces is the hippocampus. While it is not a visual area of the brain, the hippocampus is heavily implicated in memory, suggesting that it supports parallel judgments for familiarity and recollection (Squire et al., 2007). Prior fMRI studies have also demonstrated that the neural response to familiar and unfamiliar faces can be dissociated in the hippocampus (Elfgren et al., 2006; Ishai et al., 2002; O'Neil et al., 2013; Platek & Kemp, 2009; Ramon et al., 2015). In addition, neuropsychological studies have shown that cells in the hippocampus increase firing rates in response to familiar faces (Fried et al., 1997) and that removal of the amygdala and the hippocampus can impair face learning (Crane & Milner, 2002). Based on this prior evidence, we included the hippocampus in our study. Extensive prior evidence has demonstrated that face processing is right lateralised (Kanwisher et al., 1997; Landis et al., 1986; Pitcher et al., 2007; Yovel et al., 2003) so we focused our analysis on regions of interest in the right hemisphere. Face-selective areas (FFA, OFA, pSTS and amygdala) were identified in each participant using an independent localiser and the right hippocampus was identified using a subcortical structural atlas (Desikan et al., 2006).

In order to support real-world learning in our design, we used ourselves (the experimenters) as the actors for the social interactions. To achieve this, four of the experimenters became the experimental stimuli. The experimenters (henceforth called actors) each identified an individual whose face resembled their own, to act as their control. We then collected forty photographs of each actor and each foil from their own personal collection. This was done to give a wide sampling of the lighting, picture quality, pose, expression, angle and context that can vary when learning a new face in the real world. Different sets of photographs were used in the fMRI and behavioural sessions. Participants completed an initial behavioural session during which it was confirmed that they did not know, and had never seen any of the actors or foils. Then, participants returned on a subsequent day and were scanned with fMRI while viewing photographs of actors and foils. Immediately after this scan, they interacted with the actors for ten minutes. Then on subsequent days, participants repeated a second behavioural testing session, and then a second fMRI scan. Prior to each of these sessions they again interacted with the actors for ten minutes only.

Materials and Methods

Participants

Twenty-five volunteers participated in this study. Data from two participants were excluded from all analyses after we realized it was possible that they may have seen or met one of the actors before the study commenced. In addition, one participant was excluded from the main analysis because we were unable to identify the right occipital face area (OFA) from their face localiser results. The remaining 22 participants (14 women and 8 men; age range: 18 – 35 years, mean age: 21 and SD = 4) were right-handed, neurologically healthy with normal or corrected-to-normal vision. Each participant provided informed consent and was paid for their time. The study was approved by the York Neuroimaging Centre (YNiC) Research Ethics Committee at the University of York.

Stimulus Sets

Behavioural and Main fMRI Tasks

A set of 320 photographs was used. The photographs presented faces of eight identities (four actors and four foils). Four of the authors (L. S., M. E., D. O'G. and G. P.) were selected as the actors for this experiment. Each actor was asked to identify their own foil, an individual whose face resembled the actor's face. Actors were invited to consult their friends to find an individual they resembled. Similarity to the actor was confirmed by inspection of all authors. None of the foils were related to the actors.

There were 40 photographs per identity. Actors selected their own photographs from their pre-existing personal collection. Photographs of three foils were taken from the social media (e.g., Facebook or Instagram) after obtaining permission from the foils for using their photographs in the experiment and in published figures (see Figure 1 for example photographs). One foil was an Australian celebrity who is not well known in the UK. Her photographs were taken from the Google Images. Photographs were 'ambient images' (Burton et al., 2011; Sutherland et al., 2013), deliberately incorporating the range typically encountered as we recognise faces. Photographs showed faces expressing mostly positive or neutral emotions. The individuals also wore different accessories (e.g., make-up, jewellery, headwear, glasses) and hairstyles across the photographs. Faces were shown from different angles as long as more than 75% of the face was visible to make the face recognition possible. Faces were also at different distances from the camera and were presented in grayscale.

To create behavioural matching tests, a subset of 80 photographs, 10 photographs per identity, were used. Each photograph was presented twice, once it was present in the “same” trial and once in the “different” trial. The same set of 80 photographs was used across the two blocks but they were always paired with a different stimulus. Performance on pairwise face-matching tests is used as a measure of degree of familiarity (Clutterbuck & Johnston, 2004; Kramer et al., 2018). The remaining 240 photographs were divided into six different stimuli sets, containing 5 photographs per identity, to create a total of six fMRI runs (three runs per scanning session). Each of the 240 photographs was presented only once across the scan runs. This was done to ensure that participants learnt the faces of the actors rather than the photographs themselves. The order of the six fMRI runs was randomised across subjects.

Figure 1. Examples of the photographs used in the study. The four photographs on the left show one of the actors. The four photographs on the right show the foil for that actor. Note that distinguishing these two people is a difficult task for unfamiliar viewers (Hancock et al., 2000; Young & Burton, 2017).

fMRI Functional Localisation

The stimuli consisted of 3-second movie clips of faces, bodies, scenes, objects and scrambled objects. Movies of bodies, scenes, and scrambled objects were not relevant to this study hence their data are not presented. These stimuli have been successfully used for functional localisation of face-responsive areas in prior studies (Handwerker et al., 2020; Sliwinska, Bearpark, et al., 2020; Sliwinska, Elson, et al., 2020; Sliwinska & Pitcher, 2018). There were 60 movie clips for each category in which distinct exemplars appeared multiple times. Movies of faces and bodies were filmed on a black background and framed close-up to reveal only the faces or bodies of 7 children as they danced or played with toys or adults (who were out of frame). Movies of scenes included fifteen different locations which were mostly pastoral scenes filmed from a car window while driving slowly through leafy suburbs, along with some other films taken while flying through canyons or walking through tunnels that were included for variety. Movies of objects used 15 different moving objects that were selected in a way that minimizes any suggestion of animacy of the object itself or of a hidden actor pushing the object. Those included mobiles, windup toys, toy planes and tractors, and balls rolling down sloped inclines. Movies of scrambled objects were constructed by dividing each object movie clip into a 15 × 15 box grid and spatially rearranging the location of each of the resulting movie frames. Within each block, stimuli were randomly selected from within the entire set for that stimulus category. This meant that the same movie clip could appear within the same block but given the number of stimuli this did not occur frequently.

Procedure

Each participant completed four testing sessions that took place on separate days (Figure 2). Sessions one and three were behavioral sessions, sessions two and four were fMRI scanning sessions. Each behavioral session lasted approximately 15 minutes and each fMRI session lasted for approximately 1 hour. After the first behavioral session, an experimenter not performing as an actor (M.S.) ensured that the participant did not know and had never seen the actors or foils. If participants reported they recognized any of the actor or foil photographs, they were excluded from the study.

Social interactions lasted 10 minutes each and occurred immediately after session two (the first scan) and immediately prior to sessions three (the second behavioral test) and four (the second scan). The gap between sessions varied across participants due to logistical reasons but all participants completed all sessions within two weeks.

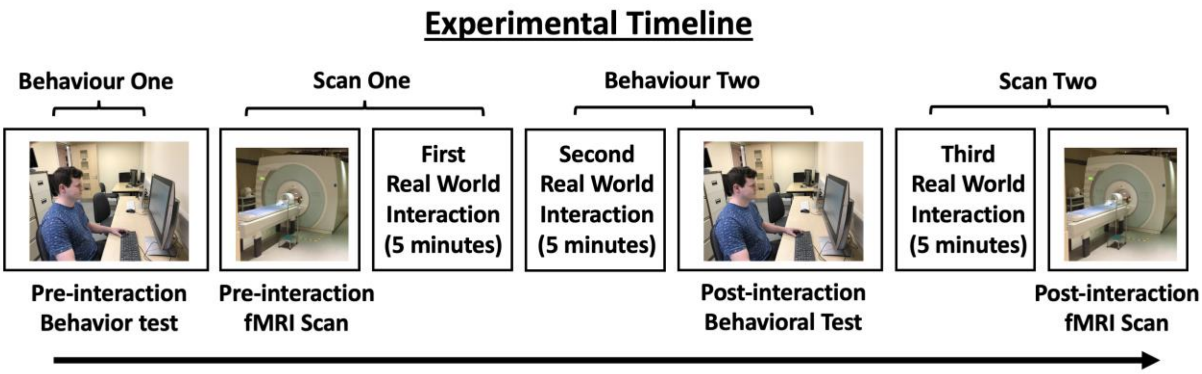

Figure 2. The experimental timeline. Testing took place on four separate days. During session one, participants completed the first behavioral task. During session two, participants were scanned using fMRI. When the scan was complete participants interacted with actors for 10 minutes immediately after leaving the scanner. During session three, participants interacted with the actors for 10 minutes and then completed the second behavioral task. During session four, participants interacted with the actors for 10 minutes and then completed the second fMRI scan.

Real-world Social Interactions

All interactions included a planned set of activities involving each participant with all four actors. These were designed to resemble real-world interactions and forced participants to pay attention to the faces of the actors.

At the first interaction, participants met the actors for the first time. The participant was greeted by the actors immediately after leaving the MRI scanner and taken by them to a separate room. For the first few minutes, the actors involved the participant in a general conversation. Then, each actor played the Rock-Paper-Scissors game ten times with the participant maintaining the eye contact with the participant.

The second interaction occurred prior to the second behavioral testing session. Actors met participants at the testing room and engaged them in a semi-structured conversation. Each actor was assigned five questions that were designed to elicit a short conversation (e.g., *What qualities do you tend to look for in a friend?; What movies have you seen recently*?; *What would your dream job be?*). After the conversation, each actor played Rock-Paper-Scissors ten times with the participant. Actors then reminded participants about the instructions for the behavioral task and left the room. All testing occurred in the absence of the actors. After the completion of the task actors re-entered the testing room and finished the interaction with a very brief conversation.

The third interaction occurred prior to the second fMRI scan. Actors engaged the participant in a short general conversation, explained the procedures of the second fMRI session and went through the MRI safety questionnaire with the participant. Each actor always asked the same set of questions. Participants were then placed in the MRI scanner by operators (not the actors).

Behavioral Sessions

Participants performed a computer-based face-matching task in which they judged whether two photographs, appearing simultaneously side-by-side on the screen, presented the same identity or two different identities (Figure 3). Participants used their right or middle finger to respond “yes” or “no”, respectively, by pressing appropriate keys on the keyboard. The response time was unlimited, but participants were asked to provide their answer as quickly and accurately as possible. No performance feedback was provided.

The face-matching task was written using E-Prime v2.0 (Psychology Software Tools, Inc.) Each trial began with a fixation cross displayed for 500 msec and followed by the stimuli presentation that was shown until the participant responded. The response triggered the next trial. Each stimuli pair was presented on a white background. Stimuli were displayed on the 22-inch CRT monitor set at 1024 × 768 resolution and refresh rate of 85 Hz. The first behavioural session was led by an experimenter who was not one of the actors, while the second behavioural session was led by all the actors (see Real-world Social Interactions for more information).

During each session, participants completed one block of the face-matching task. Each block consisted of 80 trials, with three different face-matching conditions. These were actor-actor trials (20 trials), foil-foil trials (20 trials) and actor-foil trials (40 trials). In the actor-actor and foil-foil trials, there were five trials per identity. The actor-foil trials presented a picture of an actor paired with a picture of her/his foil (10 trials per an actor). The first two conditions (actor-actor, foil-foil) constituted “same” trials while the third condition (actor-foil) constituted “different” trials. Half of the trials in the third condition had a photograph of the actor on the left side of the screen and a photograph of the foil on the right side of the screen while the reverse order of the presentation sides was applied in second half of the trials. This was done to avoid learning the identities based on a particular side of the screen. The same stimuli pairings were presented to all the participants but the trial order was randomized across participants.

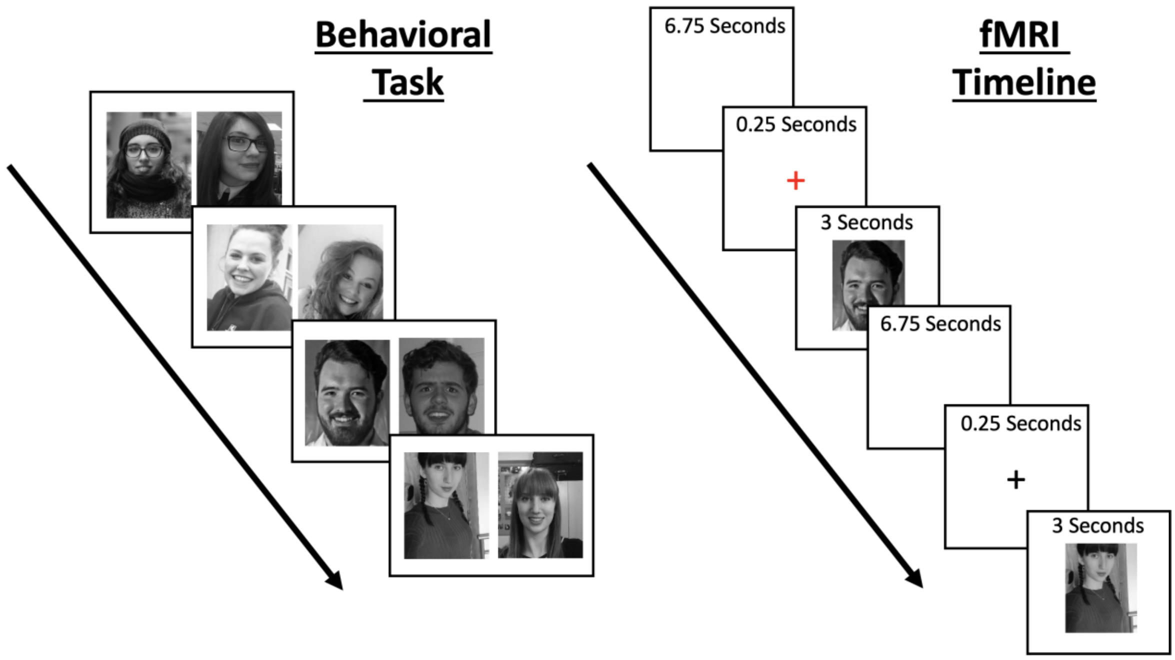

Figure 3. The timeline of the behavioural task procedure (left). Participants had to judge whether the two simultaneously presented photographs depicted the same identity or not. The presented trials represent ‘no’, ‘yes’, ‘no’, ‘yes’ trials. The timeline of the fMRI task procedure (right). Participants were scanned while photographs of actors or foils were presented for 3 seconds each. Participants were not given any instructions about the photographs. To ensure they pay attention to the stimuli, they were asked to press a response button every time the black fixation cross turned red (this occurred 30% of the time).

fMRI Sessions

In fMRI sessions, participants completed three functional runs of the main experimental tasks, followed by two runs of a face localizer task designed to identify face-selective regions of interest (ROIs). During the first fMRI session a T1-weighted structural brain scan was also collected to anatomically localise the functional data for each participant. In addition, during both sessions, a T1-FLAIR (fluid-attenuated inversion recovery) scan was acquired to improve co-registration of the functional and structural scans across the two scanning sessions.

The main fMRI task was written using PsychoPy v12.0, an open source software package. Each trial lasted 10 sec and it commenced with a gray blank screen displayed for 6.75 sec, followed by the black or red fixation cross displayed for 0.25 sec, and then by a stimulus presentation for 3 sec. Each stimulus and fixation cross were presented against a gray background. Stimuli were presented using a ProPixx LED projector (VPixx Technologies, Quebec, Canada) set at 1920 × 1080 resolution a refresh rate of 120 Hz. At the beginning and end of each run, there was a 10 sec rest period. Each run lasted 7 min. Stimuli were presented in a slow event-related design with relatively long inter-stimulus intervals to avoid the overlaps of hemodynamic response function (HRF) across trials.

In the main task, participants passively viewed photographs of faces of actors and foils that were presented sequentially at the center of the screen (5 photographs of each identity per run, 15 in total per scan). Each run contained 40 different photographs (5 photographs per identity). Each scan session contained three runs of the task. Run order was randomized across participants and across the fMRI sessions. Participants were instructed to pay attention to the screen but were not given any specific instructions regarding the faces. To ensure participants attended the screen, they were asked to press a response button when the fixation cross preceding each photograph was red instead of black. The red fixation cross occurred randomly in 3 out of 10 trials.

Face-selective ROIs were identified using a dynamic face localizer task (Pitcher et al., 2011). Data were acquired using block-design runs, lasting 234 sec each. During those runs, participants were instructed to watch videos of faces, bodies, scenes, objects, or scrambled objects, without performing any overt task. Each run contained two sets of five consecutive stimulus blocks to form two blocks per stimulus category per run. Each block lasted 18 sec and contained stimuli from one of the five stimulus categories. Each functional run also contained three 18 sec rest blocks, which occurred at the beginning, middle, and end of the run. During the rest blocks, a series of six uniform color fields were presented for 3 sec each. The order of stimulus category blocks in each run was palindromic (e.g., rest, faces, objects, scenes, bodies, scrambled objects, rest, scrambled objects, bodies, scenes, objects, faces, rest) and randomized across runs.

Brain Imaging Acquisition and Analysis

Imaging data were acquired using a 3T Siemens Magnetom Prisma MRI scanner (Siemens Healthcare, Erlangen, Germany) at YNiC utilising a twenty-channel phased array head coil tuned to 123.3 MHz. Functional images for the main and localisation tasks were recorded using a gradient-echo EPI sequence (35 interleaved slices, repetition time (TR) = 2000 msec, echo time (TE) = 30 msec, flip angle = 80°; voxel size = 3 × 3 × 3 mm; matrix size = 64 × 64; field of view (FOV) = 192 × 192 mm) providing whole brain coverage. Slices were aligned with the anterior to posterior commissure line. Structural images were acquired using a high-resolution T-1 weighted 3D fast spoilt gradient (SPGR) sequence (176 interleaved slices, repetition time (TR) = 2300 msec, echo time (TE) = 2.26 msec, flip angle = 8°; voxel size = 1 × 1 × 1 mm; matrix size = 256 × 256; field of view (FOV) = 256 × 256). T1-FLAIR scans (35 interleaved slices, repetition time (TR) = 3000 msec, echo time (TE) = 8.6 msec, flip angle = 150°; voxel size = 0.8 × 0.8 × 3.0 mm; matrix size = 256 × 256; field of view (FOV) = 192 × 192) were acquired with the same orientation as the functional scans.

Functional MRI data were analyzed using using fMRI Expert Analysis Tool (FEAT) included in the FMRIB (v6.0) Software Library (www.fmrib.ox.ac.uk/fsl). Data from the first four TRs from each run were discarded. The remaining images were slice-time corrected and realigned to the first volume of each functional run and to the corresponding anatomical scan. The volume-registered data were spatially smoothed with a 5-mm full-width-half-maximum Gaussian kernel. Signal intensity was normalized to the mean signal value within each run and multiplied by 100 so that the data represented percent signal change from the mean signal value before analysis.

The data from the main experimental task was entered into a general linear model (GLM) by convolving the standard hemodynamic response function with the regressors of interest. These were either actors (data from all four actors was averaged together) or foils (data from all four foils was averaged together). This created four conditions: actor pre-interaction, actor post-interaction, foil pre-interaction, foil post-interaction). The model was convolved using a double-gamma hemodynamic response function (HRF) to generate the main regressors. Temporal derivatives for each condition were included. First-level functional results for each participant were registered to their anatomical scan and then to the Montreal Neurological Institute (MNI) 152-mean brain using a 12 degree-of-freedom affine registration. The T1-FLAIR scan was added as an expanded functional image to help aid registration.

ROIs were identified for each participant using a contrast of greater activation evoked by dynamic faces than that evoked by dynamic objects, calculating significance maps of the brain using an uncorrected statistical threshold ($p = 0.01$). We identified four face-selective areas using anatomical landmarks, these were; the fusiform face area (FFA), occipital face area (OFA), posterior superior temporal sulcus (pSTS) and face-selective voxels in the amygdala. We were able to identify all four ROIs in the right hemisphere of twenty-two participants. In the left hemisphere the ROIs were not identified across all participants (left FFA was present in 17 participants, left OFA in 19 participants, pSTS in 15 participants and the left amygdala in 11 participants). For this reason, we focused our fMRI data analysis on the right hemisphere ROIs only as we have in prior fMRI studies of the face network using these stimuli (Pitcher et al., 2019; Pitcher et al., 2017). The peak voxel of activation was identified for each area and a 5 mm sphere was individually drawn around this point for each participant. The mean of the peak MNI coordinates for the ROIs across participants were; right FFA - 40, -52, -20, right OFA – 41, -79, -15, right pSTS – 52, -38, 3, right amygdala – 19, -5, -15. A table of the peak MNI coordinates for all ROIs across all participants in supplemental materials.

The right hippocampus was identified using the FSL mask from Harvard-Oxford Cortical and Subcortical Structural Atlas (Desikan et al., 2006). A point in the middle of the mask was located and a 5 mm sphere was individually drawn around this point for each participant.

Within each functionally defined ROI, we then calculated the magnitude of response (percent signal change from a fixation baseline) for the four experimental conditions (actor pre-interaction, actor post-interaction, foil pre-interaction, foil post-interaction).

We also conducted a whole-brain group analyses using a 2-way mixed model ANOVA (session by actor/foil) for all twenty-two participants. This analysis did not yield any activations that passed an appropriate level for a statistically significant threshold. However, at the request of the reviewers we have include the subthreshold clusters we observed in the supplemental materials for information only.

Data and Code Availability Statement

fMRI data, behavioural data, stimuli, and experimental code is available at https://osf.io/pnvs2/.

Results

Behavioral Task

Behavioral results clearly demonstrated that the real-world social interactions were sufficient to learn the faces of the previously unfamiliar actors (Figure 4). In post-interaction behavioral test, accuracy for the actor-actor trials matching condition increased, while accuracy for the foil-foil trials condition was unchanged. Performance accuracy for the actor-foil trials was also not improved by social interactions. This relatively brief encounter with a live actor appears to improve subsequent recognition (as measured in a matching test) but this improvement is evident only in the actor-actor *matching* trials. This is consistent with studies in which familiarity effects are carried by improvements in viewers' ability to cohere different images of the same face, or 'telling faces together' (Jenkins & Burton, 2011; Ritchie & Burton, 2017) and recent theoretical developments in face learning offer potential mechanisms for this process (Kramer et al., 2018) but see (Blauch et al., 2021) for an alternate view.

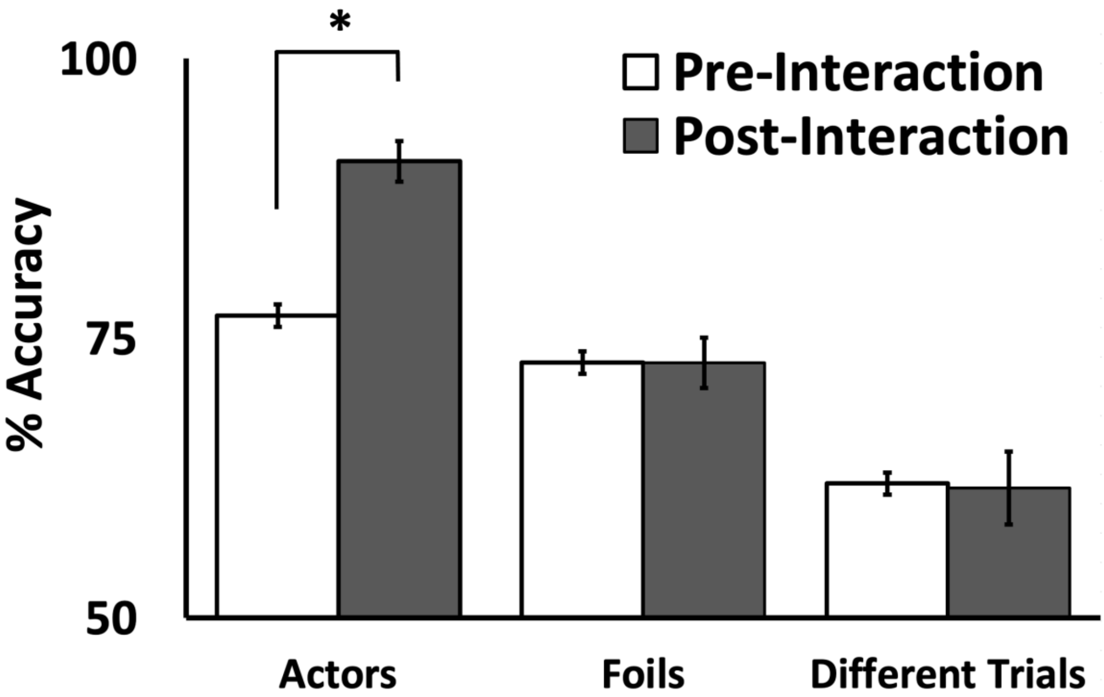

Figure 4. Mean accuracy data for the behavioural face-matching task before and after social interactions had occurred. Results showed that social interactions improved performance accuracy for the actor/actor trials only. There was a significant two-way interaction between session and trial type (p < 0.0001). Error bars show standard errors of the mean across participants. * denotes a significant difference (*p* < 0.001).

Accuracy data were entered in a two (session: pre-interaction, post-interaction) by three (trial type: actor-actor, foil-foil, actor-foil) repeated measures analysis of variance (ANOVA). Results showed a significant main effect of trial type ($F (2, 42) = 23$; $p < 0.0001$; partial η^2 =

0.5) but not of session ($F(1, 21) = 3.8$; $p = 0.065$; partial $\eta^2 = 0.16$). Crucially, there was a significant two-way interaction ($F(2, 42) = 9.25$; $p < 0.0001$; partial $\eta^2 = 0.30$).

Consistent with our hypothesis, we found a significant two-way interaction between session and trial type demonstrating that social interactions improved performance accuracy when matching actor-actor photographs. To understand what factors were driving this effect, we calculated the simple main effects. Results showed a significant difference before and after the interactions for the actor-actor trials ($F(1,21) = 19$, $p < 0.001$; partial $\eta^2 = 0.24$) but not the foil-foil trials ($F(1,21) = 0.1$, $p = 0.9$; partial $\eta^2 = 0.00$) or the actor-foil trials ($F(1,21) = 0.1$, $p = 0.7$; partial $\eta^2 = 0.00$).

Tukey's HSD (honestly significant difference) tests showed that there was no significant difference between the actor-actor and foil-foil trials pre-interaction ($p = 0.28$) but there was a significant difference post-interaction ($p < 0.0001$). There were also significant differences between actor-actor and actor-foil trials pre-interaction ($p = 0.002$) and post-interaction ($p < 0.0001$). The same pattern was also apparent for foil-foil trials and actor-foil trials both pre-interaction ($p = 0.04$) and post-interaction ($p = 0.014$).

We also analysed the reaction time (RT) data (Figure 5) in a two (session: pre-interaction, post-interaction) by three (trial type: actor-actor, foil-foil, actor-foil) repeated measures analysis of variance (ANOVA). While they showed the same pattern as the accuracy data the interaction did not reach significance. The main effects of session ($F(1, 21) = 14.48$; $p = 0.001$; partial $\eta^2 = 0.40$) and trial type ($F(2, 42) = 7.86$; $p = 0.001$; partial $\eta^2 = 0.26$) were significant. The two-way interaction between these two conditions was not significant ($F(2, 42) = 3.05$; $p = 0.06$; partial $\eta^2 = 0.12$).

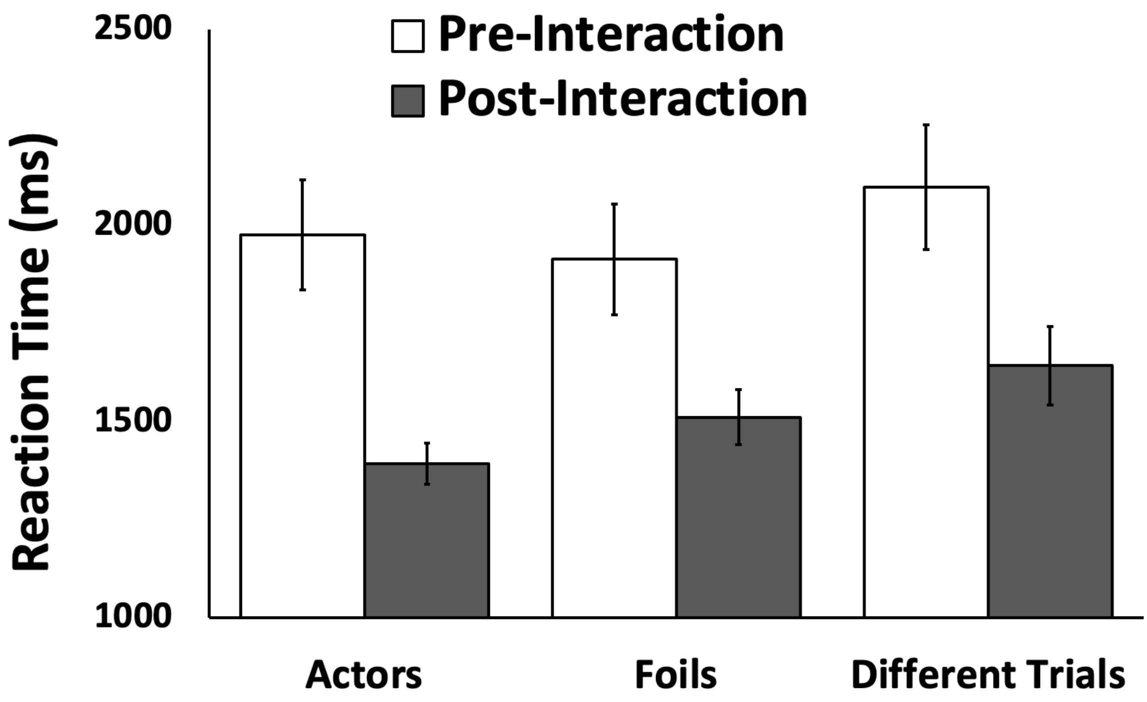

Figure 5. Mean accuracy data for the behavioural face-matching task before and after social interactions had occurred.

fMRI Results

The neural response in the right FFA, right OFA, right pSTS, right amygdala and right hippocampus to actor photographs and foil photographs only differed in scan two, after the real-world social interactions had taken place (Figure 6). The pattern across all ROIs was similar, namely the response to foil photographs was lower in scan two (after social interactions had occurred) than in scan one. In contrast, the response to the actor photographs was unchanged between scan one and scan two. These results demonstrate that the learning effect we observed for the actor photographs in the behavioral task (Figure 4) was matched by a sustained neural response. The lack of a learning effect for the foil faces was matched with a weaker neural response. This result is consistent with fMRI studies of learning in humans and non-human primates reporting a drop in the neural response to the unlearned stimulus (Kaskan et al., 2017; Op de Beeck et al., 2006).

Face-selective ROIs were identified with data from the independent functional localizer using contrast of face movies greater than object movies. We were able to localize the four face-selective ROIs in 22 of the 23 participants (one participant did not have the right OFA). The hippocampus was identified using the Harvard-Oxford Cortical and Subcortical Structural Atlas. Percent signal change data were entered into a five (ROI: rOFA, rFFA, rpSTS, right amygdala, and right hippocampus) by two (photograph: actor, foil) by two (session: pre-interaction, post-interaction) repeated measures ANOVA. We found main effects of ROI ($F (1,21) = 58$, $p < 0.0001$; partial $\eta^2 = 0.73$) but not of photograph ($F (1,21) = 1.5$, $p = 0.23$; partial $\eta^2 = 0.07$) or session ($F (1,21) = 1.6$, $p = 0.22$; partial $\eta^2 = 0.07$). There was no significant two-way interaction between ROI and photograph ($F (4,84) = 0.5$, $p = 0.72$; partial $\eta^2 = 0.02$), or between ROI and session ($F (4,84) = 0.4$, $p = 0.8$; partial $\eta^2 = 0.02$), but crucially there was a significant two-way interaction between photograph and session ($F (1,21) = 7.6$, $p = 0.012$; partial $\eta^2 = 0.26$). There was no significant three-way interaction ($F (4,84) = 0.7$, $p = 0.68$; partial $\eta^2 = 0.03$).

Consistent with our hypothesis, we found a significant two-way interaction between photograph and session demonstrating that social interactions differentially affected the neural response to actor and foil photographs. To understand what factors were driving this effect, we calculated the simple main effects between the four conditions across all ROIs: actor pre-interaction (0.56, *S.E.* = 0.05), foil pre-interaction (0.55, *S.E.* = 0.04), actor post-interaction (0.57, *S.E.* = 0.05) and foil post-interaction (0.51, *S.E.* = 0.04). There was no significant difference between actor and foil photographs pre-interaction (F (1,42) = 0.8, $p > 0.5$; partial $\eta^2 = 0.02$) but there was a significant difference post-interaction (F (1,42) = 7.5, $p = 0.009$; partial $\eta^2 = 0.15$). In addition, there was no significant difference between pre-interaction actor photographs and post-interaction actor photographs (F (1,42) = 0.05, $p > 0.5$; partial $\eta^2 = 0.03$) but there was a significant difference between pre-interaction foil photographs and post-interaction photographs (F (1,42) = 4.7, $p = 0.04$; partial $\eta^2 = 0.03$).

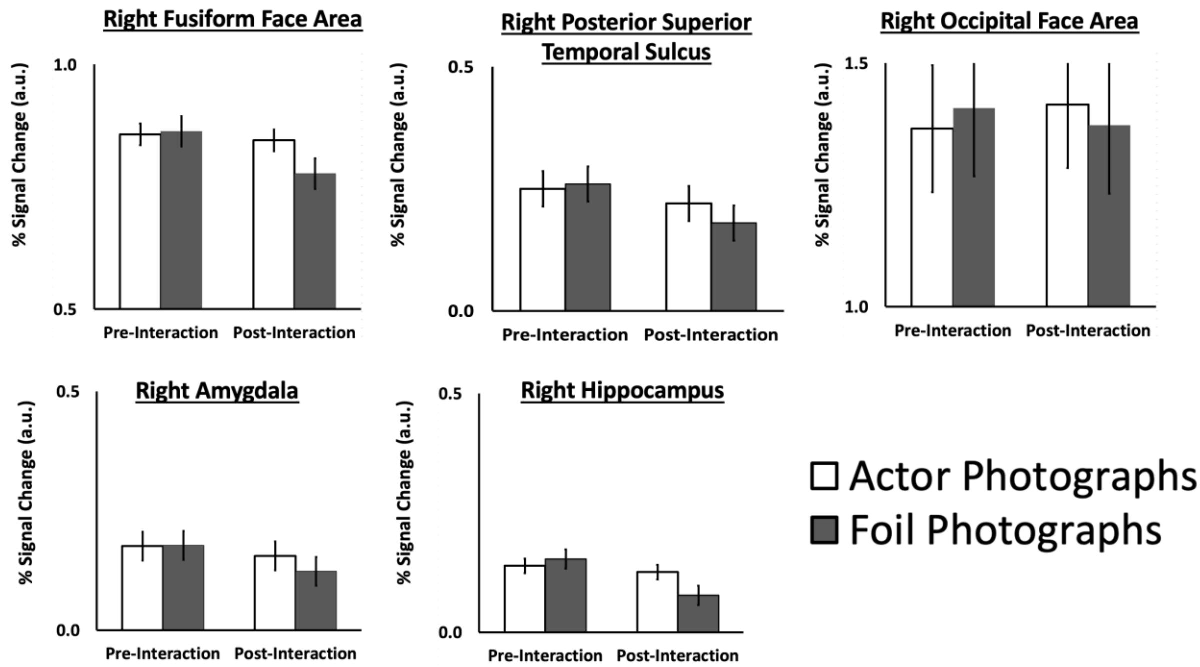

Figure 6. Percent signal change data for the actor and foil photographs before and after social interaction in the face-selective areas and the right hippocampus. We observed a significant two-way interaction ($p = 0.012$; partial $\eta^2 = 0.26$) between photograph type (actor and foil) and session (pre-interaction and post-interaction). This was driven by a reduction in the neural signal to foil photographs across all ROIs in scan two after the real-world social interactions had occurred.

Discussion

In the current study, we investigated the behavioural and neural basis of face learning using real-world social interactions. Participants were scanned on two separate days while viewing photographs of four actors or each actor's foil (the control stimuli). Between these scanning sessions the participants met and interacted with the actors for ten minutes on three separate days. In addition to the scanning sessions, participants completed a behavioural face-matching task before and after two of the real-world interactions. Behavioural results showed participants were significantly better at matching actor photographs than foil photographs but only after the social interactions had occurred (Figure 4). ROI analysis of the neuroimaging data focused on the fusiform face area (FFA), occipital face area (OFA), posterior superior temporal sulcus (pSTS), amygdala and hippocampus in the right hemisphere. There was no difference between the neural response to actor and foil photographs in all ROIs in the first scan, prior to the social interactions. In contrast, the response to foil photographs was significantly lower than the response to actor photographs during the second scan (Figure 5). These results demonstrate that short real-world social interactions are sufficient to learn the faces of previously unfamiliar individuals, and that this learning process can be detected in the nodes of the face processing network and in the hippocampus.

Prior behavioural studies have highlighted numerous differences between familiar and unfamiliar faces. It has been known for many years that memory for unfamiliar faces is severely affected by small changes in the image, for example due to lighting, expression or pose, whereas such changes barely affect familiar face recognition (Bruce, 1986; Ellis et al., 1979; Klatzky & Forrest, 1984). Simultaneous face matching is also difficult for unfamiliar faces, but trivially easy for familiar faces (Bruce et al., 1999; Burton et al., 1999). This has led some authors to argue that unfamiliar face processing (for identity) is primarily image-based (Hancock et al., 2000), whereas familiar face recognition is robust to variation because it relies on computation of the within-identity variability specific to each individual (Burton, 2013b; Jenkins et al., 2011). In the current study, we sought to incorporate within-individual variability by collecting forty photographs of each actor and their foil for use in the behavioural and fMRI tasks. These photos were taken with different cameras, from different angles, and in different lighting conditions. Despite having a range of different variations in both the actor and foil photographs we only observed a difference between the actor and foil conditions after the social interactions had occurred.

When meeting a previously unfamiliar individual, we are visually exposed to their face while we form opinions and gather semantic information from them (e.g., their name, where they are from, etc). This wealth of information is used and encoded when learning to recognise their face. In the current study, we structured the social interactions (e.g., the actors asked pre-planned questions and played short games with the participants) but there was no attempt to specifically guide the face learning. In fact, we did not even instruct the participants that the aim of the study was to learn the faces of the actors. Instead, our intention was to incorporate natural, encounter-based learning into the design, rather than use the typical lab-based controlled exposure. Perhaps surprisingly, rather little is known about the process of face learning, but it seems likely that many different factors will have contributed to the learning effect we observed after the social interactions had taken place. Prior studies suggest that these factors could include semantic information (Heisz & Shedden, 2009), multisensory information (von Kriegstein & Giraud, 2006; von Kriegstein et al., 2005) and dynamic information (Kaufmann et al., 2009) have on behavioural performance accuracy. We did not control for any of these factors in our experimental design and it seems likely that all will have contributed to our results to some degree.

However, most previous research has used a lab environment rather than in real-world social interactions. One of the challenges for future work will be to preserve naturalistic learning while also retaining the ability to manipulate potential contributions to learning in a systematic way.

The ROI analysis revealed a significant difference between actor and foil faces only after social interactions had occurred (Figure 5). This is consistent with prior neuroimaging studies that reported a higher response to familiar than unfamiliar faces in the FFA (Elfgren et al., 2006; Gobbini et al., 2004; Pierce et al., 2004; Sergent et al., 1992; Weibert & Andrews, 2015) as well as in other face-selective areas (Ishai et al., 2005; Ishai & Yago, 2006). The present study differed from these prior studies in that we scanned participants before and after real world face learning had occurred during a brief face to face conversation. While there was no significant main effect of social interaction, it is worth noting that the response to actor faces in scan two was not significantly greater than the response to actor faces in scan one. This decrease in the response to the unlearned stimulus is consistent with fMRI studies of learning in humans and non-human primates reporting a drop in the neural response to the unlearned stimulus (Kaskan et al., 2017; Op de Beeck et al., 2006). It is also possible that the reduction in the BOLD response to the foil faces in scan two was due to repetition suppression (also called adaptation) (Grill-Spector et al., 2006). The increased neural activation we observed after the real-life interactions with the actors may have reduced or eliminated the priming effect usually observed with repeated exposure to the same stimuli. This would be consistent with evidence demonstrating that that learning may bias repetition suppression effects towards repetition enhancement (Segaert et al., 2013). Finally, another possible account is the general attention of participants which may have been lower overall during the second session. However, it is important to note that neither of these factors can account for the difference in the response to actor and foil faces that we observed after the social interactions had taken place.

The behavioural results showed the same pattern as the fMRI results, namely we observed a difference in matching actors and foils only after the social interactions had occurred (Figure 4). There were only two ten-minute interactions between the behavioural testing sessions demonstrating that twenty minutes were sufficient to learn the faces of the actors. This demonstrates the effectiveness and utility of these real-world interactions in the study of how face learning and person recognition occur. Interestingly, we did not observe a significant effect of social interaction for the different trials. In fact, across the literature on face matching there is inconsistency about whether effects of familiarity or expertise are driven by “hits” (same person matching performance) or “correct rejections” (different person matching performance) (Bobak et al., 2019; Matthews & Mondloch, 2018; Ritchie et al., 2021). While studies of learning typically show improved accuracy with increased exposure (and some only report this), a break-down of these effects shows they are typically driven by only one component of performance and it has (so far) been impossible to isolate a satisfactory explanation for this. In the present study, improvement is driven by performance in same-item trials, and we suggest that this is largely consistent with previous literature using more naturalistic exposure to faces. Further, our results are consistent with studies demonstrating that telling people apart is not the same process as telling them together (Jenkins et al., 2011). These studies employed a sorting paradigm to demonstrate that unfamiliar face recognition is less tolerant of within-person variation than familiar face recognition. Our design was clearly robust enough for participants to learn the actors face but not the foil face, leading to an increased level of uncertainty when matching actor-foil photographs, alongside an improvement in ability to match different instances of the learned faces. Future studies could test this hypothesis in a two-stage real-world design in which participants are exposed to one group of actors and then a second group of actors who physically match the first group.

Conclusion

Our study has demonstrated that brief real-world social interactions are sufficient to learn the faces of previously unfamiliar individuals. Importantly, we were also able to detect the difference between learned and unlearned faces in the nodes of the face processing network and in the hippocampus. Finally, given the complexity of the social environments that we navigate on a daily basis we propose that such real-world interactions can be adopted in future neuroimaging studies.

Acknowledgements

Research supported by BBSRC grant BB/P006981/1 awarded to D.P. Thanks to Nancy Kanwisher for providing the localiser stimuli and to Joe Devlin for statistical advice.

References

Blauch, N. M., Behrmann, M., & Plaut, D. C. (2021). Computational insights into human perceptual expertise for familiar and unfamiliar face recognition. *Cognition*, *208*, 104341. https://doi.org/10.1016/j.cognition.2020.104341

Bobak, A. K., Mileva, V. R., & Hancock, P. J. B. (2019). A grey area: how does image hue affect unfamiliar face matching? *Cogn Res Princ Implic*, *4*(1), 27. https://doi.org/10.1186/s41235-019-0174-3

Bruce, V. (1982). Changing faces: visual and non-visual coding processes in face recognition. *Br J Psychol*, *73*(Pt 1), 105-116. https://www.ncbi.nlm.nih.gov/pubmed/7059746

Bruce, V. (1986). Influences of familiarity on the processing of faces. *Perception*, *15*(4), 387-397. https://doi.org/10.1068/p150387

Bruce, V., Henderson, Z., Greenwood, K., Hancock, P., Burton, A. M., & Miller, P. (1999). Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*, *5(4)*, 339-360.

Burton, A. M. (2013a). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66*(8), 1467-1485. https://doi.org/10.1080/17470218.2013.800125

Burton, A. M. (2013b). Why has research in face recognition progressed so slowly? The importance of variability. *Quarterly Journal of Experimental Psychology*, *66(8)*, 1467-1485.

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *Br J Psychol*, *102*(4), 943-958. https://doi.org/10.1111/j.2044-8295.2011.02039.x

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from Security Surveillance [Article]. *Psychological Science*, *10*(3), 243-248. https://doi.org/10.1111/1467-9280.00144

Campbell, R., Heywood, C. A., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, *28*(11), 1123-1142. https://doi.org/10.1016/0028-3932(90)90050-x

Clutterbuck, R., & Johnston, R. A. (2004). Matching as an index of face familiarity. *Visual Cognition*, *11*(7), 857-869. https://doi.org/10.1080/13506280444000021

Crane, J., & Milner, B. (2002). Do I know you? Face perception and memory in patients with selective amygdalo-hippocampectomy. *Neuropsychologia*, *40*(5), 530-538. https://www.ncbi.nlm.nih.gov/pubmed/11749983

Dalton, P. (1993). The role of stimulus familiarity in context-dependent recognition. *Mem Cognit*, *21*(2), 223-234. https://www.ncbi.nlm.nih.gov/pubmed/8469131

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Elfgren, C., van Westen, D., Passant, U., Larsson, E. M., Mannfolk, P., & Fransson, P. (2006). fMRI activity in the medial temporal lobe during famous face processing. *Neuroimage*, *30*(2), 609-616. https://doi.org/10.1016/j.neuroimage.2005.09.060

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception*, *8*(4), 431-439. https://doi.org/10.1068/p080431

Fairhall, S. L., & Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cereb Cortex*, *17*(10), 2400-2406. https://doi.org/10.1093/cercor/bhl148

Fried, I., MacDonald, K. A., & Wilson, C. L. (1997). Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron*, *18*(5), 753-765. https://www.ncbi.nlm.nih.gov/pubmed/9182800

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level (vol 12, pg 499, 2000). *Journal of Cognitive Neuroscience*, *12*(5), 912-912. <Go to ISI>://WOS:000089817500020

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*(1), 32-41. https://doi.org/10.1016/j.neuropsychologia.2006.04.015

Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, *22*(4), 1628-1635. https://doi.org/10.1016/j.neuroimage.2004.03.049

Gorno-Tempini, M. L., & Price, C. J. (2001). Identification of famous faces and buildings: a functional neuroimaging study of semantically unique items. *Brain*, *124*(Pt 10), 2087-2097. https://doi.org/10.1093/brain/124.10.2087

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects [Review]. *Trends in Cognitive Sciences*, *10*(1), 14-23. https://doi.org/10.1016/j.tics.2005.11.006

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci*, *7*(5), 555-562. https://doi.org/10.1038/nn1224

Hancock, P. J., Bruce, V. V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends Cogn Sci*, *4*(9), 330-337. https://www.ncbi.nlm.nih.gov/pubmed/10962614

Handwerker, D. A., Ianni, G., Gutierrez, B., Roopchansingh, V., Gonzalez-Castillo, J., Chen, G., Bandettini, P. A., Ungerleider, L. G., & Pitcher, D. (2020). Theta-burst TMS to the posterior superior temporal sulcus decreases resting-state fMRI connectivity across the face processing network. *Netw Neurosci*, *4*(3), 746-760. https://doi.org/10.1162/netn_a_00145

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends Cogn Sci*, *4*(6), 223-233. https://www.ncbi.nlm.nih.gov/pubmed/10827445

Heisz, J. J., & Shedden, J. M. (2009). Semantic learning modifies perceptual face processing. *J Cogn Neurosci*, *21*(6), 1127-1134. https://doi.org/10.1162/jocn.2009.21104

Hill, H., & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *J Exp Psychol Hum Percept Perform*, *22*(4), 986-1004. https://www.ncbi.nlm.nih.gov/pubmed/8756964

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage*, *17*(4), 1729-1741. https://www.ncbi.nlm.nih.gov/pubmed/12498747

Ishai, A., Schmidt, C. F., & Boesiger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Res Bull*, *67*(1-2), 87-93. https://doi.org/10.1016/j.brainresbull.2005.05.027

Ishai, A., & Yago, E. (2006). Recognition memory of newly learned faces. *Brain Res Bull*, *71*(1-3), 167-173. https://doi.org/10.1016/j.brainresbull.2006.08.017

Jenkins, R., & Burton, A. M. (2011). Stable face representations. *Philos Trans R Soc Lond B Biol Sci*, *366*(1571), 1671-1683. https://doi.org/10.1098/rstb.2010.0379

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. *Cognition*, *121*(3), 313-323. https://doi.org/10.1016/j.cognition.2011.08.001

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci*, *17*(11), 4302-4311. https://www.ncbi.nlm.nih.gov/pubmed/9151747

Kaskan, P. M., Costa, V. D., Eaton, H. P., Zemskova, J. A., Mitz, A. R., Leopold, D. A., Ungerleider, L. G., & Murray, E. A. (2017). Learned Value Shapes Responses to Objects in Frontal and Ventral Stream Networks in Macaque Monkeys. *Cereb Cortex*, *27*(5), 2739-2757. https://doi.org/10.1093/cercor/bhw113

Kaufmann, J. M., Schweinberger, S. R., & Burton, A. M. (2009). N250 ERP correlates of the acquisition of face representations across different images. *J Cogn Neurosci*, *21*(4), 625-641. https://doi.org/10.1162/jocn.2009.21080

Klatzky, R. L., & Forrest, F. H. (1984). Recognizing familiar and unfamiliar faces. *Mem Cognit*, *12*(1), 60-70. https://www.ncbi.nlm.nih.gov/pubmed/6708811

Kramer, R. S. S., Young, A. W., & Burton, A. M. (2018). Understanding face familiarity. *Cognition*, *172*, 46-58. https://doi.org/10.1016/j.cognition.2017.12.005

Landis, T., Cummings, J. L., Christen, L., Bogen, J. E., & Imhof, H. G. (1986). Are Unilateral Right Posterior Cerebral-Lesions Sufficient to Cause Prosopagnosia - Clinical and Radiological Findings in 6 Additional Patients. *Cortex*, *22*(2), 243-252. https://doi.org/Doi 10.1016/S0010-9452(86)80048-X

Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *J Neurosci*, *20*(2), 878-886. https://www.ncbi.nlm.nih.gov/pubmed/10632617

Limbach, K., Kaufmann, J. M., Wiese, H., Witte, O. W., & Schweinberger, S. R. (2018). Enhancement of face-sensitive ERPs in older adults induced by face recognition training. *Neuropsychologia*, *119*, 197-213. https://doi.org/10.1016/j.neuropsychologia.2018.08.010

Matthews, C. M., & Mondloch, C. J. (2018). Finding an unfamiliar face in a line-up: Viewing multiple images of the target is beneficial on target-present trials but costly on target-absent trials. *Br J Psychol*, *109*(4), 758-776. https://doi.org/10.1111/bjop.12301

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *J Cogn Neurosci*, *9*(5), 605-610. https://doi.org/10.1162/jocn.1997.9.5.605

Megreya, A. M., & Burton, A. M. (2006). Unfamiliar faces are not faces: evidence from a matching task. *Mem Cognit*, *34*(4), 865-876. https://www.ncbi.nlm.nih.gov/pubmed/17063917

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions [Article]. *Nature*, *383*(6603), 812-815. https://doi.org/10.1038/383812a0

Natu, V., & O'Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology*, *102*, 726-747. https://doi.org/10.1111/j.2044-8295.2011.02053.x

O'Neil, E. B., Barkley, V. A., & Kohler, S. (2013). Representational demands modulate involvement of perirhinal cortex in face processing. *Hippocampus*, *23*(7), 592-605. https://doi.org/10.1002/hipo.22117

Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *J Neurosci*, *26*(50), 13025-13036. https://doi.org/10.1523/JNEUROSCI.2481-06.2006

Parvizi, J., Jacques, C., Foster, B. L., Witthoft, N., Rangarajan, V., Weiner, K. S., & Grill-Spector, K. (2012). Electrical stimulation of human fusiform face-selective regions distorts face perception. *J Neurosci*, *32*(43), 14915-14920. https://doi.org/10.1523/JNEUROSCI.2609-12.2012

Pierce, K., Haist, F., Sedaghat, F., & Courchesne, E. (2004). The brain response to personally familiar faces in autism: findings of fusiform activity and beyond. *Brain*, *127*(Pt 12), 2703-2716. https://doi.org/10.1093/brain/awh289

Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple Dissociation of Faces, Bodies, and Objects in Extrastriate Cortex [Article]. *Current Biology*, *19*(4), 319-324. https://doi.org/10.1016/j.cub.2009.01.007

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage*, *56*(4), 2356-2363. https://doi.org/10.1016/j.neuroimage.2011.03.067

Pitcher, D., Ianni, G., & Ungerleider, L. G. (2019). A functional dissociation of face-, body- and scene-selective brain areas based on their response to moving and static stimuli [Article]. *Scientific Reports*, *9*(1), Article 8242. https://doi.org/10.1038/s41598-019-44663-9

Pitcher, D., Japee, S., Rauth, L., & Ungerleider, L. G. (2017). The superior temporal sulcus is causally connected to the amygdala: A combined TBS-fMRI study [Article]. *Journal of Neuroscience*, *37*(5), 1156-1161. https://doi.org/10.1523/JNEUROSCI.0114-16.2016

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a Third Visual Pathway Specialized for Social Perception [Review]. *Trends in Cognitive Sciences*, *25*(2), 100-110. https://doi.org/10.1016/j.tics.2020.11.006

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Curr Biol*, *17*(18), 1568-1573. https://doi.org/10.1016/j.cub.2007.07.063

Platek, S. M., & Kemp, S. M. (2009). Is family special to the brain? An event-related fMRI study of familiar, familial, and self-face recognition. *Neuropsychologia*, *47*(3), 849-858. https://doi.org/10.1016/j.neuropsychologia.2008.12.027

Puce, A., Allison, T., Spencer, S. S., Spencer, D. D., & McCarthy, G. (1997). Comparison of cortical activation evoked by faces measured by intracranial field potentials and functional MRI: two case studies. *Hum Brain Mapp*, *5*(4), 298-305. https://doi.org/10.1002/(SICI)1097-0193(1997)5:4<298::AID-HBM16>3.0.CO;2-A

Ramon, M., Vizioli, L., Liu-Shuang, J., & Rossion, B. (2015). Neural microgenesis of personally familiar face recognition. *Proc Natl Acad Sci U S A*, *112*(35), E4835-4844. https://doi.org/10.1073/pnas.1414929112

Ritchie, K. L., & Burton, A. M. (2017). Learning faces from variability. *Q J Exp Psychol (Hove)*, *70*(5), 897-905. https://doi.org/10.1080/17470218.2015.1136656

Ritchie, K. L., Kramer, R. S. S., Mileva, M., Sandford, A., & Burton, A. M. (2021). Multiple-image arrays in face matching tasks with and without memory. *Cognition*, *211*, 104632. https://doi.org/10.1016/j.cognition.2021.104632

Rotshtein, P., Henson, R. N., Treves, A., Driver, J., & Dolan, R. J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci*, *8*(1), 107-113. https://doi.org/10.1038/nn1370

Schweinberger, S. R., Huddy, V., & Burton, A. M. (2004). N250r: a face-selective brain response to stimulus repetitions. *Neuroreport*, *15*(9), 1501-1505. https://doi.org/10.1097/01.wnr.0000131675.00319.42

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Res Cogn Brain Res*, *14*(3), 398-409. https://www.ncbi.nlm.nih.gov/pubmed/12421663

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: a review of fMRI studies. *Neuropsychologia*, *51*(1), 59-66. https://doi.org/10.1016/j.neuropsychologia.2012.11.006

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, *115 Pt 1*, 15-36. https://doi.org/10.1093/brain/115.1.15

Sliwinska, M. W., Bearpark, C., Corkhill, J., McPhillips, A., & Pitcher, D. (2020). Dissociable pathways for moving and static face perception begin in early visual cortex: Evidence from an acquired prosopagnosic [Article]. *Cortex*, *130*, 327-339. https://doi.org/10.1016/j.cortex.2020.03.033

Sliwinska, M. W., Elson, R., & Pitcher, D. (2020). Dual-site TMS demonstrates causal functional connectivity between the left and right posterior temporal sulci during facial expression recognition. *Brain Stimul*, *13*(4), 1008-1013. https://doi.org/10.1016/j.brs.2020.04.011

Sliwinska, M. W., & Pitcher, D. (2018). TMS demonstrates that both right and left superior temporal sulci are important for facial expression recognition [Article]. *NeuroImage*, *183*, 394-400. https://doi.org/10.1016/j.neuroimage.2018.08.025

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nat Rev Neurosci*, *8*(11), 872-883. https://doi.org/10.1038/nrn2154

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, *127*(1), 105-118. https://doi.org/10.1016/j.cognition.2012.12.001

Tanaka, J. W., Curran, T., Porterfield, A. L., & Collins, D. (2006). Activation of preexisting and acquired face representations: the N250 event-related potential as an index of face familiarity. *J Cogn Neurosci*, *18*(9), 1488-1497. https://doi.org/10.1162/jocn.2006.18.9.1488

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cogn Affect Behav Neurosci*, *9*(1), 122-131. https://doi.org/10.3758/CABN.9.1.122

von Kriegstein, K., & Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biol*, *4*(10), e326. https://doi.org/10.1371/journal.pbio.0040326

von Kriegstein, K., Kleinschmidt, A., Sterzer, P., & Giraud, A. L. (2005). Interaction of face and voice areas during speaker recognition. *J Cogn Neurosci*, *17*(3), 367-376. https://doi.org/10.1162/0898929053279577

Weibert, K., & Andrews, T. J. (2015). Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. *Neuropsychologia*, *75*, 588-596. https://doi.org/10.1016/j.neuropsychologia.2015.07.015

Wiese, H., Tuttenberg, S. C., Ingram, B. T., Chan, C. Y. X., Gurbuz, Z., Burton, A. M., & Young, A. W. (2019). A Robust Neural Index of High Face Familiarity. *Psychol Sci*, *30*(2), 261-272. https://doi.org/10.1177/0956797618813572

Yarmey, A. D., & Barker, W. J. (1971). Repetition versus imagery instructions in the immediate- and delayed-retention of picture and word paired-associates. *Can J Psychol*, *25*(1), 56-61. https://www.ncbi.nlm.nih.gov/pubmed/5157748

Young, A. W., & Burton, A. M. (2017). Recognizing Faces. *Current Directions in Psychological Science*, *26*(3), 212-217. https://doi.org/10.1177/0963721416688114

Yovel, G., Levy, J., Grabowecky, M., & Paller, K. A. (2003). Neural correlates of the left-visual-field superiority face perception appear at multiple stages of face processing. *Journal of Cognitive Neuroscience*, *15*(3), 462-474. https://doi.org/Doi 10.1162/089892903321593162