

Learning faces as concepts improves face recognition by engaging the social brain network

Adva Shoham¹ Libi Kliger,¹ and Galit Yovel^{1,2}

¹The School of Psychological Sciences, Tel Aviv University, Tel-Aviv 6997801, Israel

²Sagol School of Neuroscience, Tel Aviv University, Tel-Aviv 6997801, Israel

Correspondence should be addressed to Adva Shoham, The School of Psychological Sciences, Tel Aviv University, Haim Levanon 30, Tel Aviv, Israel. E-mail: advashoham@mail.tau.ac.il.

Abstract

Face recognition benefits from associating social information to faces during learning. This has been demonstrated by better recognition for faces that underwent social than perceptual evaluations. Two hypotheses were proposed to account for this effect. According to the feature-elaboration hypothesis, social evaluations encourage elaborated processing of perceptual information from faces. According to a social representation hypothesis, social evaluations convert faces from a perceptual representation to a socially meaningful representation of a person. To decide between these two hypotheses, we ran a functional magnetic resonance imaging (fMRI) study in which we functionally localized the posterior face-selective brain areas and social processing brain areas. Participants watched video-clips of young adults and were asked to study them for a recognition test, while making either perceptual evaluations or social evaluations about them. During the fMRI scan, participants performed an old/new recognition test. Behavioural findings replicated better recognition for faces that underwent social then perceptual evaluations. fMRI results showed higher response during the recognition phase for the faces that were learned socially than perceptually, in the social-brain network but not in posterior face-selective network. These results support the social-representation hypothesis and highlight the important role that social processing mechanisms, rather than purely perceptual processes, play in face recognition.

Key words: social brain network; face recognition; face selective area; level of processing; trait inferences

Introduction

Faces are visual stimuli that convey very rich perceptual and social information. The perceptual information refers to our ability to evaluate whether a given face has a large nose, thick lips or large eyes (Valentine, 1991; Abudarham and Yovel, 2016). In addition, based on these perceptual features, humans also make automatic and consistent social inferences, such as how trustworthy or competent a face looks like (Willis and Todorov, 2006; Todorov and Oosterhof, 2011). Interestingly, even though both perceptual and social evaluations rely on judging the visual appearance of facial features, studies have shown that making social evaluations during face encoding significantly improves face recognition relative to perceptual evaluations (Bower and Karlin, 1974; Strnad and Mueller, 1977; Mueller *et al.*, 1978, Schwartz and Yovel, 2019a, 2019b).

Two main hypotheses were suggested to account for this social evaluation benefit in face recognition. In line with the level of processing (LOP) framework (Craik and Lockhart, 1972; Craik, 2002), Bower and Karlin (1974) suggested that trait judgments lead to deeper encoding of faces by creating a rich, semantic network of associations. In contrast to the semantic account, Winograd (1981) proposed a perceptual account for the benefit of social evaluations in face recognition, known as the feature elaboration

hypothesis, according to which trait inferences improve face recognition by encouraging elaborated processing of perceptual information and encoding more facial features.

In a recent study, Schwartz and Yovel (2019b) directly tested the feature elaboration hypothesis by measuring reaction times (RTs) for social and perceptual evaluations during the learning phase. In five different experiments that replicated the better recognition for socially than perceptually evaluated faces, they found that RTs were significantly shorter for social than perceptual evaluations. These findings imply that the social evaluation benefit in face recognition cannot be attributed to more elaborated processing of facial features during learning. With the lack of support for the feature elaboration hypothesis, Schwartz and Yovel proposed that the difference between the representation that is generated following social and perceptual evaluations is not a quantitative perceptual difference but a qualitative conceptual one. In particular, making trait inferences about faces during encoding generates a meaningful social representation of a person, thus converting faces from percepts to socially meaningful concepts, which consequently improves face recognition (Schwartz and Yovel, 2019b). We will refer to this suggestion as the social representation hypothesis.

The importance of social processing during face learning was shown in several studies (Bernstein *et al.*, 2007; Rule *et al.*,

Received: 5 February 2021; Revised: 8 July 2021; Accepted: 16 August 2021

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

2007; van Bavel and Cunningham, 2012; Hugenberg *et al.*, 2013; Wilson *et al.*, 2014; Schwartz and Yovel, 2018). In a series of studies, Hugenberg and colleagues have proposed that better recognition of own than other-race faces can be explained by social motivation and individuation of ingroup members rather than purely perceptual effects (Rodin, 1987; Levin, 1996, 2000; Maclin and Malpass, 2001; Sporer, 2001; Hugenberg *et al.*, 2010). Furthermore, Wilson and colleagues (2014) have shown that expectation for future interaction with people improves face recognition for out-group members. In that study, students who expected future interaction with peers from a different institution performed better in a recognition test of these faces than students who did not have expectations for future interactions. These studies, therefore, emphasized the contribution of social context to face recognition.

Taken together, two possible mechanisms were proposed to account for the social evaluation benefit in face recognition: feature elaboration processes or social-conceptual processes. A feature elaboration account is expected to be mediated by perceptual mechanisms, whereas a social account is expected to be mediated by social processing mechanisms. In the current study, we aim to decide between these two alternative accounts by assessing whether socially evaluated faces engage perceptual or social processing mechanisms during recognition.

To that effect, we conducted a functional magnetic resonance imaging (fMRI) study that enabled us to examine separately the response of posterior face-selective brain areas and social processing brain areas during retrieval of faces that were learned socially or perceptually. Participants performed a face learning task outside the scanner during which they evaluated faces socially or perceptually. The fMRI response to socially vs perceptually learned faces was measured during the recognition phase in posterior face-selective areas (Kanwisher and Yovel, 2006; Duchaine and Yovel, 2015; Yovel, 2016) and social processing brain areas (Alcalá-lópez *et al.*, 2018; Brothers, 1990; Ciaramidaro *et al.*, 2007; Frith, 2007; Frith and Frith, 2003; Mars *et al.*, 2012; Overwalle, 2009; Rebecca Saxe, 2006; Schmäzle *et al.*, 2017). According to the feature elaboration hypothesis, social learning encourages elaborated encoding of facial features (Winograd, 1981), predicting that the posterior face-selective network will show a higher response for socially than perceptually learned faces during the recognition stage. According to the social representation hypothesis, social evaluations convert a face image to a socially meaningful representation of a person (Schwartz and Yovel, 2019), predicting that the social processing brain areas will show a higher response for the socially than perceptually learned faces during the recognition stage. Figure 1 displays the predicted results of the two hypotheses.

Methods

Participants

Twenty participants were recruited for this study (see the Supplementary Material for sample size justification). We collected data from 22 healthy (seven women, ages 19–42 years, 19 right handed) participants, who received payment (\$15/h) for their participation. All participants were native Hebrew speakers, had a normal or corrected-to-normal vision and provided written informed consent, which was approved by the Helsinki committee of the Sheba Medical Centre and Tel Aviv University. Two participants were excluded from analysis, one because his responses were not recorded due to a technical problem and one due to a large number of volumes removed during the motion correction

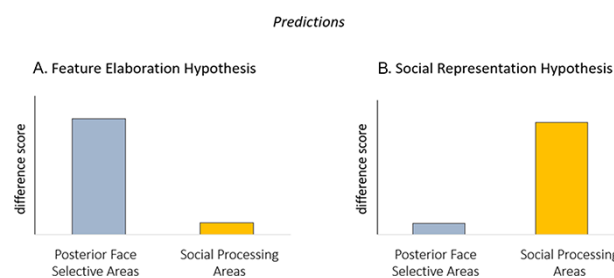


Fig. 1. Predicted results for the two hypotheses. Difference score refers to the difference between the fMRI response to socially than perceptually learned faces. A. The feature elaboration hypothesis (Winograd, 1981) predicts a higher response for the socially than perceptually learned faces in the perceptual face network. B. The social representation hypothesis predicts a higher response for the socially than perceptually learned faces in the social brain network.

procedure (see pre-processing for details), resulting in a sample of 20 participants.

Stimuli

Main experiment

Learning phase (prior to fMRI scan). Stimuli were 12 silent video-clips that depicted a short scene (average duration of 3.75 s) in which three individuals were sitting in a cafeteria, conversing with each other (see Figure 2A). Thus, a total of 36 different individuals (20 females) were presented in the videos. The race of the individuals in the video clips was Caucasian and matched the race of the participants. From the pool of 12 video clips, eight clips (24 identities) were randomly assigned to the learning phase to each participant. Of the eight clips (24 identities) that were presented during the learning phase, four clips (12 identities) were randomly assigned to the perceptual evaluation condition and the other four clips (12 identities) were assigned to the social evaluation condition. The identities in the remaining four clips (12 identities) appeared only during the recognition phase of the experiment as novel identities. The choice to use video stimuli of people during social interaction, instead of the standard static faces that were used in previous studies (Bower and Karlin, 1974; Mueller *et al.*, 1978; Schwartz and Yovel, 2019), was to extend the social evaluation benefit to more naturalistic settings (for review of the importance of studying psychological phenomenon in naturalistic settings see: Shamay-Tsoory and Mendelsohn, 2019; Snow and Culham, 2021).

Recognition phase (during fMRI scan). The recognition phase presented 36 silent colour video clips of the 24 individuals that were presented during the learning phase and 12 novel identities. Each face was cut from a video-clip and was presented on the screen, one at a time, in sequential random order (see Figure 2B). The recognition test video clips were cut from different clips of the same individuals that were taken on the same day as the learning phase video clips.

Functional localizers

Face localizer. Stimuli were silent colour video clips of faces or inanimate objects with their background removed. The face stimuli depicted a single person face down to the shoulders, talking or making facial expressions. The object stimuli were video clips of moving objects, such as a fan or a ticking clock.

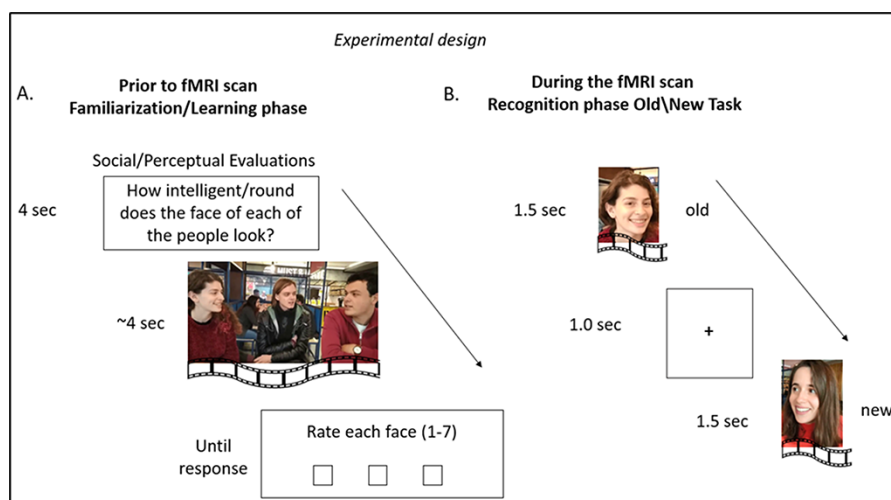


Fig. 2. The experimental design: A. participants were asked to make either social or perceptual evaluations about 24 learned identities presented in a video clip prior to the fMRI scan. B. The recognition phase (during the fMRI scan): participants were asked to indicate whether they recognized each face from the learning phase or whether that was a new face.

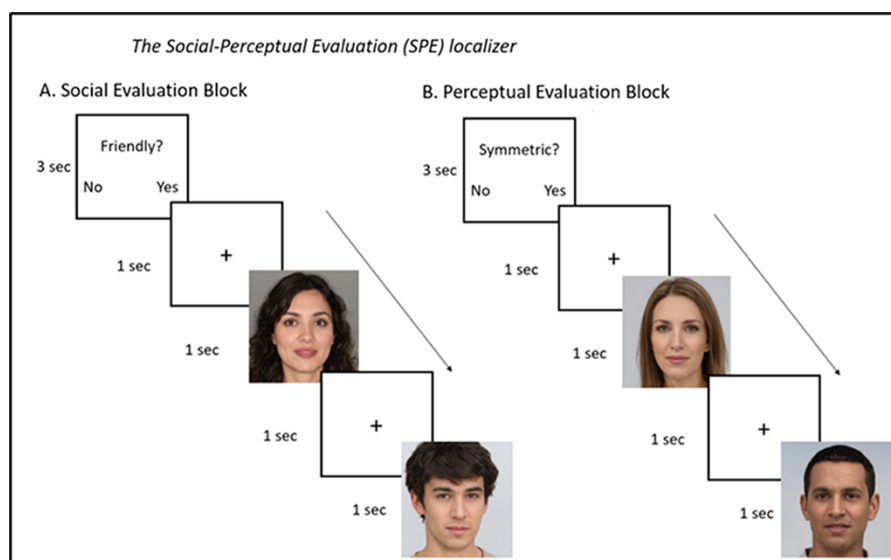


Fig. 3. The SPE localizer includes a social evaluation block (A) and a perceptual evaluation block (B). The participants were asked to judge for each face if it matches the trait that was presented at the beginning of the block, by pressing one key if it does and a different key if it does not match. (For illustration we present AI-generated faces by Generated Photos <https://generated.photos/>).

Social-perceptual evaluation localizer. Stimuli were static colour images of faces of 12 young adults (see Figure 3), all were adopted from the Database of Moving Faces and People (O'Toole *et al.*, 2005). All face images were frontal view, neutral expression with the same background.

Familiar-unfamiliar face localizer. This localizer was included to study the relationship between face familiarity and conceptual and perceptual processing and is beyond the scope of this paper.

Apparatus and procedure

Procedure

The experiment included a 15 min learning phase prior to the MRI scan, in which participants were asked to learn the identities presented in the video-clips for a later recognition test that would take place in the scanner. During the learning phase, participants were asked to make either social or perceptual evaluations about

the 24 learned identities. The learning phase was followed by a single fMRI recording session with three runs of the recognition phase, two runs of the face-object functional localizer, two runs of the social-perceptual evaluation (SPE) localizer and two runs of familiar-unfamiliar face localizer. The three recognition phase runs were presented first, followed by the localizer runs that were presented in an interleaved manner. The fMRI acquisition parameters are presented in the Supplementary Material.

Prior to fMRI scan: learning phase

During the learning phase, participants were asked to learn a total of 24 different identities, presented in eight video clips that included three identities each, and rate them based on either their perceptual appearance or their inferred personality traits. Each trial began with either a perceptual or a social question

from a set of three questions that were used in our previous studies (Schwartz and Yovel, 2019a, 2019b). Social and perceptual mini-blocks of three questions were presented in a random order. Following the clip offset, the participants were asked to rate each identity in the clip based on the question that appeared at the beginning of the trial on a scale of 1–7 (see Figure 2). The social questions were ‘How trustworthy/friendly/aggressive is the face of each of the individuals in the following clip?’. The perceptual questions were ‘How smooth/round/bright is the face of each of the individuals in the following clip?’.

fMRI scan

Recognition phase. Each of the three runs of the recognition phase included two presentations of 36 faces, of which 24 were faces of the learned individuals (half were evaluated based on personality traits and half based on perceptual features), and the remaining 12 faces were novel identities. The video clip stimuli were presented sequentially in one of six different orders that were generated using Optseq2 (Dale, 1999). Participants were asked to press one key for old faces, regardless of whether the faces were learned with social or perceptual questions and a different key for new faces (see Figure 2). Additional information regarding the recognition phase is presented in the Supplementary Material.

Functional localizers

Face localizer. Each run of the functional localizer included 21 blocks: five baseline fixation blocks and eight blocks for each of the experimental conditions: faces and objects. Each block presented 14 stimuli of 15 different video clips of which two repeated twice for a one-back task. Each stimulus was presented for 0.85 s with 0.15 s inter-stimulus interval. Each block lasted 16 s. Each run began with a 6 s fixation (six TRs) of dummy trials and lasted a total of 342 s (342 TRs).

SPE localizer. Each run of the SPE localizer included 21 blocks: five baseline fixation blocks and eight blocks for each of the experimental conditions: social evaluations and perceptual evaluations. Each block started with a social or a perceptual evaluation question from a set of four questions (that we used in our previous studies; Schwartz and Yovel, 2019a, 2019b), followed by a sequence of six different face images (see Figure 3). The participants were asked to decide for each face if it matches the trait that was presented at the beginning of the block, by pressing one key if it does and a different key if it does not. Each block lasted 16 s. The four social questions were: is the face trustworthy/friendly/aggressive/intelligent? The four perceptual questions were: is the face smooth/round/symmetric/wide? Each run began with 6 s (six TRs) of fixation and lasted a total of 342 s (342 TRs). Additional information regarding the SPE localizer procedure is presented in the Supplementary Material.

Data analyses

fMRI pre-processing

The fMRI pre-processing is described in the Supplementary Material.

Regions of interest (ROI) analysis

All extracted ROIs are presented in Table S1 and are detailed in the section below. All ROIs were defined individually for each participant using contrast *t*-maps as clusters of voxels selective to a given category based on the relevant localizer data and contrast (see Table S1).

Face Localizer

Posterior face-selective areas. The posterior face-selective areas were defined based on the face-object localizer for voxels that showed a significantly higher response for faces than objects ($P < 0.001$), in high-level visual cortex: left and right fusiform face area (FFA) within the fusiform gyrus; left and right occipital face area (OFA) in the lateral occipital cortex (LOC); left and right superior temporal sulcus (STS).

Anterior face-selective area (ATL). The anterior face-selective area (ATL) was defined based on the face-object localizer as voxels that showed a significantly higher response for faces than objects ($P < 0.001$) in the anterior temporal lobe.

Social-perceptual evaluation (SPE) Localizer

Social processing brain areas. Social processing regions of interest were defined based on the SPE localizer for the contrast social evaluations > perceptual-evaluations ($P < 0.001$). This contrast revealed the following activations in the majority of the participants (see Table S1): dorsal medial prefrontal cortex (dmPFC); ventral medial prefrontal cortex (vmPFC); precuneus; left and right temporal-parietal junction (TPJ). These areas correspond to the social brain network that was reported in many previous studies (Brothers, 1990; Frith and Frith, 2003; Mitchell et al., 2004; Saxe, 2006; Ciaramidaro et al., 2007; Frith, 2007; Dodell-feder et al., 2011; Mars et al., 2012).

Perceptual processing brain areas. Perceptual processing regions of interest were defined based on the SPE localizer data for the contrast perceptual evaluations > social evaluations ($P < 0.001$) revealed the following activations in the majority of the participants (see Table S1): left and right LOC; left and right superior parietal lobule; left and right inferior frontal gyrus; left and right supramarginal gyrus.

Brain areas defined based on a meta-analysis of an external data set (Neurosynth)

Social network brain areas. We extracted the social brain network with an independent functional localizer. We used a mask of social areas using Neurosynth (Yarkoni et al., 2011) association test of the term ‘social’. This mask reveals a large set of regions including the Amygdala, ATL, dmPFC, vmPFC, precuneus, TPJ and the fusiform gyrus. For the purpose of this analysis, we focused on the dmPFC, vmPFC, precuneus and TPJ for the following reasons: first, these regions were also revealed by our SPE localizer (social > perceptual). Second, these regions overlap with the ‘extended face network’ (Gobbini and Haxby, 2007) that was shown to be involved in social and semantic processing of faces. Third, these regions were shown to work together as a social network (Schmälzle et al., 2017). These regions therefore provide us with an independent method to test the social representation hypothesis. We do also report results in the Amygdala and ATL in the next section and show similar findings in these regions. We divided the mask into different areas using SPM, by the intersection of the mask with suitable boundaries, which are described in the Supplementary Material. We then converted all regions of interest to the native space of each subject individually. One participant was omitted from this analysis due to artefacts in frontal brain areas.

The Amygdala. The amygdala ROI extraction is described in the Supplementary Material.

Main experimental task

The main experimental task was the recognition task of the perceptually and socially evaluated faces that were learned prior to the scan (Figure 2). The time courses of the BOLD signal were extracted from each voxel for each experimental condition in the pre-defined region of interest (ROI, see Table S1) using the MarsBaR ROI toolbox of SPM (Brett et al., 2002). The mean percent signal change of the three-peak time-points (6–8 s) was calculated for each participant, for the type of evaluation (socially learned faces and perceptually learned faces) in each ROI (see Table S1 and Table S2). An analysis of variance (ANOVA) with hemisphere and ROI as factors includes only a small subset of participants who showed activations in all the pre-defined regions. Therefore, the analyses were conducted on the responses of the merged ROIs of each contrast (for more information see the Supplementary Material). A t-test was used to assess whether the effect of the type of evaluation was statistically significant for the merged ROIs of each of the pre-defined contrasts (see Table S1).

ROI selection for testing the hypotheses

To test the feature elaboration and social representation hypotheses, we first examined the response of the posterior face-selective areas (OFA, FFA and STS) and the social processing areas (social > perceptual evaluations based on the SPE localizer) to socially and perceptually learned faces during the recognition phase. Because of the similarity between the learning task and the SPE localizer task, we had to rule out an encoding-retrieval similarity account for the findings. This was done by measuring the response of areas that were defined by the opposite contrast of the SPE localizer (perceptual > social evaluations) to the socially and perceptually learned faces in the recognition test. In addition, we defined social brain areas with an independent localizer based on meta-analysis (Neurosynth) in an attempt to replicate results with our SPE localizer.

Results

Behavioural results

Accuracy

The participants performed the recognition test over three sequential runs. We calculated the averaged performance level for each run for the faces that were learned socially or perceptually (hit rate) as well as for new faces (FA rate). A repeated-measures ANOVA with type of evaluation (social and perceptual) and run (three runs) as within-subject factors, revealed a significant advantage for recognizing socially learned than perceptually learned faces, $F(1, 19) = 8.34$, $P < 0.01$, $\eta_p^2 = 0.30$. No interaction between type of evaluation and run was found $F(2, 38) = 0.81$, $P = 0.48$, $\eta_p^2 = 0.04$ (see Figure 4A) indicating that the magnitude of the social benefit was stable across all runs. Because new faces repeated three times across the three runs, we examined whether the participants did not judge the new faces as learned faces on the later runs by measuring the false alarm (FA) rate across runs: a repeated-measures ANOVA with run as within-subject factor and Huynh-Feldt Sphericity correction, revealed no differences in FA rate across the three runs, $F(2, 38) = 0.7$, $P = 0.48$, $\eta_p^2 = 0.04$ (see Figure 4B). RT analysis revealed no differences between the conditions (see Supplementary Material).

fMRI results

To decide between the feature-elaboration and the social-representation hypotheses, we compared fMRI responses during

the recognition task for the socially learned faces and perceptually learned faces in the posterior face-selective areas and the social processing areas.

The response of posterior face-selective areas and social processing brain areas to socially and perceptually learned faces

According to the feature elaboration hypothesis, social evaluations involve elaborated processing of the perceptual information of faces, predicting higher response of posterior face-selective areas to faces learned with social evaluations compared to perceptual evaluations. According to the social-representation hypothesis, social evaluations convert faces from a perceptual representation to a socially meaningful representation of a person, predicting higher response of social processing brain areas to faces learned with social evaluations compared to perceptual evaluations (see Figure 1). We found that the difference in response to socially relative to perceptually learned faces was larger in social processing brain areas ($M = 0.07$, $s.d. = 0.11$) than the posterior face-selective areas ($M = 0.019$, $s.d. = 0.05$), $t(19) = 2.85$, $P = 0.01$, $d = 0.63$, 95% CI [0.015 0.10] (Figures 5 shows the social-perceptual difference score, Figure 6A shows the response to each condition). A repeated-measures ANOVA with condition (social, perceptual) and brain areas (social processing areas and posterior face-selective areas) as within-subject factors revealed a main effect of condition, $F(1, 19) = 8.71$, $P < 0.01$, $\eta_p^2 = 0.31$ and a significant interaction between condition and brain areas, $F(1, 19) = 8.15$, $P < 0.02$, $\eta_p^2 = 0.30$. Post-hoc tests revealed significantly higher fMRI response for socially learned faces than perceptually learned faces in the social processing areas ($t(19) = 4.01$, $P < 0.001$), and no difference in the posterior face-selective areas ($t(1) = 0.99$, $P > 0.3$). These results are consistent with the social representation hypothesis (see Figure 1). Additional analysis, which also includes the new faces, revealed no difference between perceptually learned faces and new faces, that were both lower than social-learned faces, consistent with a social account (see the Supplementary Material).

The SPE localizer that was used to define the social processing areas is based on a task that is similar to the task that participants performed during the learning stage (social vs perceptual evaluations). Thus, these results may merely reflect an encoding-retrieval similarity effect (for review, see Danker and Anderson, 2010). To rule out this possibility, we performed two additional analyses: we first examined whether the same effect, in the reversed direction, is found in perceptual processing areas that show higher response in the perceptual evaluation than the social evaluation conditions of the SPE localizer task. Second, we examined if the same effect is also found in social areas that were defined with an independent set of coordinates taken from independent data sets (based on Neurosynth, Yarkoni et al., 2011).

Perceptual processing brain areas

If the social effect in social processing areas, defined by the SPE localizer, reflects encoding-retrieval similarity, we expect to find the same effect in the reverse direction (a perceptual effect) in areas that were extracted from the opposite contrast (perceptual evaluations > social evaluations): a paired t-test of type of evaluation (social, perceptual) revealed no difference between perceptually learned faces ($M = 0.08$, $s.d. = 0.13$) and socially learned faces ($M = 0.087$, $s.d. = 0.116$), $t(19) = 0.38$, $P = 0.70$, $d = 0.086$, 95% CI = [-0.025, 0.037] (Figure 6B, left). Therefore, the similarity between the localizer and the learning task does not account

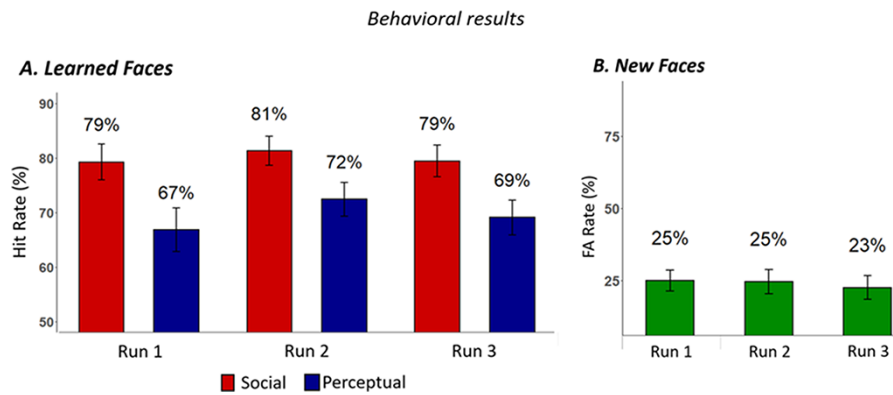


Fig. 4. Behavioural results. A. Recognition level (hit rate) of faces following social and perceptual evaluations. B. False alarm rate for new faces.

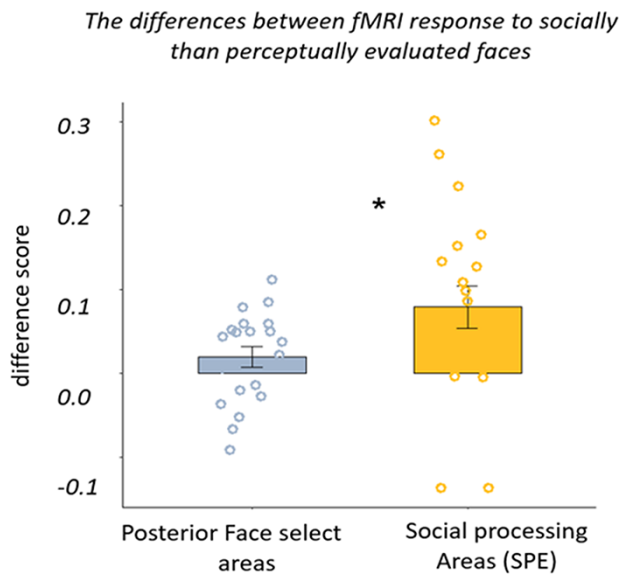


Fig. 5. Difference score refers to the difference between fMRI responses to socially than perceptually learned faces. The difference was significantly larger in the social processing brain areas than the posterior face-selective brain areas, consistent with the social representation hypothesis (see Figure 1). * $P < 0.05$.

for the difference between socially and perceptually learned faces in social processing brain areas.

Social brain areas defined based on an independent localizer (Neurosynth)

To examine if the social effect is also found in social brain areas that are defined by an external localizer, we compared between the response to socially and perceptually learned faces in social brain areas extracted from a social localizer using Neurosynth (see Methods): a paired *t*-test of type of evaluation (social, perceptual) revealed a significantly higher fMRI signal for socially learned faces ($M = 0.002$, $s.d. = 0.087$) than perceptually learned faces ($M = -0.03$, $s.d. = 0.088$), $t(18) = 2.36$, $P < 0.029$, $d = 0.54$, 95% CI [0.004, 0.069] (Figure 6B, right). Therefore, the social effect in social brain areas is not specific to the areas that were extracted using our SPE localizer.

We next examined two additional brain areas that are involved in semantic-emotional processing of faces, the anterior temporal lobe face area (for review, see Wong and Gallate, 2012;

Olson, et al., 2013; Collins and Olson, 2014) and the Amygdala (for review, see Adolphs, 2010). If socially learned faces indeed engage a social-emotional-semantic processing areas the ATL and Amygdala are expected to show a similar preference to socially learned faces (Figure 7). Both ATL and Amygdala showed significantly higher fMRI response for socially learned faces than perceptually learned faces. These analysis are fully described in the Supplementary Material.

Finally, we performed an additional exploratory analysis to examine whether the magnitude of the social effect in social brain areas is correlated with the magnitude of the social effect in memory across participants. To that effect, we computed the Pearson correlation between the difference in recognition performance (hit rate) and fMRI response to socially vs perceptually evaluated faces in the social processing brain areas. This correlation was moderate and not significant ($r(18) = 0.26$, $P > 0.25$). Thus, we found no evidence that individual differences in the magnitude of the social effect in social brain areas is associated with individual differences in the social benefit in recognition performance. Nevertheless, given that socially evaluated faces activated the social brain areas more than perceptually evaluated and new faces during a recognition phase, does suggest that the social brain is engaged during the retrieval of socially learned faces.

Discussion

Social processing of faces during learning improves face recognition relative to perceptual processing (Bower and Karlin, 1974; Strnad and Mueller, 1977; Mueller et al., 1978; Schwartz and Yovel, 2016, 2019b). However, this improved recognition performance does not inform us about the nature of the representation that is generated following these different types of learning strategies. The goal of the current study was to decide between two different mechanisms that were proposed to account for this social evaluation benefit in face recognition. The feature elaboration hypothesis proposes that the recognition benefit is a result of a perceptual enhancement to the face representation (Winograd, 1981), whereas the social representation hypothesis suggests that faces are converted from a perceptual to a socially meaningful representation (Bower and Karlin, 1974; Schwartz and Yovel, 2016, 2019b). To that effect, participants learned faces while making perceptual or social evaluations outside the scanner and performed an old-new face recognition task in the scanner. Behavioural results on the face recognition task replicated the social evaluation benefit, showing better recognition

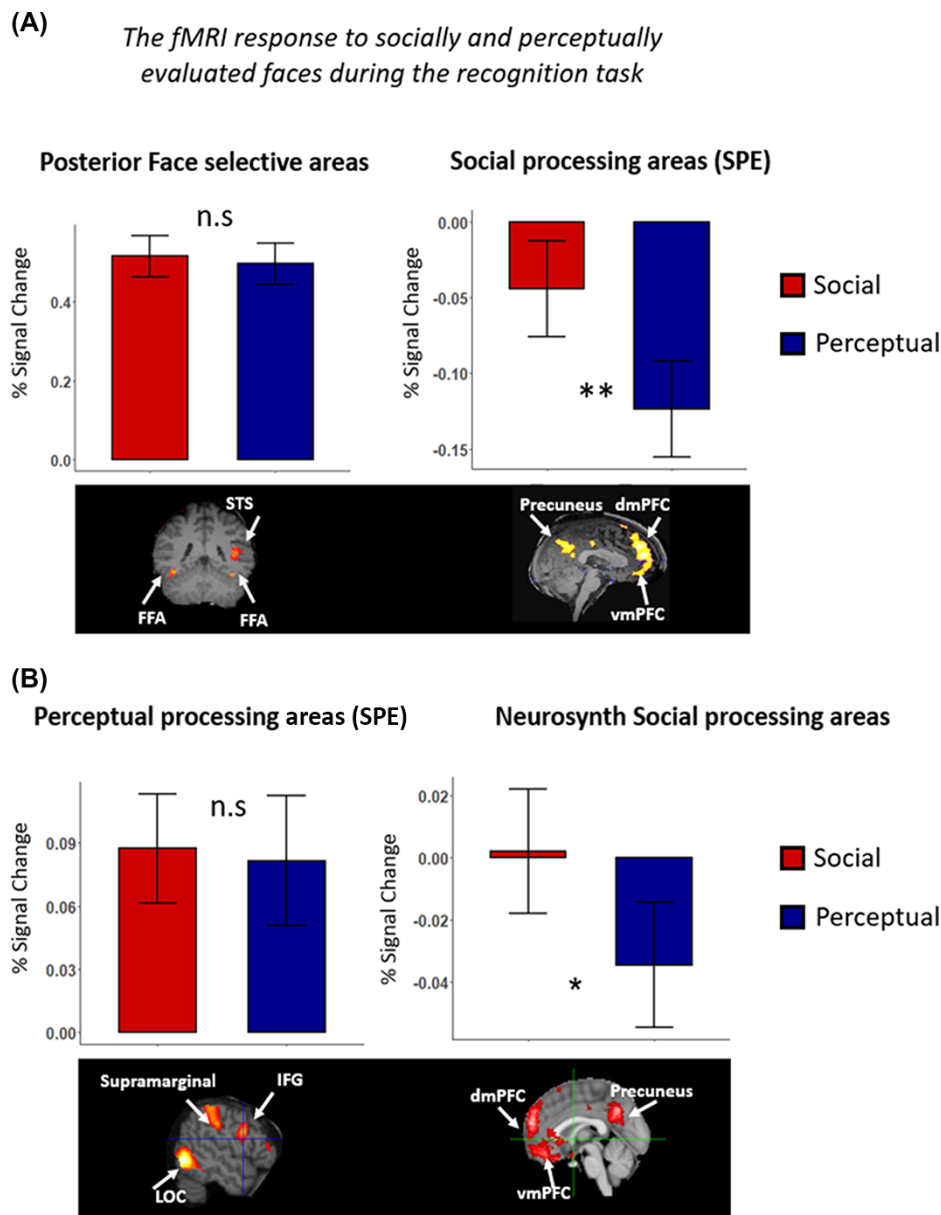


Fig. 6. fMRI results. The differences between fMRI responses during the recognition task, for faces that were learned socially and perceptually. A: In posterior face selective areas (left) In social processing areas defined using the SPE localizer (right). B: In perceptual processing areas defined using the SPE localizer (left). In social brain areas defined by Neurosynth (right). n.s.: non-significant; * $P < 0.05$; ** $P < 0.01$.

for faces that were learned socially than perceptually. Examination of the response of social and perceptual brain areas to these faces during recognition revealed higher response to socially learned than perceptually learned faces in social processing brain areas but not in posterior face-selective areas. These findings are consistent with the social representation hypothesis suggesting that social evaluations convert faces to socially meaningful representations rather than enhance their perceptual representations.

The social processing brain areas in our study were defined as voxels that showed a higher response to socially evaluated than perceptually evaluated faces. This contrast revealed in each participant the known social brain network, including the dmPFC, vmPFC, precuneus and TPJ (Brothers, 1990; Frith and Frith, 2003; Saxe, 2006; Ciaramidaro et al., 2007; Frith, 2007; Mars et al., 2012). Nevertheless, since the participants made social evaluations both

in the study phase of the main task and in the SPE localizer task, it was essential to assure that the effect we found is not due to an encoding-retrieval similarity effect (for review, see Danker and Anderson, 2010). Additional analyses ruled out this interpretation. First, brain regions that showed higher response during perceptual than social evaluations in the SPE localizer showed no difference between perceptually learned and socially learned faces (Figure 6B, left). Second, we found a higher response to socially than perceptually learned faces in social brain areas that were extracted using an independent localizer, based on meta-analysis of previous studies that activated the social brain network (Neurosynth: Yarkoni et al., 2011) (Figure 6B, right). Furthermore, the same effect was also found in social/semantic-related areas, including the amygdala and the face-selective ATL (Figure 7). Thus, we conclude that social evaluations during learning activates social and semantic mechanisms, changing

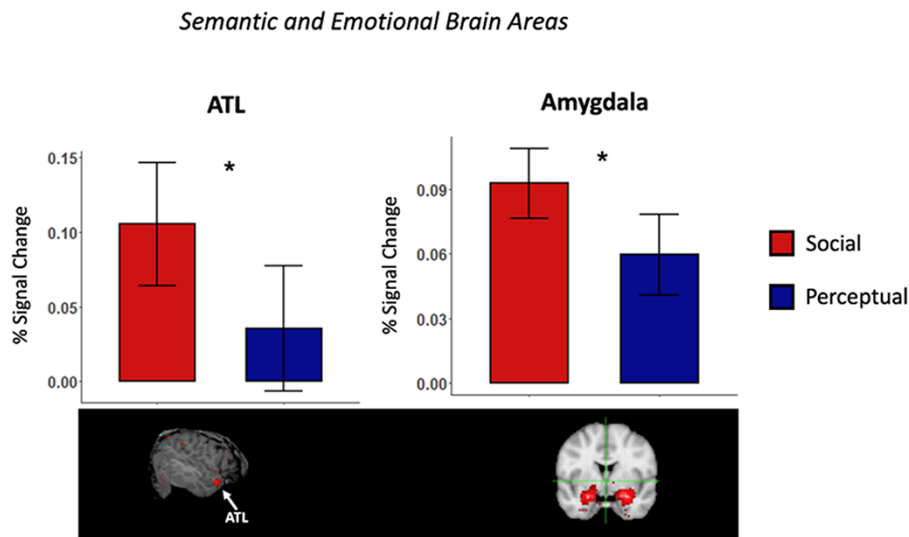


Fig. 7. fMRI results. The differences between fMRI responses during the recognition task, for faces that were learned socially and perceptually. In the anterior temporal lobe (ATL) face area (left) and the Amygdala (right). * $P < 0.05$.

the representation of faces from perceptual images to meaningful social concepts.

Several alternative explanations to the social benefit in face recognition, which were addressed in our previous studies (Schwartz and Yovel, 2019), are noteworthy. First, the improved performance for socially learned than perceptually learned faces may reflect an interference of perceptual evaluations rather than the benefit for social evaluations. This possibility was ruled out by adding a no evaluation condition during learning. Results showed no difference in recognition between faces learned with perceptual or no evaluations, which were both lower than the socially learned faces. This is also consistent with our findings that the response of social brain areas to perceptually learned and new faces was similar (see Supplementary) indicating no evidence for interference for perceptually learned faces. Another alternative account for a social benefit may be that social evaluations generate more variable ratings than perceptual evaluations. Results of our previous as well as our current study (see Supplementary) revealed similar variability of social and perceptual ratings. Finally, perceptual evaluations in previous studies involved local shape-based features (e.g. eye shape and lips), whereas social evaluations may involve global processing of facial information. Our previous study shows that the social evaluation benefit is as large relative to global perceptual evaluations (e.g. face symmetry and skin tone), which were also used in the current study.

The social brain areas that were found in the current study are also reported in studies that have examined the neural response to familiar faces. In particular, Gobinni and colleagues revealed that the TPJ, the precuneus and vmPFC, showed higher response to familiar than unfamiliar faces (Gobinni and Haxby, 2007; di Oleggio Castello et al., 2017). These areas were considered part of the extended face processing system, which processes semantic and emotional information about familiar faces. We propose that social evaluation during encoding mimics the process by which we become familiar with socially relevant faces, as familiar faces are typically encoded in a social context.

The question of whether perceptual or social mechanisms underlie face recognition abilities has been discussed in the context of other well-established face recognition phenomena. For

example, whereas the other race effect, better recognition of own than other race faces, was primarily attributed to perceptual mechanisms (Rhodes et al., 1989; Meissner and Brigham, 2001; Tanaka et al., 2004; Crookes and Rhodes, 2017), later studies have proposed a social account that involves individuation and motivation (Rodin, 1987; Levin, 1996, 2000; Maclin and Malpass, 2001; Sporer, 2001; Hugenberg et al., 2010). In most of these studies, faces were provided with social context during study and show better recognition during test. Nevertheless, it was unclear in what way the representation of these faces changes during social encoding. Here, we show that social encoding engages the social brain network during retrieval, highlighting the important contribution of social processing mechanisms for face recognition.

The roles of social or perceptual mechanisms in face recognition were also discussed in studies that reported impaired face processing in autism spectrum disorder (ASD). ASD is associated with deficits in social interaction (for review see Rao et al., 2008) and with social cognition dysfunction (for review see Pelphrey et al., 2004). One hypothesis that had been proposed to account for the face recognition deficit is that atypical eye contact of individuals with ASD (for review see: Jones et al., 2008; Senju and Johnson, 2009) might lead to perceptual deficits in processing of the eye region of the face (Weigelt et al., 2012). Other studies have emphasized a social motivation account for face processing deficits in ASD (Dawson et al., 2002). According to this hypothesis, impairment in social motivation results in reduced attention to faces as well as to all other social stimuli such as the human voice, hand gestures, and so on (Dawson et al., 2002). This hypothesis is consistent with studies that found that people with ASD show reduced activity in components of the social brain network during social tasks (Fulvia et al., 2002) as well as generally reduced functional connectivity between these areas (Gotts et al., 2012; von dem Hagen et al., 2013). Based on our findings, we predict that reduced activity of the social brain network during face learning and recognition may mediate ASD's poor performance in face recognition, a hypothesis that should be tested in future investigations.

The current study presents a new functional localizer for defining the social brain network in an individual-based manner. There are several tasks that were used in previous studies to define brain areas that are involved in social cognition including the

false-belief task (Saxe and Kanwisher, 2003; Dodell-feder et al., 2011) and the why/how task (Spunt et al., 2014). Both tasks were shown to activate brain regions that are commonly associated with theory of mind and mentalizing tasks, which are also part of the social brain network. For the purpose of the current study, we created the new SPE localizer to individually localize areas that are involved in social and perceptual processing of facial information. The SPE localizer reveals the main areas of the social brain network in each subject individually and can be therefore used in future studies to define the social brain areas in an individual-based manner, taking advantage of the improved statistical power and better localization of the functional ROI approach (Saxe et al., 2006; Kanwisher, 2017).

In summary, the current study aimed to decide between two hypotheses that were suggested to account for the social evaluation benefit in face recognition. We found that social encoding engages the social cognition system, rather than enhancing perceptual processing. Social processing is an integral aspect of face processing as we commonly encode faces for the purpose of social interactions. Thus, social processing should be considered and further explored to fully understand the mechanisms of face recognition. More generally, we show how different types of encoding strategies significantly modify the nature of the representation of the same stimuli, as evident by the different levels of engagement of task-relevant brain systems.

Acknowledgements

None declared.

Funding

The research was partly funded by a joint NSFC-ISF.

Conflict of interest

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

Supplementary data

Supplementary data are available at SCAN online.

References

- Abudarham, N., Yovel, G. (2016) Reverse engineering the face space: discovering the critical features for face identification. *Journal of Vision*, **16**(3), 1–18.
- Adolphs, R. (2010) What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences*, **1191**(1), 42.
- Alcalá-lópez, D., Smallwood, J., Jefferies, E., et al. (2018) Computing the social brain connectome across systems and states. *Cerebral cortex*, (May 2017), **28**(7), 2207–32.
- Bernstein, M.J., Young, S.G., Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, **18**(8), 706–12.
- Bower, G.H., Karlin, M.B. (1974) Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology*, **103**(4), 751–7.
- Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B. (2002) Region of interest analysis using an SPM toolbox. In *8th international conference on functional mapping of the human brain*, **16**(2), 497.
- Brothers, L. (1990) The neural basis of primate social communication. *Concepts Neuroscience*, **14**(2), 81–91.
- Ciaramidaro, A., Adenzato, M., Enrici, I., et al. (2007) The intentional network: how the brain reads varieties of intentions. *45*(13), 3105–13.
- Collins, J.A. & Olson, I.R. (2014) Beyond the FFA: the role of the ventral anterior temporal lobes in face processing. *Neuropsychologia*, **61**, 65–79.
- Craik, F.I.M. (2002) Levels of processing: past, present and future? *Memory*, **10**(5–6), 305–18.
- Craik, F.I.M., Lockhart, R.S. (1972) Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, **11**(6), 671–84.
- Crookes, K., Rhodes, G. (2017) Poor recognition of other-race faces cannot always be explained by a lack of effort.
- Dale, A.M. (1999) Optimal experimental design for event-related fMRI. *Human Brain Mapping*, **8**(23), 109–14.
- Danker, J.F., Anderson, J.R. (2010) The ghosts of brain states past: remembering reactivates the brain regions engaged during encoding. *Psychological Bulletin*, **136**(1), 87–102.
- Dawson, G., Carver, L., Meltzoff, A.N., Panagiotides, H., McPartland, J., Webb, S.J. (2002) Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development*, **73**(3), 700–17.
- di Oleggio Castello, M.V., Halchenko, Y.O., Guntupalli, J.S., Gors, J.D., Gobbini, M.I. (2017) The neural representation of personally familiar and unfamiliar faces in the distributed system for face perception. *Scientific Reports*, **7**(1), 1–14.
- Dodell-feder, D., Koster-hale, J., Bedny, M., Saxe, R. (2011) NeuroImage fMRI item analysis in a theory of mind task. *NeuroImage*, **55**(2), 705–12.
- Duchaine, B., Yovel, G. (2015) A revised neural framework for face processing. *Annual Review of Vision Science*, **1**(1), 393–416.
- Frith, C.D. (2007) The social brain? 671–8.
- Frith, U., Frith, C.D. (2003) Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **358**(1431), 459–73.
- Fulvia, C., Chris, F., Francesca, H., Uta, F. (2002) Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, **125**, 1839–49.
- Gobbini, M.I., Haxby, J.V. (2007) Neural systems for recognition of familiar faces. *Neuropsychologia*, **45**(1), 32–41.
- Gotts, S.J., Simmons, W.K., Milbury, L.A., Wallace, G.L., Cox, R.W., Martin, A. (2012) Fractionation of social brain circuits in autism spectrum disorders. *Brain*, **135**(9), 2711–25.
- Hugenberg, K., Young, S.G., Bernstein, M.J., Sacco, D.F. (2010) The categorization-individuation model: an integrative account of the other-race recognition deficit. *Psychological Review*, **117**(4), 1168–87.
- Hugenberg, K., Wilson, J.P., See, P.E., Young, S.G. (2013) Towards a synthetic model of own group biases in face memory. *Visual Cognition*, **21**(9–10), 1392–417.
- Jones, W., Carr, K., Klin, A. (2008) Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, **65**(8), 946–54.
- Kanwisher, N. (2017) The quest for the FFA and where it led. *Journal of Neuroscience*, **37**(5), 1056–61.
- Kanwisher, N., Yovel, G. (2006) The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **361**(1476), 2109–28.

- Levin, D.T. (1996) Classifying faces by race: the structure of face categories. *22*(6), 1364–82.
- Levin, D.T. (2000) Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *129*(4), 559–74.
- Maclin, O.H., Malpass, R.S. (2001) Racial categorization of faces the ambiguous race face effect. *7*(1), 98–118.
- Mars, R.B., Neubert, F.X., Noonan, M.A.P., Sallet, J., Toni, I., Rushworth, M.F.S. (2012) On the relationship between the “default mode network” and the “social brain”. *Frontiers in Human Neuroscience*, *6*(2012), 1–9.
- Meissner, C.A., Brigham, J.C. (2001) Own-race bias in memory for faces a Meta-Analytic Review. *7*(1), 3–35.
- Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2004) Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience*, *24*(21), 4912–7.
- Mueller, J.H., Carlomusto, M., Goldstein, A.G. (1978) Orienting task and study time in facial recognition. *Bulletin of the Psychonomic Society*, *11*(5), 313–6.
- Olson, I.R., McCoy, D., Klobusicky, E., Ross, L.A. (2013) Social cognition and the anterior temporal lobes: a review and theoretical framework. *Social Cognitive and Affective Neuroscience*, *8*(2), 123–33.
- O’Toole, A.J., Harms, J., Snow, S.L., et al. (2005) A video database of moving faces and people. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(5), 812–6.
- Overwalle, F.V. (2009) Social cognition and the brain: a meta-analysis. *858*, 829–58.
- Pelphrey, K., Adolphs, R., Morris, J.P. (2004) Neuroanatomical substrates of social cognition dysfunction in autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*(4), 259–71.
- Rao, P.A., Beidel, D.C., Murray, M.J. (2008) Social skills interventions for children with Asperger’s syndrome or high-functioning autism: A review and recommendations. *Journal of Autism and Developmental Disorders*, *38*(2), 353–61.
- Rhodes, G., Brake, S., Taylor, K., Tan, S. (1989) Expertise and configural coding in face recognition. *British journal of psychology*, *80*(3), 313–31.
- Rodin, M.J. (1987) Who is memorable to whom: a study of cognitive disregard. *Social Cognition*, *5*(2), 144–65.
- Rule, N.O., Ambady, N., Adams, R.B., Macrae, C.N. (2007) Us and them: memory advantages in perceptually ambiguous groups. *Psychonomic Bulletin & Review*, *14*(4), 687–92.
- Saxe, R., Kanwisher, N. (2003) People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *NeuroImage*, *19*(4), 1835–42.
- Saxe, R. (2006) Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*(2), 235–9.
- Saxe, R., Brett, M., Kanwisher, N. (2006) Divide and conquer: a defense of functional localizers. *NeuroImage*, *30*(4), 1088–96.
- Schmälzle, R., Brook, M., Donnell, O., et al. (2017) Brain connectivity dynamics during social interaction reflect social network structure. *114*(20), 5153–8.
- Schwartz, L., Yovel, G. (2016) The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General*, *145*(11), 1493–511.
- Schwartz, L., Yovel, G. (2019a) Learning faces as concepts rather than percepts improves face recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Schwartz, L., Yovel, G. (2019b) Independent contribution of perceptual experience and social cognition to face recognition. *Cognition*, *183*(2018), 131–8.
- Senju, A., Johnson, M.H. (2009) Atypical eye contact in autism: models, mechanisms and development. *Neuroscience and Biobehavioral Reviews*, *33*(8), 1204–14.
- Shamay-Tsoory, S.G., Mendelsohn, A. (2019) Real-life neuroscience: an ecological approach to brain and behavior research. *Perspectives on Psychological Science*, *14*(5), 841–59.
- Snow, J.C., Culham, J.C. (2021) The treachery of images: how realism influences brain and behavior. *Trends in Cognitive Sciences*.
- Sporer, S.L. (2001) Recognizing faces of other ethnic groups an integration of theories. *7*(1), 36–97.
- Spunt, R.P., Adolphs, R., Perner, J., et al. (2014) Validating the why/how contrast for functional MRI studies of theory of mind. *NeuroImage*, *99*(January), 301–11.
- Strnad, B.N., Mueller, J.H. (1977) Levels of processing in facial recognition memory. *Bulletin of the Psychonomic Society*, *9*(1), 17–8.
- Tanaka, J.W., Kiefer, M., Bukach, C.M. (2004) A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *93*, 1–9.
- Todorov, A., Oosterhof, N. (2011) Source of visual information for identifying people and reading their emotional. *Signal Processing Magazine, IEEE*, *117*(March), 117–22.
- Valentine, T. (1991) A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, *43*(2), 161–204.
- van Bavel, J.J., Cunningham, W.A. (2012) A social identity approach to person memory: group membership, collective identification, and social role shape attention and memory. *Personality & Social Psychology Bulletin*, *38*(12), 1566–78.
- von dem Hagen, E.A.H., Stoyanova, R.S., Baron-Cohen, S., Calder, A.J. (2013) Reduced functional connectivity within and between ‘social’ resting state networks in autism spectrum conditions. *Social Cognitive and Affective Neuroscience*, *8*(6), 694–701.
- Weigelt, S., Koldewyn, K., Kanwisher, N. (2012) Face identity recognition in autism spectrum disorders: a review of behavioral studies. *Neuroscience and Biobehavioral Reviews*, *36*(3), 1060–84.
- Willis, J., Todorov, A. (2006) First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–8.
- Wilson, J.P., See, P.E., Bernstein, M.J., Hugenberg, K., Chartier, C. (2014) Differences in anticipated interaction drive own group biases in face memory. *PLoS One*, *9*, 3.
- Winograd, E. (1981) Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 181–90.
- Wong, C., Gallate, J. (2012) The function of the anterior temporal lobe: a review of the empirical evidence. *Brain Research*, *1449*, 94–116.
- Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D. (2011) Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, *8*(8), 665–70.
- Yovel, G. (2016) Neural and cognitive face-selective markers: an integrative review. *Neuropsychologia*, *83*, 5–13.