Title: Love and friendship at first sight: Rapid neural representations of personal relevance

Mareike Bayer*[1], Oksana Berhe[1,2], Isabel Dziobek[1], Tom Johnstone[3,4]

1 Berlin School of Mind and Brain, Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany
2 Zentralinstitut für Seelische Gesundheit, Mannheim, Germany
3 Centre for Integrative Neuroscience and Neurodynamics, School of Psychology and Clinical Language Sciences, The University of Reading, Reading, UK
4 School of Health Sciences, Swinburne University of Technology, Melbourne, Australia

Corresponding author

Dr. Mareike Bayer
Berlin School of Mind and Brain
Department of Psychology
Humboldt-Universität zu Berlin
Unter den Linden 6, 10099 Berlin
Germany
Email: mareike.bayer@hu-berlin.de

Abstract

Faces are a primary source of social information, but little is known about the sequence of neural processing of personally relevant faces, such as those of our loved ones. We applied representational similarity analyses to EEG-fMRI measurement of neural responses to faces of personal relevance to participants - their romantic partner and a friend - compared to a stranger. Faces expressed fear, happiness or no emotion. Shared EEG-fMRI representations started 100ms after stimulus onset not only in visual cortex, but also regions involved in social cognition, value representation and autobiographical memory, including ventromedial prefrontal cortex, temporoparietal junction and posterior cingulate. According to established models of face recognition, these activations precede the stage of structural face encoding at around 170 ms after stimulus onset. Representations in fusiform gyrus, amygdala, insular cortex and N. accumbens were evident after 200 ms. Representations related to romantic love emerged after 400ms in subcortical brain regions associated with reward. Our results point to the prioritized processing of personal relevance with extensive cortical representation as soon as 100 ms after stimulus onset; preceding the stage of structural face encoding.
Keywords: Personal relevance, emotion, faces, EEG-fMRI, Representation

Significance Statement

Models of face processing, supported by studies using EEG or fMRI, postulate that recognition of face identity takes place with structural encoding in the fusiform gyrus around 170 ms after stimulus onset. We provide evidence, based on simultaneous measurement of EEG and fMRI, that the high personal relevance of our friends' and loved ones' faces is detected prior to structural encoding. These effects start as early as

100 ms after stimulus onset and are not confined to visual cortex, encompassing brain regions involved in value representation, self-relevant processing and social cognition. Our findings imply that our brain can 'bypass' full structural encoding of face identity in order to prioritise the processing of faces most relevant to us.

Faces are arguably the most important social and emotional stimuli we encounter in daily life. They tell us an enormous amount about our fellow humans, including whether they are strangers, friends, enemies, or loved ones, how old they are, how healthy they are, and how they are feeling, both generally, as well as towards us specifically. So important are faces to us, that we appear to be experts at processing them, and have developed specialised neural circuitry to do so (1). Based on a large body of behavioural and neuroimaging research, theories of how different types of face information are extracted from the visual stream have been proposed, with separate processing pathways being suggested for transient (e.g. emotion) and stable (e.g. identity) aspects of faces (2). Yet testing these models is hampered by the lack of temporal resolution needed to delineate several rapidly occurring processing stages, and the almost exclusive focus on faces that carry little or no personal relevance to the observer. Personal relevance is a core aspect of emotion. Appraisals of the relevance of stimuli or situations to an individual's goals and needs are at the heart of appraisal theories of emotion, with the intensity of emotions largely correlated with appraisals of personal relevance (3). The amygdala serves as a general rapid detector of stimulus relevance (4); amygdala responses to emotional stimuli including facial expressions can occur rapidly and outside of awareness (5). The amygdala initiates changes to cognitive, autonomic and sensory systems that heighten awareness, prepare for any necessary response and enhance cortical processing of relevant information (6). For example, the amygdala seems to mediate the amplification of emotional content in early visual areas and the fusiform cortex within 100 ms after stimulus onset (7). Two brain regions that share structural connections with the amygdala, the insula and orbitofrontal cortex, have also been identified as key regions in representing the emotional relevance of faces (8, 9). Other perceived face characteristics, such as trustworthiness (10) appear also to rapidly

engage the amygdala without conscious processing. Behavioural evidence suggests an even broader range of impressions related to relevance, such as threat, likeability, attractiveness, trustworthiness, are consistently made based on very brief (40-100ms) exposure to face stimuli (11, 12).

However, studies examining rapid responses to face stimuli have typically examined responses to unfamiliar faces, which it can be argued are of no *personal* relevance to the participants in an experiment. How quickly personal relevance can be extracted from faces, and where in visual or higher level processing streams this occurs, is largely unknown. FMRI studies on romantic love have compared responses to faces of loved ones with responses to friends or other acquaintances, finding involvement of reward-related subcortical areas like the ventral tegmental areas, caudate and putamen, but also of insula and anterior cingulate cortex, as well as of occipital and fusiform gyrus, superior temporal gyrus and dorsolateral middle frontal gyrus (for review, see 13). Evidence for amygdala involvement is heterogeneous, with some studies reporting activations and others deactivations in response to a loved one's face (14, 15). These studies have not provided evidence on how quickly or in what sequence these neural effects occur. EEG findings suggests that earliest effects of personal relevance of faces might occur at the stage of structural face encoding around 170 ms after stimulus onset (for review, see 16); but were more robustly reported at higher-order processing stages, modulating the amplitudes of the P3 component (17). Findings from the language domain, however, have shown even earlier effects of personally relevant contexts during sensory processing around 100 ms after word onset (18), raising the possibility that information on the personal relevance of faces might also be extracted at an early stage. While the personal relevance of faces has not been the primary focus of previous research, face identity and face familiarity have been widely investigated. Studies using

machine learning have shown that EEG and MEG signals within 100 ms after stimulus onset contain sufficient information for decoding of face identity based on physical features alone (19, 20). Whether face familiarity can be decoded this quickly, based on stimulus properties alone, is unclear. Familiar faces are recognized in a highly robust and effortless way, whereas recognition of unfamiliar faces is more error-prone and relies heavily on image matching (16, 21). Neuroimaging research suggests that this advantage involves a core face processing network comprised of occipital face area (OFA), fusiform face area (FFA), posterior superior temporal sulcus (pSTS), anterior STS, and inferior frontal gyrus, as well as an extended face processing network, which includes brain regions involved in episodic memories (precuneous, anterior temporal cortex), person knowledge (temporoparietal junction (TPJ), medial prefrontal cortex) and emotion processing (amygdala, insula); (22, 23). The extended network shows stronger activation for familiar compared to unfamiliar faces (23–25), which might be the neural basis for behavioural findings of more robust face detection and recognition for familiar faces (16).

While face familiarity in everyday life overlaps considerably with personal relevance (most familiar faces we see are family, friends and colleagues), familiarity does not imply personal relevance. For example, while famous faces can be highly familiar, they are rarely highly relevant for the individual. So far, few studies have contrasted familiarity and personal relevance, but existing evidence suggests that brain activation and behavioural performance are more influenced by the personal relevance for the observer than familiarity per se, with strongest effects for relevant others and loved ones, followed by friends and other personally familiar people, and with weakest effects for famous faces (16, 26). It is unknown at what stage during face processing the brain can detect personal relevance, and specifically whether determining the personal

relevance of a face requires structural face encoding, as assumed by models of face identity processing.

In the present study, we aimed to investigate not only the brain network underlying processing of the emotional and personal relevance of faces, but also the temporal unfolding of brain activity, revealed through simultaneous recordings of EEG and fMRI. We presented heterosexual females in a stable romantic relationship with pictures of their romantic partner, a close male friend, and a male stranger, displaying fearful, happy and neutral facial expressions. We combined EEG and fMRI data using representational similarity analyses (RSA; 27, 28), a subtype of multivariate pattern analysis (MVPA). Instead of investigating amount of activation in a brain region (as typical for fMRI), or activity at a given time (as with EEG/ERP), RSA focuses on the *representational structure* of activations to describe the information space captured by a given measure. This structure is described by representational dissimilarity matrices (RDMs; see Figure 1), composed of pairwise distances between condition-specific activations, reflecting abstractions in information space. Because they are abstracted from the physical properties of the original measures, RDMs can be compared across different modalities, such as EEG and fMRI, despite profound differences in temporal and spatial resolution. Furthermore, data-derived RDMs can be compared to conceptual, model-derived RDMs, in order to test and contrast theoretical predictions. In this study, we compared RDMs centred around each voxel in fMRI data and for each time point in the EEG, thus allowing for the temporally and spatially precise examination of processing of personal relevance and emotional expression. Additionally, we tested fMRI RDMs against model-based RDMs reflecting the processing of personal relevance and emotional expressions.

We predicted that personally relevant faces would elicit increased BOLD activation in the extended face network and increased ERP amplitudes at later processing stages. We further predicted that the fMRI RDM structure in these extended face processing regions would correlate with model-based RDMs representing personal relevance, and with RDM structure of EEG at later processing stages.

Emotional relevance was predicted to increase BOLD responses in amygdala, insula and orbitofrontal cortex, and ERP amplitudes at sensory and later stages (P1 and P3, respectively). We predicted that fMRI RDM structure in these brain regions would correlate with model-based RDMs representing emotional relevance, and with RDM structure of EEG as early as 100ms.

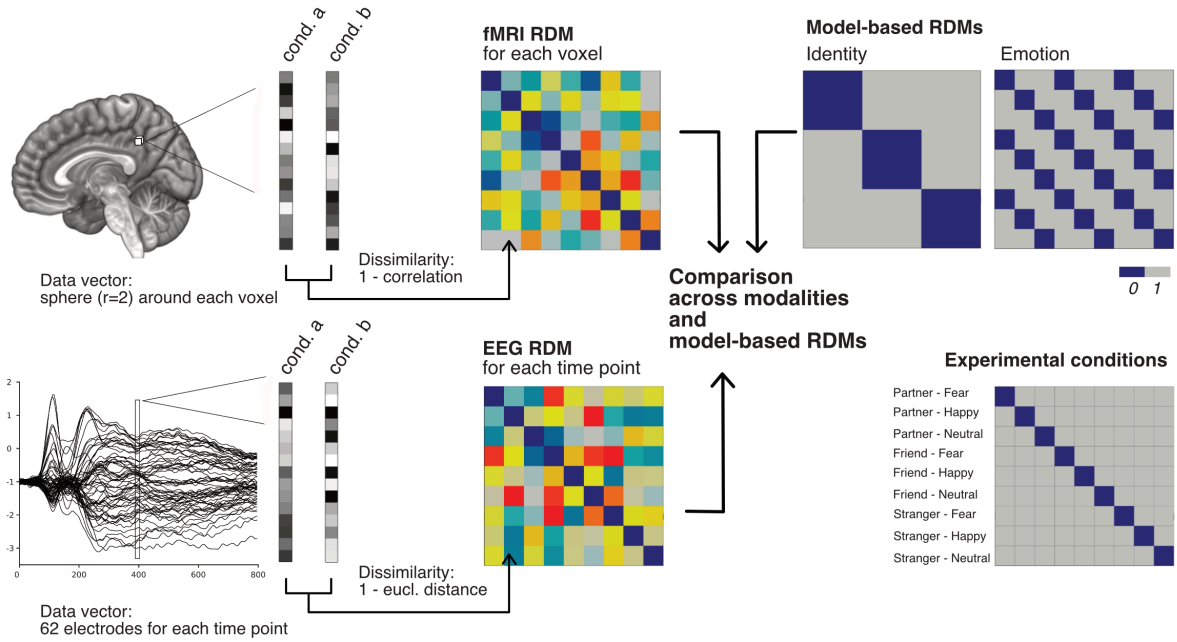

Fig. 1: **Representational Similarity Analyses (RSA).** Representational dissimilarity matrices (RDMs) are constructed for each voxel in the brain and for each data point in grand averaged ERPs by comparing pairwise, condition-specific activations. RDMs are symmetric with a diagonal of zeros, and their size corresponds to the number of experimental conditions, here 9 x 9. Model-based RDMs reflect theoretical predictions.

After construction, RDMs are compared both across modalities and with model-based RDMs.

Results

Unimodal analyses

fMRI

Significantly increased activation for Partner > Stranger and Friend > Stranger was observed across multiple brain regions, including the right fusiform cortex, precuneus, posterior cingulate cortex, anterior cingulate cortex (ACC), right middle temporal gyrus and posterior superior temporal sulcus (STS; Fig. 2). The contrast Partner > Stranger showed additional activation in the bilateral orbitofrontal cortex, while Friend > Stranger yielded additional activation in the ventromedial prefrontal cortex. There was no significant difference for Partner vs. Friend, and no increased activation for the Stranger condition in any of the comparisons. See supplementary table 1 one for a complete list of activated brain regions.

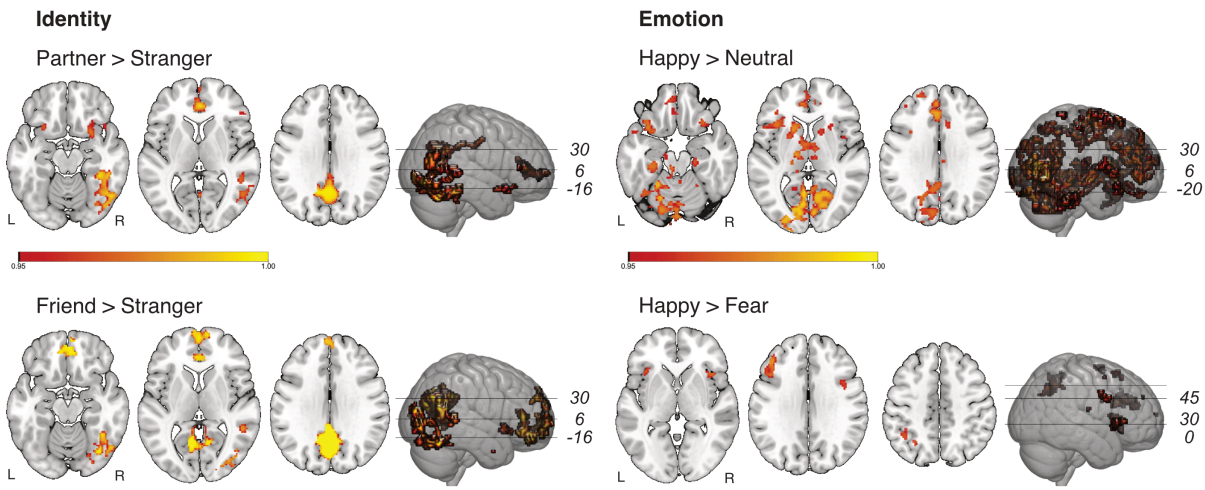

Fig. 2: **Unimodal fMRI results.** Bold activations for Identity comparisons (Partner > Stranger and Friend > Stranger) and Emotion comparisons (happy > neutral, happy >

fear). Activation thresholded with FWE < 0.05 using permutations and threshold free cluster enhancement (TFCE).

For the comparison of emotion conditions (Fig. 2) there was widely distributed activation for Happy > Neutral, including a large cluster in the parieto-occipital cortex (including the intracalcarine cortex, lingual gyrus and precuneus), cerebellum and brain stem. Further clusters were located in anterior brain regions, including the medial and ventromedial prefrontal cortex, orbitofrontal cortex, ACC, inferior and superior frontal gyrus, insular cortex. Finally, significant activation was seen subcortically, including the left amygdala, bilateral thalamus and left caudate.

In the contrast Happy > Fear, significant activation was found in bilateral insula and frontal operculum, and left-lateralized in inferior and middle frontal gyrus and lateral occipital cortex.

Region of interest analyses in amygdala, VTA, putamen, caudate and N. accumbens showed an effect of Identity on bilateral VTA activations, Fs(2,42) > 4.60, ps < .016, η_p^2s > .107, based on increased activations for Partner > Stranger, p = .01 (L VTA). In the left VTA, there was an interaction between Identity and Emotion reflecting that the activation in response to the Friend's face was modulated by the emotional expression, with increased activation for happy expressions compared to fearful and neutral expressions, Fs(1,21) > 7.21, ps < .042. In bilateral caudate activity, a main effect of Emotion, Fs(2,42) > 2.86, ps = .025, η_p^2s > .161, was based on numerically higher activations in response to happy faces, but post-tests were not significant. ROI analyses of Amygdala, putamen and N. accumbens did not reveal significant results. Full statistical analyses are reported in supplementary table 2.

EEG

Analyses of P1 amplitudes showed a main effect of Identity, $F(2,68) = 4.1$, $p = .025$, $\eta_p^2 = .194$, based on higher amplitudes for Partner as compared to Stranger, $p = .017$; no other comparison reached significance (Fig. 3). Likewise, N170 amplitudes were modulated by the factor Identity, $F(2,34) = 9.17$, $p = .001$, $\eta_p^2 = .350$, with higher amplitudes for Partner and Friend compared to Stranger, $ps < .02$. The P3 also showed a main effect of Identity, $F(2,34) = 13.86$, $p < .001$, $\eta_p^2 = .449$; post-tests showed higher amplitudes for Partner compared to both Friend and Stranger, $ps < .014$. A similar pattern was observed for LPC amplitudes, where a main effect of Identity, $F(2,34) = 21.60$, $p < .001$, $\eta_p^2 = .560$, was based on higher amplitudes for Partner compared to Friend and Stranger, $ps < .003$ (Fig. 4). Full statistical analyses are reported in supplementary table S3.

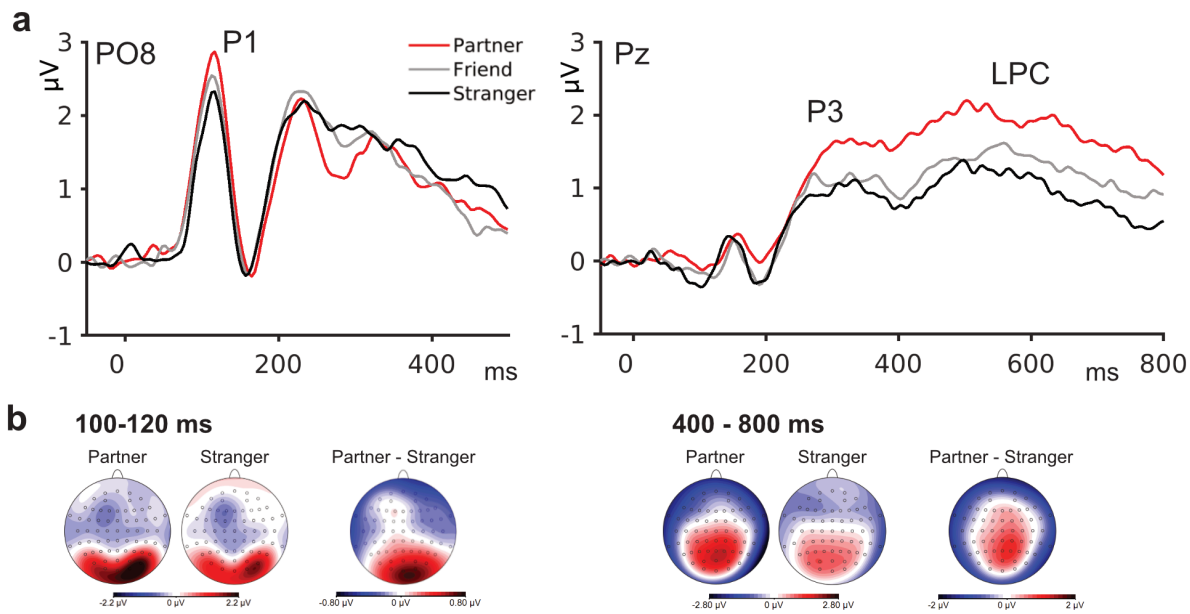

Figure 3: **Effects of personal relevance in event-related potentials. a**, Grand mean ERP waveforms for the factor Identity at electrodes PO8 and Pz, showing increased amplitudes for Partner compared to Friend/Stranger in P1 and P3/LPC amplitudes. **b**, Scalp distributions for Partner, Stranger, and their difference topography at indicated

time intervals show increased activation at sensory processing and higher-order stages of face processing.

Representational Similarity Analyses

Conceptual model RDMs

Face Identity

Using a model-RDM for face identity as a searchlight in voxelwise fMRI analyses corroborated the findings from standard analyses, revealing widespread representations in the core and extended face processing network (see Fig 5). Analyses using a theoretical RDM for familiarity (Partner and Friend vs. Stranger) showed a high level of consistency with representations of Identity, suggesting that face familiarity seems to account for significant amount of the effects observed. In contrast, analysis with an RDM reflecting romantic Love (Partner vs. Friend and Stranger) revealed more focal activations, which were located in bilateral insula, left amygdala, putamen, opercular cortex, lateral occipital cortex and superior temporal gyrus. See supplementary table 1 for a complete list of brain regions.

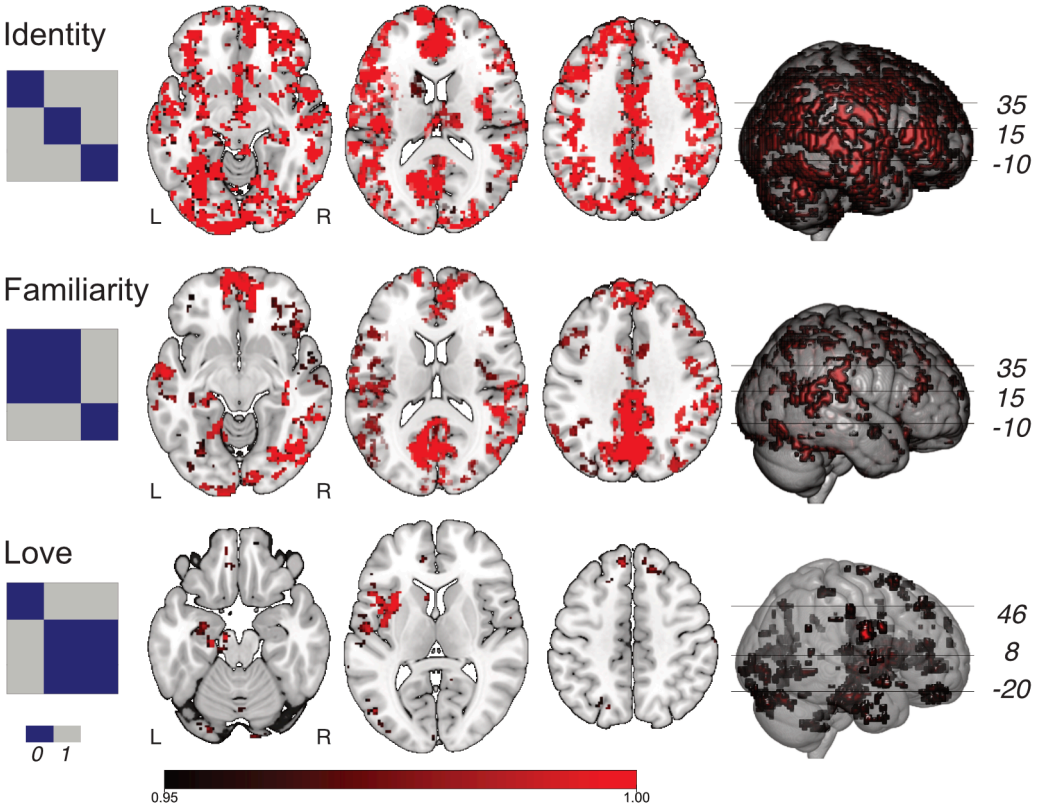

Fig 4: **Representations of conceptual model RDMs for Identity, Familiarity and Love.** Activation thresholded with FWE < 0.05 using permutations and TFCE.

Emotional expression

Comparisons with the Emotion RDM revealed representations that were less widespread than effects of Identity, but included bilateral amygdala, right putamen, the orbitofrontal and ventromedial prefrontal cortex, the temporoparietal junction (TPJ) and left inferior frontal gyrus. A theoretical RDM for valence (distance = 1 from neutral to both fear and happy, distance = 2 between fear and happy) showed activation in most areas of the core and extended face network, including bilateral amygdala, insula, hippocampus, ventromedial and orbitofrontal PFC (Fig. 5). The analyses of theoretical emotional arousal (assuming increased values for fearful and happy compared to neutral faces) revealed no significant representations.

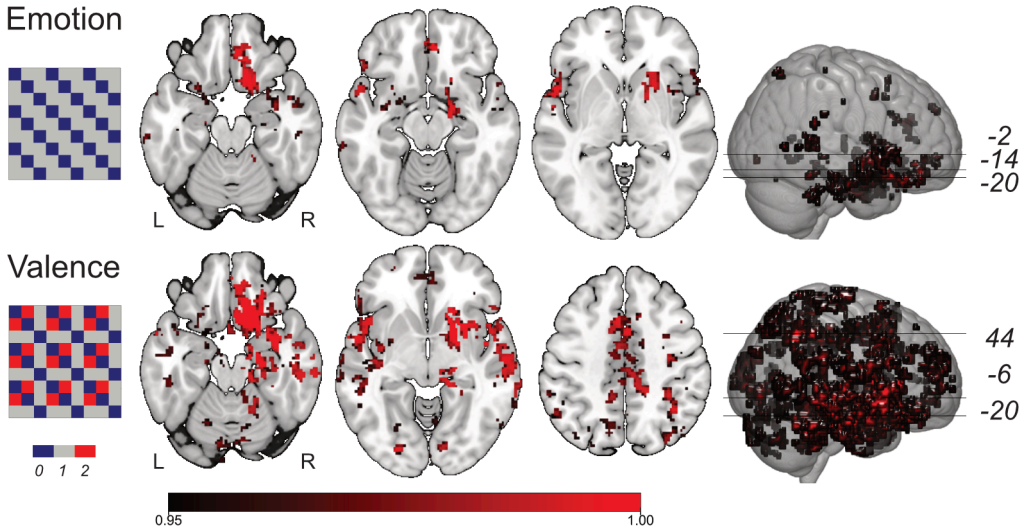

Fig. 5: **Representations of conceptual model RDMs for Emotion and Valence.** Activation thresholded with FWE < 0.05 using permutations and TFCE.

Combined EEG-fMRI RSA Analyses

Peaks in EEG RDM structure were evident at 52 ms, 108 ms, 204 ms, 308 ms, 428 ms, and 660 ms. Accordingly, these RDMs were used as searchlight RDMs in the whole-brain analyses on single-subject fMRI data (Fig. 6).

For the EEG RDM at 52 ms after stimulus onset there were no brain regions with significantly correlated spatial RDMs. EEG representations at 108 ms after stimulus onset mainly represented large distances between responses to Stranger faces, particularly fear, and Partner and Friend faces, with lowest global field power for Stranger fear faces. The RDM showed correspondence to BOLD representations in the extended face network, including the precuneous, posterior cingulate cortex and TPJ, but also in the ventromedial PFC and in the visual cortex (intracalcarine cortex and lingual gyrus). At 204 ms, EEG RDM represented mostly differential activation for familiar faces (Partner and Friend) and Stranger. Corresponding EEG-fMRI representations at this time point were more widespread, additionally including the

ACC, fusiform gyrus, amygdala, insula and N. accumbens. From 308 ms onward, EEG RDMs reflected similar representations for Partner and fearful Friend on the one hand, and Stranger and (happy and neutral) Friend on the other. In EEG-fMRI, corresponding representations at 308 ms were more focal and most evident in the precuneus, posterior cingulate and superior frontal gyrus. Finally, EEG-fMRI representations at 428 ms and 660 ms after stimulus onset were again more extended, with additional significant representations in the middle temporal gyrus, TPJ, in the inferior and superior frontal gyri, as well as in regions of the midbrain. See supplementary table 1 for a complete list of brain regions.

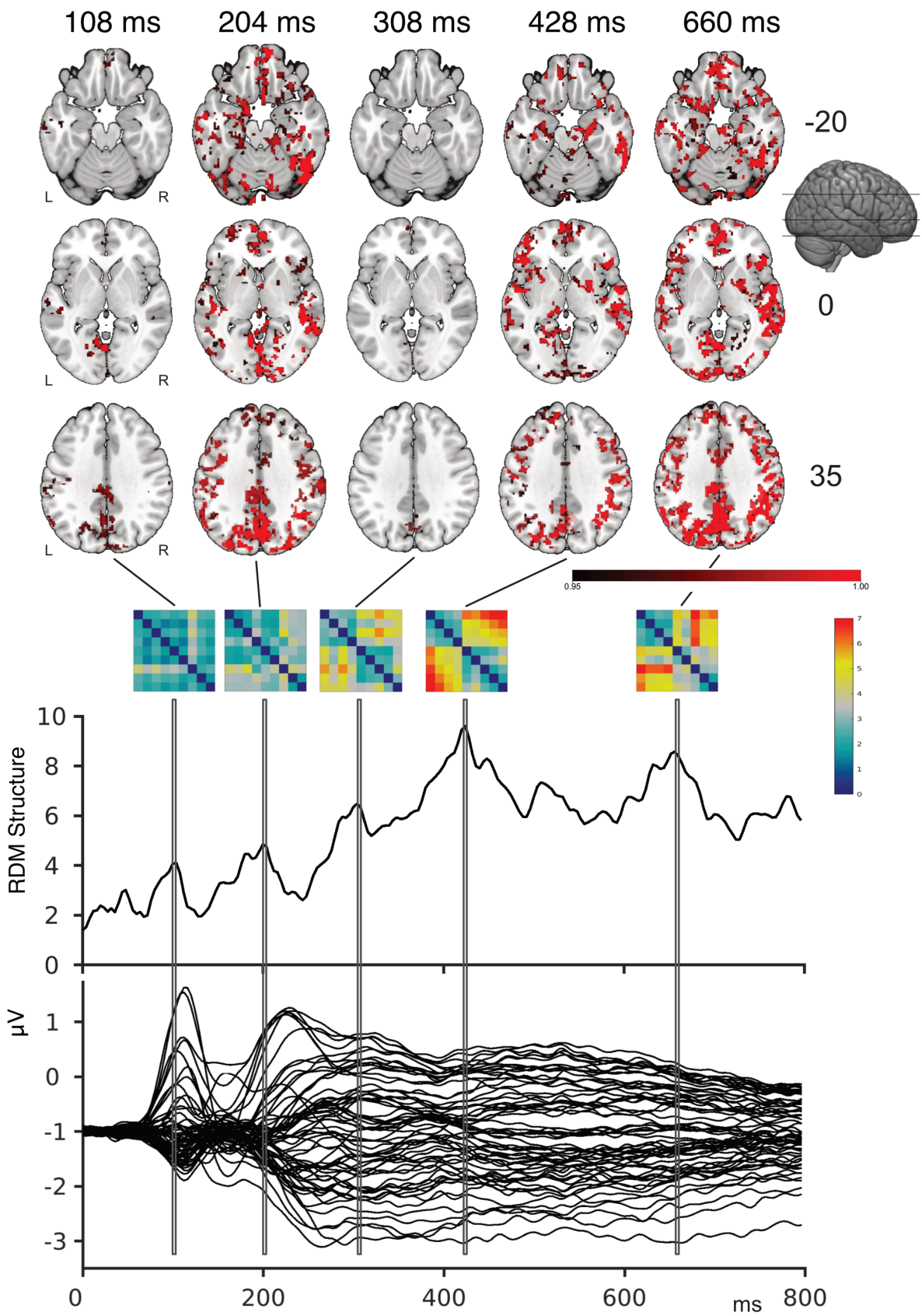

Fig 6: **Combined EEG-fMRI representations.** Brain regions in which fMRI RDMs corresponded to representations in the EEG at the corresponding time points. RDMs

show EEG representations used as target RDMs in the combined EEG-fMRI analyses, selected on the basis of peaks in the RDM structure (top graph). Each line in the lower graph represents the mean time series for one electrode.

Discussion

While it is accepted that emotional facial expressions are processed extremely rapidly, studies of the perception of face familiarity have focussed on the more elaborate and slower processing involved in face recognition. Yet the faces of those closest to us are arguably more personally relevant than the emotional expressions of strangers, and thus might receive a high degree of processing priority. In this study, we investigated the time course of neural processing of emotional and personal relevance in faces using simultaneous EEG-fMRI and representational similarity analyses. Our results point to the strong and rapid impact of personal relevance on face processing, with increased activation in the core and extended face processing network, as well as fast attention allocation in ERPs, and shared representations between EEG and fMRI spatial activation patterns in multiple cortical regions starting as early as 100 ms after stimulus onset.

The effects of emotional facial expressions on fMRI activation in this study were limited to happy compared to neutral and fearful faces, which elicited increased BOLD activation in a network previously associated with emotion processing, including the insular and orbitofrontal cortex, amygdala, and thalamus. These results suggest the preferential processing of happy facial expressions in a personally relevant (experimental) context in healthy subjects. This result is corroborated with our finding of increased model-based representations of stimulus valence (dimensional coding

fearful – neutral - happy) in BOLD activation patterns, rather than undirected emotion categories or arousal values.

The personal relevance of the faces had a far more substantial impact on ERP and fMRI measures than their emotional content. In accordance with previous research, our unimodal fMRI analyses showed that personal relevance increased activation in the core and extended face processing network (16, 22, 25), including fusiform gyrus, posterior STS and inferior frontal gyrus [core network], and in the precuneus, posterior cingulate cortex, anterior cingulate cortex, medial prefrontal cortex as parts of the extended face network.

We augmented our unimodal fMRI analyses with model-based RSA to identify neural activation corresponding to *representations* of personal relevance. Personal relevance was examined both in terms of personal familiarity (Partner/Friend vs. Stranger), and romantic love (Partner vs. Friend/Stranger). Face familiarity accounted for the majority of effects of personal relevance across the core and extended face processing network. In contrast, representations corresponding to the romantic love RDM were more focal and mostly outside the face network, including subcortical regions in putamen, caudate, amygdala, thalamus, cerebellum, and regions of the brain stem including corticospinal and corticobulbar tracts and substantia nigra, but also cortical regions including the insular cortex, medial and ventromedial prefrontal cortex, inferior frontal gyrus and in the intracalcarine cortex. These results are consistent with previous findings that the processing of a loved one's face engages areas of the brain's dopaminergic reward circuitry in the dorsal striatum (putamen and caudate) (13) and substantia nigra. Consistent with this, region-of-interest analyses revealed increased BOLD activation for Partner compared to Stranger in the ventral tegmental area, which is the major source of dopaminergic neurons in the mesolimbic dopamine system. In the left VTA, analyses

also showed an interaction of personal relevance and emotional expressions: While activity was generally increased in response to the partner's face compared to a stranger, activation to the Friend's face was increased only for happy expressions, showing a coding of both identity and emotional expression that reflects the degree of personal reward value.

Event-related potentials revealed that personal relevance modulated sensory perception (P1), structural encoding (N170), and higher-order processing (P3/LPC), with increased amplitudes for the Partner's face. The finding of increased P1 amplitudes is especially noteworthy, since it is related to perceptual processing at around 100 ms after stimulus onset in the extrastriate cortex, and thus precedes structural face encoding as indexed by the N170 component (29), as well as subsequent face identification processes (2). Increased activation in visual areas within 100 ms after stimulus onset have been reported for stimuli with emotional and motivational relevance (30), and have recently been related to associative learning of physical stimulus properties (31–33). Thus, our results are not necessarily in conflict with models of face recognition based on structural encoding and associative memory (2), but might reflect reward value associated with (the physical features of) a loved one's face, extracted prior to structural encoding mechanisms. Corroborating previous findings, analyses of ERP amplitudes of higher-level processing as indexed by the P3 and late positive complex also revealed increased activation for Partner compared to Friend and Stranger. Again, these findings demonstrate that our experimental effects reflect the *personal* relevance of presented faces rather than simply familiarity or identity.

While ERP analyses suggest early modulation of visual perceptual processes by personal relevance, the combination of EEG and fMRI data using RSA analyses allows for the time-resolved analyses of representations within the fMRI. These analyses suggest fast modulation of neural processing across a far more widespread collection of cortical regions. Shared EEG-fMRI representations were first apparent as early as 100 ms after stimulus onset not only in the visual cortex, but also in the ventromedial and medial prefrontal cortex, regions involved in value encoding and self-referential processing (34, 35). Interestingly, the medial prefrontal cortex with its ability to encode social value and reward, was recently also discussed as a guiding structure in infants' cortical face specialisation (36). Further early representations were observed in regions linked to multimodal sensory integration and theory of mind (TPJ) (37), and episodic and autobiographical memory (posterior cingulate cortex) (38). The fast and widespread activation of brain areas involved in social cognition and reward encoding serves to highlight the prioritization of social information, especially personally relevant information. At 200 ms, shortly after the stage of structural face encoding, representations additionally included the fusiform gyrus, but also amygdala, insular cortex and N. accumbens. This time course of representations of face relevance suggests that the response of subcortical relevance- and reward related structures like the amygdala and the N. accumbens might rely on the output of structural face encoding at around 170 ms after stimulus onset. Earlier observed effects of personal relevance at around 100 ms after stimulus onset, might thus be based on fast cortico-cortical connections involving the prefrontal cortex (33).

Finally, at the stage of higher-order processing from approximately 400 ms after stimulus onset, representations were identified in all regions of the core and extended face processing network, including amygdala, insular and orbitofrontal cortex (22), but

also in regions identified with the theoretical RDM for romantic love, like putamen, cerebellum and regions of the brain stem. Consistent with EEG RDMs, ERPs showed evidence for differential processing of romantic partners at a similar time scale. Across modalities and analyses types (unimodal analyses, multimodal and theoretical RSA), our results suggest that early processing mainly reflects the amplification of familiarity (i.e., Partner and Friend compared to Stranger); whereas effects specific for romantic love only emerge at a higher-order processing stage. However, it is important to keep in mind that the terms 'familiarity' and 'love' in this case are not exclusively related to Friend and Partner, respectively, but reflect different aspects of close relationships (39): Both Partner and Stranger are highly personally relevant; and one can feel friendship love towards a close friend. On the other hand, a romantic partner can also serve the role of a friend. Thus, romantic love in our study refers to representations that are unique to a romantic relationship, over and above a close friendship.

In this study, emotional expressions were associated with relatively weak effects; while effects related to personal relevance clearly dominated. This should not come as a surprise. Being able to rapidly identify and respond to those closest to us is a fundamental ability present from early infancy (40). It also suggests that emotional expressions are not processed in a rigid, automatic way, but that brain responses rather reflect context-specific relevance of social stimuli. Emotional facial expressions, even of a friend or loved one, devoid of meaningful context (e.g. in a lab) are of limited direct relevance. De-contextualized emotional expressions on the faces of *strangers,* as used in the vast majority of studies on facial emotion, are even less relevant to the individual. Our results therefore present a compelling case for including stimuli tailored to the

individual in future research, rather than relying on standardised datasets of largely personally irrelevant stimuli.

Understanding the time course and neural circuitry of processing both emotional and personal relevance might be of special importance in clinical conditions, where processing of emotional or social information is specifically disrupted. Autism spectrum conditions (ASC) are one such example: A vast body of research reports atypical processing of (emotional) faces; (41); however, research also suggests typical processing of personally relevant faces (42) . Therefore, atypical (emotional) face processing might reflect altered personal relevance attached to strangers, rather than a dysfunction of the neural face processing architecture, with potential consequences for the design of targeted interventions.

What is clear from this study is that in basic social neuroscience, as well as in clinical research, there are compelling reasons to study responses to individualised social stimuli. After all, by far the most pervasive social stimuli we encounter in our daily lives, the ones our brains are most attuned to, are our close friends and loved ones.

Methods

The study was reviewed and approved by the University of Reading research ethics committee. All participants provided informed consent before taking part in the study.

Participants

Data were collected from 22 female participants (mean age = 19.8 years, sd = 0.9 years). For four participants, no EEG data was available (due to technical problems and insufficient data quality). All participants were in a heterosexual romantic relationship

at the time of data collection (mean duration = 20.0 months, sd = 14.6 months; friendship: mean duration = 35.9 months, sd = 26.5 months). Participants received a mean score of 105.9 points (sd = 10.7) of 135 points on the Passionate Love Scale (43). They reported high content both with their relationship (mean = 8.8/10, sd = 1.3) and their friendship (mean = 8.3, sd = 1.3). All participants had normal or corrected-to-normal vision; 20 participants were right-handed. Participants were recruited through the Undergraduate Research Panel and Internet ads; they received course credit or £25 for participation.

Stimuli

Stimuli consisted of portraits of the participant's boyfriend, a male close friend, and a male stranger, displaying fearful, happy, and neutral facial expressions (3 x 3 design). All stimuli were obtained by taking screen shots during a Skype session prior to the experimental session. At the start of this session, information was about the study and the possibility to withdraw from participation. The session started by taking portraits with a neutral expression. Prior to happy and fearful expressions, volunteers were asked to remember a happy/frightening scene or memory in order to induce the respective mood state. Portraits were edited using Affinity Photo software; they were cut to size, adjusted in luminance, and received a light-grey background. Portraits of the male stranger were obtained in the same way; all participants were presented with the same stranger.

Participants completed ratings of stimulus valence and arousal on the experimental stimuli, as well as on attractiveness (only on the neutral expression) using 7-point Likert scales. Valence ratings confirmed differences between emotion categories, $F(2,42) = 69.93$, $p < .001$, $\eta_p^2 = .769$, reflecting expected differences between all categories (Happy

> Neutral > Fearful), all *ps* < .002. Furthermore, valence ratings showed a main effect of Identity, $F(2,42) = 18.43$, $p < .001$, $\eta_p^2 = .467$, based on more positive ratings for Partner compared with Friend, $p = .006$, and Stranger, $p < .001$, as well as for Friend compared with Stranger, $p = .021$. Results also showed an interaction between Identity and emotion categories, $F(4,84) = 7.08$, $p < .001$, $\eta_p^2 = .252$. Post-tests of Emotion within Identity reflected that the difference between fearful and neutral faces was only significant for Stranger, p < .001 (for full details see supplementary table 4). Arousal ratings showed a main effect of emotion category, $F(2,32) = 14.02$, $p < .001$, $\eta_p^2 = .400$, with higher arousal ratings for happy compared to fearful and neutral faces, *ps* < .015. A main effect of Identity, $F(2,42) = 23.37$, $p < .001$, $\eta_p^2 = .527$, reflected higher arousal ratings for Partner compared with Friend and Stranger, *ps* < .001. Finally, there was an interaction between emotion category and Identity, $F(4,84) = 4.10$, $p = .004$, $\eta_p^2 = .163$; for follow-up analyses see supplementary table 4.

Attractiveness ratings revealed an effect of Identity, $F(2,42) = 32.71$, $p < .001$, $\eta_p^2 = .609$; participants rated their Partner as more attractive than their Friend and the Stranger, *ps* < .001, whereas Friend and Stranger were rated as equally attractive. Full statistical analyses are reported in supplementary materials; rating values are reported in table 1.

Table 1. Rating values (means and standard deviations) of experimental stimuli

	Attractiveness *(1-7)*	**Valence** *(-3 -3)*			**Arousal** *(1-7)*		
		Fearful	**Happy**	**Neutral**	**Fearful**	**Happy**	**Neutral**
Partner	6.1 (1.1)	-0.4 (1.5)	2.3 (0.7)	0.0 (1.8)	3.2 (1.7)	5.4 (1.5)	3.7 (1.8)
Friend	4.1 (1.4)	-1.1 (1.1)	1.6 (0.8)	-0.6 (1.1)	2.1 (1.6)	2.9 (1.9)	2.5 (1.6)
Stranger	4.1 (1.2)	-2.0 (1.1)	0.5 (1.1)	-0.6 (1.2)	1.6 (1.1)	3.2 (2.0)	2.5 (1.6)

Procedure

After receiving information about the study, participants provided informed consent and were fitted with the EEG cap. Inside the scanner, participants performed a passive face-viewing paradigm with occasional 1-back trials.

Face stimuli were presented for 1s. Every face stimulus was presented 40 times, resulting in 360 experimental trials. The sequence of stimuli and the jittering of the inter-trial interval was determined by the program fMRI simulator (44), in order to maximise statistical power (mean ITI = 3000 ms, min ITI = 2500 ms, using an exponential function as described in 45). During the ITI, a fixation cross was presented in the middle of the screen. Stimuli were presented in 10 blocks of 36 stimuli each; additionally, each block contained four 1-back trials in order to ensure participants' attention to the faces. In these trials, a question mark was presented directly after a face, followed by another face. Participants had to indicate by button press whether the face presented after the question mark was identical to the one presented before the question mark (both in identity and emotional expression) or not. The face remained on the screen until participants had given a response using the response pad. Short breaks were provided between blocks. Stimuli were presented on a Nordic Neuro Labs goggle system at 60 Hz on an 800 x 600 pixel screen using EPrime software (Psychology Software Tools, Inc.).

Data acquisition and pre-processing

fMRI

Data were collected on a 3T Siemens Trio MRI scanner using the standard 12 channel head coil. Functional images were acquired continuously with a T2*-weighted gradient EPI echo sequence (40 interleaved axial slices, phase encoding P to A, 2.5 x 2.5 mm

voxels, slice thickness = 2.5 mm, interslice gap = 0.25mm, 100x100 in-plane matrix, TR = 2500 ms, TE = 30 ms, Flip Angle: 90°). A high-resolution T1-weighted whole-brain structural image was acquired using an MPRAGE sequence (1mm isotropic voxels, FOV = 160 x 256 x 256 mm, Flip Angle: 9°).

Data were processed with FSL 5.0 (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Brain extraction was performed using the FSL Brain Extraction Tool (46). Prior to statistical analyses, we applied denoising (Marchenko-Pastur Principal Component Analyses) and motion correction (MCFLIRT; 47). Data were spatially smoothed using a 5 mm full-width at half-maximum Gaussian kernel (SUSAN; 48), and grand-mean intensity normalized. Motion artefacts were removed with ICA-AROMA (49) and data were high-pass filtered with a cutoff frequency of 100 s. Single subject 4D data were registered to the subject's high-resolution structural image with FLIRT (50) using a BBR cost function. Registration of the individual high-resolution structural image to the Montreal Neurological Institute (MNI) template (2mm isotropic MNI 152 standard space image) was performed using ANTS (51). Finally, transformations were combined for the registration of the functional data to standard space.

EEG

Continuous EEG data were collected from 64 Ag-AgCl electrodes attached to an MRI-compatible EEG cap (BrainCap MR, BrainProducts) simultaneous to fMRI acquisition using an MRI-compatible amplifier system (BrainAmp MR plus, BrainProducts). An additional electrode was placed on the back of the participant (left of the spinal column) for recording the electrocardiogram (ECG). Online, data were referenced to electrode FCz; electrode AFz was used as ground. Electrode impedances where kept below 20 kΩ; data were recorded with a high frequency cutoff of 250 Hz, a low frequency cutoff of 0.1

Hz, and a sampling rate of 5000 Hz. A sync box (BrainProducts) was used to synchronise the MRI and EEG computer clocks to ensure optimal synchronisation of EEG recording and MRI slice acquisition.

Offline, MR gradient artefacts were identified using synchronisation markers from the scanner and removed using a modified version of the template subtraction algorithm (52) as implemented in BrainVision Analyzer 2 (BrainProducts). Gradient artefacts were removed from continuous, baseline corrected data (using the whole artefact for baseline correction) with a sliding window of 21 artefacts. After correction, data were down-sampled to 250 Hz and low-pass filtered using an FIR filter of 70 Hz cutoff. Ballistocardiographic artefacts were identified using the ECG channel with a semiautomatic template matching procedure; correction was performed using a template subtraction approach based on a sliding window of 15 pulse intervals. In order to identify additional artefacts caused by eye blinks, eye movements and residual ballistocardiographic artefacts, we performed a restricted Infomax ICA (53) and removed artefact components based on their topography and time course. Data was re-referenced to average reference and segmented into epochs from -100 ms to 800 ms relative to stimulus onset, and baseline-corrected using a 100 ms pre-stimulus baseline. Trials with activity exceeding ± 100 μV or voltage steps larger than 100 μV were marked as artefact trials (0.6 % of trials) and excluded from further analyses. Data were averaged per participant and experimental condition.

Data analyses

All voxelwise fMRI statistical tests were one-sided. All other tests were two-sided unless otherwise stated.

Unimodal analyses

fMRI

A first level GLM was applied with separate regressors for emotion by identity conditions (9 regressors) and for 1-back trials; regressors were created by convolving the temporal profile of each experimental condition with the double gamma haemodynamic response function in FSL. Additional nuisance regressors without convolution were included to model breaks (between blocks) and artefacts (framewise displacement > 1mm). Contrasts of interest included effects of Identity and Emotion (pairwise comparisons of identity conditions and emotion conditions, respectively) and were entered into a mixed-effects GLM. Whole-brain analyses were thresholded using permutation tests and threshold-free cluster enhancement to ensure a corrected Familywise Error (FWE) of < .05. During permutation tests, we used a grey matter mask (MNI gray matter tissue priors thresholded at 0.3) and applied variance smoothing (2mm). Additionally, we performed region-of-interest (ROI) analyses for bilateral activations of amygdala, Nucleus accumbens, putamen and caudate (masks created from Harvard-Oxford subcortical structural atlas) and ventral tegmental area (VTA; 54; mask retrieved from neurovault.org/collections/1380). All masks were thresholded at 25% and binarised. Within these regions of interest, we extracted subject-wise covariation parameter estimates for all experimental conditions and performed repeated-measures ANOVAs with the factors Identity and Emotion (3 x 3). Huynh-Feldt correction was applied for violations of sphericity in rm-ANOVAs; in post-tests, p-values were corrected for multiple comparisons using Bonferroni correction.

EEG

EEG analyses were performed on four ERP components of interest, P1, N170 and P3 and LPC. P1 amplitudes were quantified on a group of occipital electrodes (PO8, PO4, POz, PO3, PO7, O1, Oz, and O2). Semi-automatic peak detection was performed for each condition on the average of the region of interest in the time window from 90 to 130 ms after stimulus onset (mean peak latency = 107ms). N170 peak amplitudes were detected on averaged electrodes TP9, TP7, TP8, TP10, P7 and P8 in the time window of 150 to 220 ms after stimulus onset (mean peak latency = 172ms). In order to account for differences in the preceding P1 component, N170 amplitudes were subtracted from P1 amplitudes for each condition before statistical analyses were performed. P3 and LPC amplitudes were analysed at electrodes, CP1, CPz, CP2, P1, Pz, P2 and POz in the time windows of 300 to 400ms (P3) and 400 to 800 ms (LPC). Analyses were performed with repeated-measures ANOVAs including the factors Identity (3), Emotion (3) and, for P3 and LPC, electrode (7). Huynh-Feldt correction was applied for violations of sphericity in rm-ANOVAs; in post-tests, p-values were corrected for multiple comparisons using Bonferroni correction.

RSA analyses

RSA analyses were performed using the CoSMoMVPA toolbox (55) on Matlab R2017b. For plotting, we further used the toolbox for representational similarity analyses (56).

fMRI

Representational dissimilarity matrices (RDMs) were constructed for each voxel in the brain, separately for each subject, based on z-values of each of the 9 experimental conditions included in the mixed-effects GLM. For each voxel and condition, we created a vector (based on a sphere around each voxel, using a radius of 2 voxels). We then quantified the dissimilarity between each pair of experimental conditions as 1 –

Pearson's R of the two corresponding vectors. This resulted in a 9 x 9 RDM for each voxel and each participant.

EEG

RDM matrices were constructed using grand-averaged waveforms. For each time point from stimulus onset to 800 ms after stimulus onset, the distance between pairs of experimental conditions was quantified as their Euclidean distance across all 62 scalp electrodes. Euclidean distance was used in order to take amplitude differences into account, which convey essential information in event-related potentials. Analyses resulted in a 9 x 9 RDM for each time point, see supplementary figure 1.

In order to select EEG time points as target RDMs in the fMRI analyses, we derived a measure of internal structure of each time point's RDM by computing the Euclidean distance of RDM cell values to the arithmetic mean of cell values. As a result, time points with pronounced differences of dissimilarity values (RDM cell entries) between condition pairs receive high values, whereas smaller differences result in low values. For these calculations, we used one half of the (symmetrical) RDM, excluding the diagonal zeros.

Conceptual model RDMs

Conceptual model RDMs were used as searchlight with voxel-wise fMRI RDMs in order to identify brain regions representing different aspects of the face stimuli. Conceptual model RDMs for Emotion and Identity were created based on the assumption of high similarity within experimental categories and low similarity across categories (coded as 0 and 1, respectively). We also computed conceptual model RDMs for face familiarity (Partner/Friend vs. Stranger) and romantic love (Partner vs. Friend/Stranger; see Figure 4, and for emotional valence (distance between happy/fearful vs. neutral = 1,

distance happy vs. fearful = 2; Figure 5) and arousal (distance happy/fearful vs. neutral = 1, distance happy vs. fearful = 0).

Joint EEG-fMRI

For combination of EEG and fMRI data, we performed representational similarity analyses using EEG RDMs as a searchlight with each participant's individual fMRI RDMs for each voxel using Pearson's R. Resulting brain maps of similarity were combined across subjects using permutation tests and threshold-free cluster enhancement in order to ensure a corrected Familywise Error (FWE) of < .05.

Acknowledgements

We thank Michael Lindner and Catriona Scrivener for assistance with data collection and Luca Brivio for assistance with data analysis. M.B. was supported by the Berlin School of Mind and Brain (Humboldt-Universität zu Berlin). T.J. was supported through a travel grant from Berlin School of Mind and Brain (Humboldt-Universität zu Berlin).

Author contributions

M.B., T.J. and I.D. designed the experiment, M.B. and T.J. collected data, conducted analyses and drafted the manuscript, M.B. and O.B. analysed EEG data. All authors provided feedback and revised the manuscript.

Competing interests

The authors declare no competing financial interest.

Code availabaility

The code used for the analyses is available from the corresponding author upon request.

Data availability

Data from this study are available from the corresponding author upon request.

References

1. Kanwisher N (2000) Domain specificity in face perception. *Nat Neurosci*. doi:10.1038/77664.
2. Bruce V, Young A (1986) Understanding face recognition. *Br J Psychol*. doi:10.1111/j.2044-8295.1986.tb02199.x.
3. Scherer KR (2001) Appraisal considered as a process of multilevel sequential checking. *Appraisal Processes in Emotion: Theory, Methods, Research*, eds Scherer KR, Schorr A, Johnstone T, pp 92–120.
4. Sander D, Grafman J, Zalla T (2003) The human amygdala: an evolved system for relevance detection. *Rev Neurosci* 14(4):303–316.
5. Liddell BJ, et al. (2005) A direct brainstem-amygdala-cortical "alarm" system for subliminal signals of fear. *Neuroimage*. doi:10.1016/j.neuroimage.2004.08.016.
6. Whalen PJ, et al. (2013) Neuroscience and Facial Expressions of Emotion: The Role of Amygdala-Prefrontal Interactions. *Emot Rev* 5(1):78–83.
7. Vuilleumier P, Driver J (2007) Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. *Philos Trans R Soc London B Biol Sci* 362(1481).
8. Fusar-Poli P, Landi P, O'Connor C (2009) Neurophysiological response to emotional faces with increasing intensity of fear: A skin conductance response

study. *J Clin Neurosci* 16(7):981–982.

9. Vuilleumier P, Pourtois G (2007) Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia* 45(1):174–194.
10. Freeman JB, Stolier RM, Ingbretsen ZA, Hehman EA (2014) Amygdala Responsivity to High-Level Social Information from Unseen Faces. *J Neurosci*. doi:10.1523/JNEUROSCI.5063-13.2014.
11. Bar M, Neta M, Linz H (2006) Very first impressions. *Emotion*. doi:10.1037/1528-3542.6.2.269.
12. Willis J, Todorov A (2006) First impressions: Making up your mind after a 100-ms exposure to a face. *Psychol Sci*. doi:10.1111/j.1467-9280.2006.01750.x.
13. Cacioppo S, Bianchi-Demicheli F, Hatfield E, Rapson RL (2012) Social neuroscience of love. *Clin Neuropsychiatry* 9(1):3–13.
14. Zeki S (2007) The neurobiology of love. *FEBS Lett* 581(14):2575–2579.
15. Acevedo BP, Aron A, Fisher HE, Brown LL (2012) Neural correlates of long-term intense romantic love. *Soc Cogn Affect Neurosci*. doi:10.1093/scan/nsq092.
16. Ramon M, Gobbini MI (2018) Familiarity matters: A review on prioritized processing of personally familiar faces. *Vis cogn*. doi:10.1080/13506285.2017.1405134.
17. Guerra P, et al. (2012) Affective processing of loved familiar faces : Integrating central and peripheral electrophysiological measures. *Int J Psychophysiol* 85(1):79–87.
18. Bayer M, Ruthmann K, Schacht A (2017) The impact of personal relevance on emotion processing: Evidence from event-related potentials and pupillary responses. *Soc Cogn Affect Neurosci*. doi:10.1093/scan/nsx075.

19. Vida MD, Nestor A, Plaut DC, Behrmann M (2017) Spatiotemporal dynamics of similarity-based neural representations of facial identity. *Proc Natl Acad Sci*. doi:10.1073/pnas.1614763114.

20. Nemrodov D, Niemeier M, Mok JNY, Nestor A (2016) The time course of individual face recognition: A pattern analysis of ERP signals. *Neuroimage*. doi:10.1016/j.neuroimage.2016.03.006.

21. Young AW, Burton AM (2017) Recognizing Faces. *Curr Dir Psychol Sci*. doi:10.1177/0963721416688114.

22. Gobbini MI, Haxby J V (2007) Neural systems for recognition of familiar faces. 45(1):32–41.

23. Natu V, O'Toole AJ (2011) The neural processing of familiar and unfamiliar faces: A review and synopsis. *Br J Psychol*. doi:10.1111/j.2044-8295.2011.02053.x.

24. Haxby J V., Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4(6):223–233.

25. Visconti Di Oleggio Castello M, Halchenko YO, Guntupalli JS, Gors JD, Gobbini MI (2017) The neural representation of personally familiar and unfamiliar faces in the distributed system for face perception. *Sci Rep*. doi:10.1038/s41598-017-12559-1.

26. Sugiura M (2014) Neuroimaging studies on recognition of personally familiar people. *Front Biosci*. doi:10.2741/4235.

27. Kriegeskorte N (2008) Representational similarity analysis – connecting the branches of systems neuroscience. *Front Syst Neurosci*. doi:10.3389/neuro.06.004.2008.

28. Kriegeskorte N, Kievit RA (2013) Representational geometry: Integrating cognition, computation, and the brain. *Trends Cogn Sci*.

doi:10.1016/j.tics.2013.06.007.

29. Eimer M (2000) The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11(10):2319–24.

30. Pourtois G, Schettino A, Vuilleumier P (2013) Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biol Psychol* 92(3):492–512.

31. Rossi V, et al. (2017) Motivational Salience Modulates Early Visual Cortex Responses across Task Sets. *J Cogn Neurosci* 29(6):968–979.

32. Bayer M, Grass A, Schacht A (2018) Associated valence impacts early visual processing of letter strings: Evidence from ERPs in a cross-modal learning paradigm. *Cogn Affect Behav Neurosci*. doi:10.3758/s13415-018-00647-2.

33. Steinberg C, Bröckelmann AK, Rehbein M, Dobel C, Junghöfer M (2013) Rapid and highly resolving associative affective learning: Convergent electro- and magnetoencephalographic evidence from vision and audition. *Biol Psychol*. doi:10.1016/j.biopsycho.2012.02.009.

34. Berridge KC, Kringelbach ML (2015) Pleasure Systems in the Brain. *Neuron*. doi:10.1016/j.neuron.2015.02.018.

35. Northoff G, et al. (2006) Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *Neuroimage* 31(1):440–457.

36. Powell LJ, Kosakowski HL, Saxe R (2018) Social Origins of Cortical Face Areas. *Trends Cogn Sci*. doi:10.1016/j.tics.2018.06.009.

37. Patel GH, Sestieri C, Corbetta M (2019) The evolution of the temporoparietal junction and posterior superior temporal sulcus. *Cortex*. doi:10.1016/J.CORTEX.2019.01.026.

38. Leech R, Sharp DJ (2014) The role of the posterior cingulate cortex in cognition

and disease. *Brain* 137(1):12–32.

39. Hendrick C, Hendrick S (1986) A Theory and Method of Love. *J Pers Soc Psychol*. doi:10.1037/0022-3514.50.2.392.

40. Pascalis O, de Schonen S, Morton J, Deruelle C, Fabre-Grenet M (1995) Mother's face recognition by neonates: A replication and an extension. *Infant Behav Dev*. doi:10.1016/0163-6383(95)90009-8.

41. Philip RCM, et al. (2012) A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neurosci Biobehav Rev* 36(2):901–942.

42. Pierce K (2004) The brain response to personally familiar faces in autism: findings of fusiform activity and beyond. *Brain* 127(12):2703–2716.

43. Hatfield E, Sprecher S (1986) Measuring passionate love in intimate relationships. *J Adolesc* 9(4):383–410.

44. Rorden C (2011) fMRI Simulator. Available at: https://www.mccauslandcenter.sc.edu/crnl/tools/fmristim.

45. Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8(2–3):109–14.

46. Smith SM (2002) Fast robust automated brain extraction. *Hum Brain Mapp* 17(3):143–155.

47. Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17(2):825–41.

48. Smith SM, Brady JM (1997) SUSAN—A New Approach to Low Level Image Processing. *Int J Comput Vis*. doi:10.1023/A:1007963824710.

49. Pruim RHR, et al. (2015) ICA-AROMA: A robust ICA-based strategy for removing

motion artifacts from fMRI data. *Neuroimage*. doi:10.1016/j.neuroimage.2015.02.064.

50. Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. *Med Image Anal*. doi:10.1016/S1361-8415(01)00036-6.

51. Avants BB, Tustison N, Song G (2009) Advanced Normalization Tools (ANTS). *Insight J*. doi:http://hdl.handle.net/10380/3113.

52. Allen PJ, Josephs O, Turner R (2000) A Method for Removing Imaging Artifact from Continuous EEG Recorded during Functional MRI. *Neuroimage* 12(2):230–239.

53. Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. *Neural Comput* 7(6):1129–59.

54. Smittenaar P, Kurth-Nelson Z, Mohammadi S, Weiskopf N, Dolan RJ (2017) Local striatal reward signals can be predicted from corticostriatal connectivity. *Neuroimage*. doi:10.1016/j.neuroimage.2017.07.042.

55. Oosterhof NN, Connolly AC, Haxby J V. (2016) CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Front Neuroinform*. doi:10.3389/fninf.2016.00027.

56. Nili H, et al. (2014) A Toolbox for Representational Similarity Analysis. *PLoS Comput Biol*. doi:10.1371/journal.pcbi.1003553.