Getting to Know You: Emerging Neural Representations during Face Familiarization

[©]Géza Gergely Ambrus,¹ Charlotta Marina Eick,¹ Daniel Kaiser,² and Gyula Kovács¹

¹Institute of Psychology, Friedrich Schiller University Jena, Leutragraben 1, D-07743 Jena, Germany, and ²Department of Psychology, University of York, Heslington, York, YO10 5DD, United Kingdom

The successful recognition of familiar persons is critical for social interactions. Despite extensive research on the neural representations of familiar faces, we know little about how such representations unfold as someone becomes familiar. In three EEG experiments on human participants of both sexes, we elucidated how representations of face familiarity and identity emerge from different qualities of familiarization: brief perceptual exposure (Experiment 1), extensive media familiarization (Experiment 2), and real-life personal familiarization (Experiment 3). Time-resolved representational similarity analysis revealed that familiarization quality has a profound impact on representations of face familiarization. Across all experiments, we found no enhancement of face identity representation, suggesting that familiarity and identity representations emerge independently during face familiarization. Our results emphasize the importance of extensive, real-life familiarization for the emergence of robust face familiarity representations, constraining models of face perception and recognition memory.

Key words: EEG; familiarity; identification; multivariate pattern analysis; person recognition

Significance Statement

Despite extensive research on the neural representations of familiar faces, we know little about how such representations unfold as someone becomes familiar. To elucidate how face representations change as we get familiar with someone, we conducted three EEG experiments where we used brief perceptual exposure, extensive media familiarization, or real-life personal familiarization. Using multivariate representational similarity analysis, we demonstrate that the method of familiarization has a profound impact on face representations, and emphasize the importance of real-life familiarization. Additionally, familiarization shapes representations of face familiarity and identity differently: as we get to know someone, familiarity signals seem to appear before the formation of identity representations.

Introduction

Across our lifetimes, we continuously become familiar with new faces. Being able to recognize familiar people is critical for social interactions. Determining the facial identity of a person, however, is a computationally challenging endeavor. This is because every face shares the same basic features; thus, telling the faces of different people apart is difficult (for review, see Peterson and Rhodes, 2006). Further, the facial features of an individual show

The authors declare no competing financial interests.

Correspondence should be addressed to Géza Gergely Ambrus at geza.ambrus@uni-jena.de or Gyula Kovács at gyula.kovacs@uni-jena.de.

https://doi.org/10.1523/JNEUROSCI.2466-20.2021

Copyright © 2021 the authors

enormous variability across time because of illumination, viewpoint, facial expression, hairstyle, makeup, and age changes, thereby also making the task of "telling people together" very difficult (Jenkins et al., 2011; Andrews et al., 2015). Given these difficulties, it is not surprising that special brain mechanisms have developed for the identification of familiar faces (Haxby et al., 2000; Calder and Young, 2005; Gobbini and Haxby, 2006, 2007; Duchaine and Yovel, 2015; Ramon and Gobbini, 2018; Young and Burton, 2018; Rapcsak, 2019). Recent studies investigating the temporal dynamics of face processing suggested that both earlier, feedforward and later, feedback processes are important for identification (Visconti di Oleggio Castello and Gobbini, 2015; Dobs et al., 2019; Wiese et al., 2019). Cortical processing differs between different face identities within the first 200 ms, for both unfamiliar (Nemrodov et al., 2016, 2018; Vida et al., 2017) and familiar faces (Ambrus et al., 2019; Dobs et al., 2019). However, the earliest representations of facial identity seem to be explained by visual features shared between people, whereas later representations (from

Received Sep. 21, 2020; revised Feb. 22, 2021; accepted Apr. 5, 2021.

Author contributions: G.G.A. and G.K. designed research; G.G.A. and C.M.E. performed research; G.G.A., D.K., and G.K. analyzed data; G.G.A., C.M.E., D.K., and G.K. wrote the paper.

This work was supported by Deutsche Forschungsgemeinschaft Grants K03918/5-1 and KA4683/2-1. We thank Sophie-Marie Rostalski, Louisa Fortwengel, Morgana Dalla Palma, Alexia Dalski, and Bettina Kamchen for help in participant recruitment and data acquisition.

400 ms onwards) are related to the invariant coding of identity (Ambrus et al., 2019; Wiese et al., 2019).

Despite extensive research on how we represent familiar faces, we know much less about how the neural representation of a person unfolds as they become familiar.

Face familiarity can be divided into three distinct qualities (Natu and O'Toole, 2011): (1) During perceptual familiarization, a face is presented repeatedly without additional information. (2) During media (or other forms of parasocial) (Horton and Wohl, 1956) interactions, additional person-specific biographical semantic information is made available. One gradually gains knowledge about the person's characteristics, without meeting the person in real life. This level is typically studied experimentally using faces of famous people. (3) During personal familiarization, we encounter the person in real life and the biographical information gets detailed; personal episodic memories, as well as emotional knowledge, become available, too. Consequently, during the process of getting to know someone, a predominantly perceptual face representation gradually turns into representation of the person in declarative memory. Therefore, the identification of a person involves both the processing of perceptual information and the matching of this information to previously existing person knowledge stored in declarative recognition memory. To date, no study has compared how representations of face identity and familiarity emerge when faces are familiarized through perceptual, media, and personal familiarization. Therefore, we currently do not know what it takes to get to know a person to a level that reliably forges novel face representations in the brain.

Here, we elucidate how neural representations of familiarity and identity emerge after different qualities of familiarization. In three EEG experiments, participants viewed "ambient" images (Jenkins et al., 2011) of unfamiliar faces and faces familiarized by a brief perceptual familiarization (Experiment 1), an extensive media familiarization (Experiment 2), or a real-life, personal familiarization (Experiment 3). To track representations, we used time-resolved representational similarity analysis (RSA) (Kriegeskorte and Kievit, 2013; Cichy et al., 2014). We were particularly interested in the emerging familiarity representations, reflecting processing differences between familiarized and unfamiliar faces, and face identity representations, reflecting differences among familiarized identities, compared with differences among unknown identities.

Our study yielded three key insights into how face representations change as we get to know someone. First, our results reveal cortical representations of familiarity, emerging after \sim 400 ms over the right temporal cortex. Second, we find that the quality of familiarization has a profound impact on familiarity representations: they were strongly visible after personal familiarization (Experiment 3), weaker after media familiarization (Experiment 2), and completely absent after perceptual familiarization (Experiment 1). Third, we find that familiarity representations emerge independently of enhanced identity representations: Although significant face identity representations were found in all experiments, the degree of familiarization did not impinge on their quality. Together, these results emphasize the importance of long, real-life familiarization for the emergence of robust representations of face familiarity, constraining models of face perception and recognition memory.

Materials and Methods

Participants

Forty-two, 24, and 23 participants (8, 2, and 7 males; average \pm SD, age: 22.01 \pm 3.06, 22.81 \pm 2.01, and 23.65 \pm 4.07 years, respectively) took

part in the perceptual, media, and personal familiarization experiments, respectively. Sample sizes were similar to or exceeded those in previous EEG-multivariate pattern analysis studies on cortical face representations (Cauchoix et al., 2014; Nemrodov et al., 2018; Acunzo et al., 2019; Ambrus et al., 2019; Smith and Smith, 2019; Kaiser and Nyga, 2020). All participants were right-handed and had normal or corrected-to-normal visual acuity. Participants received partial course credits or monetary compensation. The experiments were conducted in accordance with the guidelines of the Declaration of Helsinki and with the approval of the ethics committee of the University of Jena. Written informed consent was acquired from all participants.

Stimuli and procedures for Experiment 1

For Experiment 1, the stimuli were ambient, color photographs of 4 Hungarian female celebrities (CA: Adél Csobot; NA: Anikó Nádai; SK: Kata Sarka; SE: Erika Szabó), unknown to our participants. The experiment involved a perceptual familiarization phase, a subsequent EEG recording phase, and a final face matching test phase. For each participant, we selected 2 of the 4 celebrities, balanced across participants. We familiarized participants with these two identities in a sequential sorting task (Ambrus et al., 2017). On each trial, participants were instructed to assign a photograph to one of the two identities via button press. Trials were separated by a 150 ms intertrial interval. No feedback was given to the participants. Overall, 120 sorting trials were recorded, with 30-30 images of the two to-be-familiarized identities repeated twice.

The subsequent EEG recording phase was executed immediately after the familiarization phase. During the EEG experiment, 1760 trials (including 160 target trials; see below) were presented. The 1600 nontarget trials, 1 of 40 face images (10 per identity) was presented for 600 ms (each image was repeated 40 times, amounting to 400 repetitions of each face identity). In target trials, images were rotated 10° clockwise or counterclockwise (10% of trials); participants were instructed to press a button in these trials. After the EEG recording, participants were asked to perform a face-matching task, in which they had to indicate whether two photographs were of the same or different identities.

Stimuli and procedures for Experiment 2

For Experiment 2, 20 images of the leading actors and actresses of the television series, "The Americans" (Keri Russel and Matthew Rhys) and "The Bridge" (Sofia Helin and Thure Lindhardt), were downloaded from the Internet. We selected highly variable images of these actors, depicting them during public appearances and avoided using snapshots from their films.

Experiment 2 involved a pre- and a post-familiarization EEG recording phase. The experimental design was identical to Experiment 1, except for the stimuli used. In each session, a total of 1066 trials (1040 nontargets and 26 targets) were presented. Here, in the 1040 nontarget trials, 1 of 80 face images (20 per identity) was presented for 600 ms (each image was repeated 13 times, amounting to 260 repetitions of each face identity). Different face images were used in the pre- and postfamiliarization sessions. For Experiment 2, the pre- and post-familiarization EEG recordings were separated by a media familiarization, during which participants watched one season of one of the series (allocated in a counterbalanced manner across participants). On average, participants required 2 weeks to complete the season. To facilitate the acquisition of person-specific information, we sent daily text messages to the mobile phone of our participants, including biographical trivia and personal details of the actors. The main actor and actress of this series served as familiarized face stimuli of the EEG recording sessions, while those of the other series served as unfamiliar faces for this experiment. Participants were unaware of the real aims of the experiment, and they were asked to keep the procedures of the experiment confidential for the duration of the entire study.

Stimuli and procedures for Experiment 3

For Experiment 3, 10 images of four young women (two student assistants in our laboratory and two similar-age friends of laboratory members) were collected with variable facial expressions, hairstyles, and lightning.

As in Experiment 2, participants completed a pre- and a post-familiarization EEG session. During these sessions, participants viewed images of the four identities. Experimental procedures were the same as in Experiment 2. In each session, a total of 968 trials (880 nontargets and 88 targets) were presented. Here, in the 880 nontarget trials, 1 of 40 face images (20 per identity) was presented for 600 ms (each image was repeated 22 times, amounting to 220 repetitions of each face identity). The pre-familiarization EEG recording took place on a Monday. On the following 3 d of the week, participants were asked to visit the laboratory and to have an \sim 1-h-long informal discussion and to play a quiz game with the two student assistants. The two assistants were instructed to discuss as many of their autobiographical details with the participants as possible but otherwise behave naturally. Participants were unaware of the real aims of the experiment and were told that the experiment tests social cognition. The post-familiarization EEG recording always took place on the Friday after the third familiarization session. The experiments were concluded with a short quiz, testing the person-related semantic knowledge of our participants, and a familiarity rating.

EEG preprocessing

All recording and analysis procedures were identical for the three experiments, unless otherwise stated. The EEG was recorded by a 64-channel Biosemi Active II system (512 Hz sampling rate; bandwidth: DC to 120 Hz) in a dimly lit, electrically shielded, and sound-attenuated chamber. The distance between the eyes and the computer screen was set to 96 cm via a chin rest. Electrooculogram was recorded from the outer canthi of the eyes and from above and below the left eye. EEG was notch-filtered at 50 Hz, bandpass-filtered between 0.1 and 40 Hz, segmented from -200 to 1300 ms relative to stimulus onset, and baseline-corrected with respect to the first 200 ms. The resulting data were downsampled to 100 Hz to increase the signal-to-noise ratio in the multivariate analyses (Grootswagers et al., 2017). No further processing of the data was performed, given that, compared with mass averaging done in univariate analysis, EEG classification techniques are quite robust to incidental noise (by means of suppressing noisy data during training) (Grootswagers et al., 2017). The preprocessing pipeline was implemented in MNE-Python (Gramfort et al., 2013, 2014).

RSA

Multivariate analysis methods were identical for all three experiments. To model the neural organization of face representations, we performed an RSA (Kriegeskorte, 2008) on the EEG data. The neural dissimilarity between all pairs of face images (i.e., between all individual images) was extracted by performing linear classification analysis, where pairwise decoding accuracies were used as a measure of representational dissimilarity. Linear discriminant analysis classifiers (implemented in MNI-Python 0.19) were trained and tested on response patterns across clusters of electrodes (see below), separately for each pair of images and for each time point of the epoch. Training and testing were done in a leave-oneout scheme: classifiers were trained on all but one trial for each of the two conditions, and tested on the remaining trials. This procedure was repeated until each trial was left out once, and classification accuracy was averaged across these repetitions. Classification time courses were then arranged into representational dissimilarity matrices (RDMs; with empty diagonal), separately for each time point. As we were mainly interested in the gradual buildup of neural representations of face familiarity and identity, we averaged the neural RDMs in discrete bins of 50 ms. This averaging further increased the signal-to-noise-ratio of our data while retaining sufficient temporal precision.

To also provide an approximation of the origin of neural representations, we constructed RDMs separately for six electrode clusters across the scalp (as in Ambrus et al., 2019). Linear discriminant analysis classifiers were trained and tested on response patterns recorded in six electrode clusters over right and left posterior (13 electrodes each), central (12 electrodes each), and anterior cortex (12 electrodes each) (see Fig. 1*C*). All subsequent analyses were performed separately for each of these clusters.

To model the neural dissimilarity, we created two categorical predictor RDMs. Each predictor RDM covered 40×40 (Experiment 1 and 3) or 80 × 80 (Experiment 2) elements and contained zeros where the entries represented comparisons of similar images (i.e., similar on the dimension of interest, see below) and contained ones where the entries reflected comparisons of dissimilar images. To quantify correspondence between the predictor RDMs and the neural RDMs, we unfolded the lower off-diagonal elements of the matrices into two vectors (i.e., the diagonal of both matrices was discarded) and correlated the vectors using Spearman's correlation coefficients. These correlations were computed separately for each time point, leading to a time series of correlations that reflected the correspondence of the neural data and the predictor. Individual-participant correlations were Fishertransformed. In order to control for low-level stimulus properties, we constructed pixel-dissimilarity matrices (Cichy et al., 2017; Ambrus et al., 2019) and partialed out their effects in all subsequent analyses.

Modeling familiarity information. For assessing the effect of face familiarity (i.e., the differences between familiarized and unfamiliar faces), all comparisons within the same familiarization condition (i.e., both familiarized or both unfamiliar faces) were marked as similar (0), and all comparisons between the different familiarization conditions (i.e., one familiarized and one unfamiliar face) were marked as dissimilar (1). To not confound familiarity information with identity information, all comparisons within the same identity (i.e., two different images of the same person) were excluded from this analysis (see Fig. 1D).

Modeling identity information. For assessing the effect of face identity (i.e., the differences between different people), all comparisons within the same identity (i.e., both images showing the same person) were marked as similar (0), and all comparisons between two different identities (i.e., two images showing different people) were marked as dissimilar (1). Such identity predictor RDMs were constructed separately for the two familiarized and for the two unfamiliar faces (see Fig. 1D), allowing us to independently track identity information for faces that have been familiarized and faces that are unknown.

Statistical testing

To identify significant effects across time, we used cluster permutation tests (as implemented in MNE-Python), with 10,000 iterations. Statistical testing was one-sided (p < 0.05) with the assumption that correlations between neural RDMs and model RDMs were significantly larger than zero.

In order to compare familiarity information across the different modes of familiarization (see Fig. 1), we performed an *F* test-based cluster permutation test (10,000 permutations, two-tailed, p < 0.05) on the familiarity model correlations of the three experiments. As a *post hoc* test, in time intervals flagged as significant, model correlation values were averaged and compared across experiments using two-tailed independent-samples *t* tests (p < 0.05)

Data availability

The conditions of our ethics approval do not permit public archiving of the study data. The data will be made available to interested researchers following completion of a data sharing agreement and approval by the local ethics committee. Readers seeking access to the data should contact the corresponding author.

Results

To track the emergence of neural representations of face familiarity and face identity, we conducted three complementary EEG experiments. We particularly focused on what it takes to forge such representations for people who were initially completely unknown. We therefore gradually increased the quality of familiarization across our experiments, starting from perceptual familiarization (Experiment 1) to extensive media familiarization using a TV series (Experiment 2) and real-life personal familiarization (Experiment 3).

A Paradigm



B MVPA Analysis





Figure 1. Experimental paradigm and EEG analysis methods. *A*, All experiments covered faces of four identities: two identities that were familiarized during the experiment (F1, F2) and two identities that remained unfamiliar to participants (U1, U2). Each experiment comprised EEG recording sessions (pre- and/or post-familiarization) during which participants viewed a diverse set of face photographs of the familiarized und unfamiliar identities. The three experiments differed in the quality of face familiarization: Experiment 1 (E1) involved brief perceptual familiarization in an identity sorting task. Experiment 2 (E2) involved media familiarization with the main actors of one of two TV series ("The Americans," "The Bridge") as well as the delivery of biographical information about these actors. Experiment 3 (E3) involved a 3-d-long personal familiarization with two student assistants. After each experiment, familiarity was assessed in a behavioral test. *B*, Multivariate pattern analysis. EEGs were segmented between -200 and 1300 ms relative to stimulus onset. For each time point separately, linear dassification analyses on EEG response patterns across electrodes of a duster were performed for each combination of individual images, using a leave-one-trial-out scheme. This procedure resulted in time-resolved RDMs, which reflected pairwise dissimilarity for each combination of faces at each time point across the epoch. *C*, Separate neural RDMs were created based on response patterns from six different electrode dusters over the right and left posterior, central, and anterior cortex. Central electrodes were added to both right and left dusters. *D*, To track representations of face familiarity, we correlated neural RDMs with a familiarity predictor RDM (example shows the RDM of Experiment 1), in which faces of the same familiarity were coded as similar. To track representations of face identity, we correlated neural RDMs with identity predictor RDMs separately for the familiarized and unfamiliarized neural RD



Figure 2. Demonstration of the between-identity similarity of the identities and the within-identity variability of the stimulus sample in the three experiments. Top, Landmark-based, within-identity morphs. Middle, Pixelwise averages of the images for each identity, separately. Bottom, Variance across the images within one identity, with warmer colors representing higher pixelwise variance. These average images were not part of the actual experiment and merely illustrate the stimuli and their properties.

Experiment 1: perceptual familiarization

In Experiment 1, we probed emerging face representations using a simple, perceptual familiarization task. Before the EEG recording session, participants (n = 42) completed a sequential facesorting task with 30 ambient photographs of 2 previously unknown women (Fig. 1A), which captured extensive withinperson variability (Fig. 2A). Participants were instructed to sort photographs of 2 similarly aged unfamiliar women, presented sequentially and in a randomized order, into two identities, using the left and right buttons of a computer keyboard (for details, see Materials and Methods). Such tasks have been successfully used to facilitate invariant face recognition in previous behavioral studies (Jenkins et al., 2011; Burton, 2013). A face-matching test, conducted after the experiment (Andrews et al., 2015; Ambrus et al., 2017), confirmed that the perceptual familiarization indeed facilitated image-independent face recognition: Participants were more accurate in discriminating the identities of the familiarized faces than the ones of similar, but unknown faces (Fig. 3A).

After the perceptual familiarization, participants completed an EEG testing session, during which they viewed 10 images of the 2 familiarized persons and 10 images of 2 visually similar but unknown persons (Fig. 1*A*). During this session, participants performed an orthogonal task (see Materials and Methods). Images that were presented during the familiarization phase were not repeated during the EEG session. To track the cortical representations of the faces across time, we used an RSA framework whereby the organization of cortical representations is captured by the pairwise (dis)similarity in responses to the different stimuli, manifest in neural RDMs. Such neural RDMs were constructed from time-resolved decoding analyses (Fig. 1B). For each time point across the EEG epochs (200 ms pre-stimulus until 1300 ms post-stimulus), we performed decoding analyses between each possible pair of faces. In these analyses, we trained and tested classifiers on response patterns across EEG electrodes from six different topographic clusters, which spanned the right and left parietal-occipital, central-temporal, and frontal cortex (Fig. 1*C*). Classification accuracy between two faces was taken as a measure of their neural discriminability (with higher classification accuracy indexing greater neural dissimilarity). Aggregating these pairwise classification accuracies yielded a neural RDM (40-by-40 entries) for each time point across the epochs (100 Hz resolution).

By modeling the neural RDMs with different predictor RDMs (Fig. 1*D*), we were able to track cortical representations of face familiarity and identity.

To track representations of face familiarity, neural RDMs at each time point were correlated with a predictor RDM in which entries were coded as similar when both faces were familiarized or when both faces were unknown, and as dissimilar when one face was familiarized, and the other face was unknown. To not confound familiarity with identity, all comparisons between



Figure 3. Behavioral results. Face matching performance for the perceptually familiarized (F1, F2) faces exceeded that of the unfamiliar (U1, U2) or novel (Unseen) faces in Experiment 1 (A). This suggests that the sorting task led to the learning of the identities involved (Ambrus et al., 2017). Average familiarity estimations after media (B) and personal (D) familiarization for the familiarized identities are larger compared with those of the unfamiliar faces. Media (C) and personal (E) familiarization also reduced the estimated number of presented identities between pre- and post-familiarization sessions (note that the correct answer is 4 in both cases). This suggests that the personal familiarization procedures facilitated the creation of a stable identity representation. Error bars indicate SE. *p < 0.05. ***p < 0.001. n.s., not significant.

images of the same identity were excluded from the familiarity predictor RDM.

When correlating the neural data with the familiarity predictor RDM, we found that, except for a very brief interval between 600 and 800 ms in the right posterior ROI, familiarity did not significantly predict neural responses in any of the electrode clusters (Fig. 4*A*). This result suggests that a short perceptual familiarization is insufficient for shaping neural representations of facial familiarity. This finding is surprising, as perceptual familiarization tasks are used in a variety of studies as a tool for familiarizing participants with different face identities (Dubois et al., 1999; Leveroni et al., 2000; Gobbini and Haxby, 2007; Kriegeskorte et al., 2007; Nestor et al., 2011; Anzellotti et al., 2014), and our participants showed behavioral signs of incidental learning of face identity (Fig. 3*A*).

To track representations of face identity, neural RDMs were correlated with two predictor RDMs at each time point: (1) a predictor RDM capturing identity for the familiarized faces; in this RDM, entries were coded as similar when the two faces stemmed from the same familiarized identity, and as dissimilar when both faces stemmed from different familiarized identities; and (2) a predictor RDM capturing identity for the unfamiliar faces; in this RDM, entries were coded as similar when the two faces stemmed from the same unfamiliar identity, and as dissimilar when both faces stemmed from different unfamiliar identities.

These analyses revealed representations of face identity for both the unfamiliar and the familiarized faces. For the unfamiliar faces, face identity was represented from \sim 600 ms over all electrodes, with late anterior clusters bilaterally (Fig. 5*B*). For the familiarized faces, face identity was represented 250-1300 ms over bilateral central and posterior sites (Fig. 5*A*).

Next, we directly compared identity information in the EEG signals for the unfamiliar and the familiarized faces. Representations of face identity that emerge during familiarization training should be indexed by higher correlations between the familiarized identity predictor RDM and the neural RDMs, compared with the correlations between the unknown identity predictor RDM and the neural RDMs; this is because familiarization should selectively enhance the differentiation between the familiarized identities, but not between the unknown identities.

However, we did not find any significant differences between identity representations for the unfamiliar and the familiarized faces. This result suggests that perceptual familiarization does not lead to a measurable increase in the quality of existing identity representations in the brain. The identity representations uncovered here are independent of learning and may rather be related to visual feature differences that allow for an incidental representation of facial identity, even if the person is effectively unknown.

Experiment 2: media familiarization

As perceptual familiarization neither led to strong representations of face familiarity, nor to emerging representations of familiarized face identity, we sought to increase the quality of face familiarization. In Experiment 2, we used a media familiarization, in which we familiarized participants (n = 24) with the two leading actors appearing in a TV series. We selected two series on Netflix ("The Americans" and "The Bridge") which are relatively unknown to our potential pool of participants. Both series had a female and a male protagonist with approximately equal screen time. Next, we allocated one of the series to each participant for the familiarization phase. The series allocation was counterbalanced across participants. The participants were instructed to watch an entire season of their allocated series at home, within 14 d. Additionally, to ensure the best possible familiarization, a daily SMS was sent to their mobile phone about important biographical details of the main actors. The faces of the main actor and actress of this series served as familiarized faces, whereas those of the other, unseen series served as unfamiliar faces. EEG recordings took place in two sessions, before and after the familiarization phase (Fig. 1A). During the EEG sessions, previously unseen, ambient images of the two main actors of the familiarized series were intermixed with images of the two main actors of the other, unseen series. Participants again performed an orthogonal task (see Materials and Methods). Familiarity ratings collected after the post-familiarization EEG session confirmed that participants felt familiar with the two actors they had seen in the series (Fig. 3B,C).

We analyzed the data in the same RSA framework as in Experiment 1, tracking neural representations of face familiarity and identity. However, the inclusion of a pre- and post-familiarization EEG recordings allowed us to compare the EEG



Experiment 1: Perceptual familiarization Experiment 2: Media familiarization

Figure 4. Neural representations of face familiarity across the three experiments. Each panel represents the correlation between the time-resolved neural RDMs and the familiarity predictor RDM in all electrodes and each of the six electrode clusters with image pixel dissimilarity partialed out. A, Perceptual familiarization in Experiment 1 did not lead to robust neural signatures of face familiarity. B, Media familiarization in Experiment 2 led to a measurable representation of face familiarity in bilateral electrode clusters over the temporal cortex: after familiarization, familiarized, and unfamiliar faces were better discriminable than they were before familiarization. C, A similar pattern of results emerged in Experiment 3. After personal familiarization, we found even stronger signatures of face familiarity that spanned electrode clusters over the bilateral occipital and temporal cortex. Together, this suggests that representations of face familiarity emerge in the occipitotemporal cortex at ~400 ms of processing. Critically, our results indicate that such representations strongly depend on the quality of familiarization. Horizontal lines indicate statistical significance (p < 0.05, corrected for multiple comparisons across time). Shaded ranges represent SEM.

responses to images of the very same person, either when the person remained completely unknown to the participant or was extensively seen in the TV series.

We did not expect to find any familiarity information before the familiarization, as all four actors (and their film characters) were unfamiliar to our participants at this stage. As expected, evidence for representations of face familiarity was largely absent in the EEG signals (Fig. 4B). Only in a short time window and only over all electrodes (350-550 ms; peak: 450 ms), we found a difference between the unfamiliar and the to-be-familiarized faces, suggesting that this weak initial representation of familiarity may have been caused by incidental visual similarities among images of actors that stage in the same series. After familiarization, we found a markedly different pattern of results emerging in the electrode clusters spanning the temporal cortices (Fig. 2B). In the post-familiarization EEG, neural representations of face familiarity were evident over all electrodes between 200 and 650 ms (peak: 600 ms), two short time periods (200-300 and 1150-1250 ms) in the left central cluster and from 200 ms to 700 ms (peak at 600 ms) and between 850 and 1050 ms (peak at 950 ms) in the right central cluster. When directly comparing EEG recordings before and after familiarization, face familiarity predicted neural responses more strongly post- compared with prefamiliarization in the right central cluster between 300 ms and 400 ms (peak: 300 ms), between 450 ms and 600 ms (peak: 600 ms), and between 850 ms and 1150 ms (peak: 950 ms). Together, these results suggest that media familiarization produces reliable representations of face familiarity. These representations emerge most prominently in EEG signals recorded over the temporal cortex and after 400 ms of processing. This spatiotemporal profile suggests that these representations are related to a convergence of perceptual face information and memory representations formed during familiarization.

Similar to Experiment 1, the representation of face identity was comparable before and after media familiarization (Fig. 5C, D). The strongest identity information was found in the posterior electrode clusters, starting from 100 ms post-stimulus, with multiple subsequent peaks (~150 ms as well as between 300 and 650 ms). Interestingly, the spatiotemporal distribution of the identity and familiarity representations effects are markedly different. The strongest familiarity effects emerged in electrode clusters over the right temporal cortex, whereas the strongest identity effects emerged in electrode clusters over the bilateral occipital cortex. The occipital source of identity representations suggests that these representations are indeed mostly related to visual differences between identities, rather than genuine representations stored at the interface of perception and memory. In line with this interpretation, we found no differences in identity representations for the familiarized faces when comparing EEG recordings before and after the familiarization.

Experiment 3: personal familiarization

Unlike perceptual familiarization in Experiment 1, the media familiarization in Experiment 2 led to reliable representations of face familiarity. However, we were still unable to uncover an enhanced representation of face identity resulting from familiarization. We therefore increased the quality of face familiarization even further. In Experiment 3, we familiarized our participants



Figure 5. Neural representations of face identity across the three experiments. Each panel represents the correlation between the time-resolved neural RDMs and the identity predictor RDMs for the familiarized (A, C, E) or the unfamiliar faces (B, D, F) in each of the six electrode clusters, with image pixel dissimilarities partialed out. Across all the experiments, we found that EEG signals across the whole scalp, but most prominently over occipital cortex, allowed for discriminating both familiarized and unfamiliar face identities. Familiarization did not improve the quality of face identity representation in any of the experiments: identity information was comparable between familiarized and unfamiliar faces and between the pre- and post-familiarization EEG recordings. The similarity in identity discrimination between familiarized and unfamiliar faces suggests that these effects are driven by visual features correlated with face identity rather than by a genuine representation of familiar face identity. This suggests that strong neural signals of face familiarity can be observed in the absence of measurable representations of familiar face identity. Horizontal lines indicate statistical significance (p < 0.05, corrected for multiple comparisons across time). Shaded ranges represent SEM.

(n = 23) with people they had never met before via real-life, personal interactions (Fig. 1A). During the pre-familiarization EEG session, participants viewed ambient images of four unfamiliar women (Fig. 2), which included two student assistants of the laboratory and two randomly selected, unknown women of similar age and appearance, while performing an orthogonal task. In the following 3 d, participants were personally familiarized with the two student assistants. They were asked to visit the laboratory and have an \sim 1-h-long informal discussion with them while playing a quiz game ("Black Stories") (Berger, 2014). During the post-familiarization EEG session, participants again viewed images of the four identities they had seen during the pre-familiarization session, including the two women they now had previously met personally. Familiarity ratings collected after the experiment confirmed that participants felt familiar with the two student assistants they had met in person (Fig. $3D_{2}E$).

We analyzed the data in the same RSA framework as used in Experiment 2, again tracking neural representations of face familiarity and identity.

In accordance with the fact that all four persons were originally unfamiliar to our participants, no familiarity information could be detected before the personal familiarization phase (Fig. 4C). Personal familiarization, however, changed this picture drastically. After familiarization, strong and long-lasting familiarity information was found over the bilateral central and posterior electrode clusters (Fig. 4C). Over all electrodes, we observed a strong familiarity effect between 450 and 850 ms (peak at 600 ms). In the right central cluster, face familiarity more strongly predicted neural responses after familiarization between 550 and 600 ms (peak: 600 ms), whereas in the posterior clusters, familiarization enhanced representation of face familiarity between 550-1000 and 400-1150 ms (peaks at 550 ms) for the left and right hemispheres, respectively. These results strongly corroborate the results from Experiment 2, showing that experimental familiarization can produce robust representations of face familiarity. These representations are found over the occipital and temporal cortices, and most strongly at \sim 400 ms after stimulus onset. Together, our results show a striking dependence of these representations on the quality of familiarization: neural signatures of face familiarity are very strong for people we got to know in real life, considerably weaker for people we have encountered in the media, and undetectable for people whose pictures we have only seen during perceptual tasks.

In Experiment 3, representations of facial identity were found in the bilateral posterior and central electrode clusters (Fig. 5*E*, *F*). The strongest correlations with the identity predictor RDMs were found after 200 ms over bilateral posterior and right central and anterior electrode clusters. However, similar to the results of Experiment 2, the identity representation of familiarized faces did not change significantly when comparing EEG recordings before and after the familiarization. This result demonstrates that strong neural representations of acquired face familiarity can emerge in the absence of detectable representation of familiar face identity.

In order to compare familiarity information across the different familiarization methods (Fig. 1*A*), an *F* test-based cluster permutation test was performed on the Familiarity model correlations in the three experiments (Fig. 6). This analysis shows that, at bilateral posterior sites, as well as overall the entire scalp, significant clusters emerged in the 450-1000 ms time window. At these clusters flagged as significant, pairwise, independent-sample *t* tests were applied to compare the averaged model correlation values between experiments. These comparisons revealed a significant difference between both Media (p = 0.001) and Personal (p < 0.0001) compared with Perceptual familiarization in the 450-700 ms interval, and between Media (p = 0.004) and Personal (p < 0.007) familiarization at 850-1000 ms for the entire scalp. In the bilateral posterior ROIs, both Perceptual and Media familiarizations differed from Personal familiarization (between 450-750 ms on the right and 500-1000 ms on the left, all pvalues ≤ 0.003). Together, these results suggest that the neural representation of face familiarity gradually increases with familiarization quality; and under the strongest familiarization, direct personal contact, representations of face familiarity are found most strongly in electrodes over the lateral occipital cortex.

Discussion

When we see a person's face, we often immediately know whether we have or have not seen them previously, leading to a feeling of familiarity. At the same time, we may also recognize who they are (i.e., identify the person), provided we know them. However, face familiarity and identity do not always go hand in hand in our brains. We show that neural representations of face familiarity emerge after 400 ms over the right temporal cortex, and they are strongly dependent on the quality of the familiarization. They are mostly absent for people encountered during brief perceptual familiarization, stronger for people that are familiarized through extensive media exposure, and most robustly visible when we personally get to know someone. These representations of face familiarity emerge independently of identity representations for familiar faces: In our experiments, familiarization did not enhance the quality of the identity representations.

As the key result, our study revealed robust representations of face familiarity that emerge during face familiarization, particularly after prolonged and deep familiarization through extensive media exposure and personal real-life contact. These emerging representations of face familiarity may originate from two complementary processing stages. On one hand, they may reflect changes in perceptual processing, with face-selective brain mechanisms becoming more responsive to familiar faces, compared with faces that are unknown to us. This interpretation is in line with previous fMRI studies showing enhanced activation in faceselective regions of the occipitotemporal cortex for familiar, compared with unfamiliar, faces (e.g., Gobbini et al., 2004; Ramon et al., 2015; Weibert and Andrews, 2015; for review, see Natu and O'Toole, 2011). On the other hand, representations of face familiarity may reflect a convergence between perceptual information and information stored in declarative recognition memory. This interpretation is consistent with studies of face memory, which have reported enhanced BOLD activations for highly familiar, compared with less familiar faces, in the prefrontal, parietal, and medial temporal cortices (Leveroni et al., 2000; Kosaka et al., 2003; Leube et al., 2003; Gobbini and Haxby, 2006; Bobes et al., 2013; Silson et al., 2019)

Pinpointing neural representations of familiarity to the interface of face perception and recognition memory is consistent with the timing of the effects in our study: Representations of face familiarity emerged only after 400 ms of processing, much later than perceptual responses to faces (Eimer et al., 2011). This timing is consistent with a previous ERP study that reported differences between familiar and unfamiliar faces in averaged waveforms after 400 ms (Wiese et al., 2019). Their relatively late timing suggests that representations of face familiarity emerge from post-perceptual interactions that are governed by feedback from memory systems. But which memory processes underlie



Figure 6. Direct comparison of the neural representations of face familiarity across the three experiments for the familiarized faces in the post-familiarization phase. Plots represent the correlation between the time-resolved neural RDMs and the familiarity predictor RDM in all electrodes and each of the six ROIs with image pixel dissimilarity partialed out. Representations of face familiarity differed significantly across experiments. For all electrodes, Media and Personal familiarization led to significantly greater familiarity information than Perceptual familiarization. Personal familiarization particularly led to stronger familiarity information in posterior electrodes, compared with the other types of familiarization methods. Black horizontal lines indicate statistical significance (p < 0.05; *F* test-based cluster permutation test). Top, Color lines indicate significant differences across conditions (p < 0.05; independent-samples *t* test on the mean model correlations in the significant cluster intervals). Shaded ranges represent SEM.

these interactions? Recognition memory is often conceived as consisting of two functionally and anatomically different components: familiarity and recollection (Yonelinas, 2002; Rugg and Yonelinas, 2003; Renoult et al., 2019). Familiarity serves with the vague feeling of "knowing" that the person has already been encountered, and it is assumed to depend largely on anterior temporal and perirhinal cortical functions. Recollection, on the other hand, serves person recognition by providing episodic details and person-associated information and is assumed to depend on medial temporal lobe structures, such as the hippocampus, as well as on prefrontal and parietal areas (for review, see Brown and Banks, 2015). Our results support such a dissociation between familiarity and recollection: Although we found strong correlates of face familiarity, we did not find any evidence for recollection in the form of identity-specific representations. Observing representations of face familiarity in the absence of clear identity representations provides a novel neural correlate for the well-known psychological phenomenon of feeling familiar with a person without being able to correctly identify them (Mandler, 1980).

Contrary to the strong representations of face familiarity, we did not find representations of face identity that arise from familiarization with a person. Even when the faces were familiarized through real-life exposure, we found no stronger neural representation of face identity for familiar than for unfamiliar faces. This finding is surprising, given that previous fMRI (Kriegeskorte et al., 2007; Axelrod and Yovel, 2015; Jeong and Xu, 2016; Visconti Di Oleggio Castello et al., 2017) and MEG/ EEG (Ambrus et al., 2019; Dobs et al., 2019) studies have identified robust neural signatures of familiar face identity. There are multiple reasons why such representations were not found here. First, representations of face identity may take longer to consolidate, and the applied familiarization times, specifically in the personal familiarization experiment, may not be sufficient to shape robust identity signals; in real life, we typically meet people often and with longer timespans between encounters. Second, after such relatively brief familiarization periods, face identity signals may not be strong enough or emerge on a too fine spatial scale to be readily decodable from scalp EEG recordings. Finally, emerging identity representations may be differently visible in neural recordings when participants are prompted to identity people during the experiments, rather than just passively viewing the images.

However, our data show that faces of different individuals are generally discriminable from EEG data. Across all experiments, EEG signals carried information about face identity, starting within the first 200 ms after stimulus onset. Such early identity signals are consistent with previous work on reconstructing unfamiliar face identity from EEG signals (Vida et al., 2017; Nemrodov et al., 2018; but see Dobs et al., 2019) for finding no traces of unfamiliar identity representations. Our results expand these findings by suggesting that the signal carries sufficient information for identity discrimination across several and highly variable natural images. However, these identity signals were already observed before familiarization and were highly similar for the unfamiliar and familiarized faces. It thus seems that the identity representations observed in the current study are related to visual feature differences between faces of different identities that allow for an incidental representation of facial identity, even if the person is effectively unknown. Although these representations are markedly different from identity representations in a genuine sense, they may serve as a stepping stone toward the acquisition of genuine representations of face identity (Johnston and Edmonds, 2009). This idea is consistent with recent computational work (Blauch et al., 2021): When training a deep neural network to discriminate face identity, representations in the early layers of the network (akin to the early neural representations of visual features) stay largely unchanged, and only later network representations (akin to more advanced representations of face identity in the brain) change. At this point, more work is needed to uncover where and when such genuine identity representations are formed in the human brain as we become familiar with a person.

Together, our findings provide new insights into how brain representations change as we get to know new people. They highlight that the most prominent signs of face familiarization are emerging representations of face familiarity that strongly depend on the quality of familiarization. Curiously, we show that neural representations of familiarity emerge in the absence of measurable identity information, providing a novel explanation for why in real-life situations we sometimes have a strong feeling of knowing a person, without being able to tell who they are. Better understanding the diverging trajectories of familiarity and identity representations as we get to know new people has the potential to inform more accurate models of face perception as well as recognition memory.

References

- Acunzo D, MacKenzie G, van Rossum MC (2019) Spatial attention affects the early processing of neutral versus fearful faces when they are taskirrelevant: a classifier study of the EEG C1 component. Cogn Affect Behav Neurosci 19:123–137.
- Ambrus GG, Windel F, Burton AM, Kovács G (2017) Causal evidence of the involvement of the right occipital face area in face-identity acquisition. Neuroimage 148:212–218.
- Ambrus GG, Kaiser D, Cichy RM, Kovács G (2019) The neural dynamics of familiar face recognition. Cereb Cortex 29:4775–4784.

- Andrews S, Jenkins R, Cursiter H, Burton AM (2015) Telling faces together: learning new faces through exposure to multiple instances. Q J Exp Psychol 68:2041–2050.
- Anzellotti S, Fairhall SL, Caramazza A (2014) Decoding representations of face identity that are tolerant to rotation. Cereb Cortex 24:1988–1995.
- Axelrod V, Yovel G (2015) Successful decoding of famous faces in the fusiform face area. PLoS One 10:e0117126.
- Berger N (2014) Black stories: Medizin edition, Vol 3. Kempen, Germany: Moses Verlag.
- Blauch NM, Behrmann M, Plaut DC (2021) Computational insights into human perceptual expertise for familiar and unfamiliar face recognition. Cognition 208:104341.
- Bobes MA, Lage Castellanos A, Quiñones I, García L, Valdes-Sosa M (2013) Timing and tuning for familiarity of cortical responses to faces. PLoS One 8:e76100.
- Brown MW, Banks PJ (2015) In search of a recognition memory engram. Neurosci Biobehav Rev 50:12–28.
- Burton AM (2013) Why has research in face recognition progressed so slowly? The importance of variability. Q J Exp Psychol 66:1467–1485.
- Calder AJ, Young AW (2005) Understanding the recognition of facial identity and facial expression. Nat Rev Neurosci 6:641–651.
- Cauchoix M, Barragan-Jason G, Serre T, Barbeau EJ (2014) The neural dynamics of face detection in the wild revealed by MVPA. J Neurosci 34:846–854.
- Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time. Nat Neurosci 17:455–462.
- Cichy RM, Kriegeskorte N, Jozwik KM, Bosch JJ, van den Charest I (2017) Neural dynamics of real-world object vision that guide behaviour. bioRxiv 147298. doi.org/10.1101/147298.
- Dobs K, Isik L, Pantazis D, Kanwisher N (2019) How face perception unfolds over time. Nat Commun 10:1258.
- Dubois S, Rossion B, Schlitz C, Bodart JM, Michel C, Bruyer R, Crommelinck M (1999) Effect of familiarity on the processing of human faces. Neuroimage 9:278–289.
- Duchaine B, Yovel G (2015) A revised neural framework for face processing. Annu Rev Vis Sci 1:393–416.
- Eimer M, Gosling A, Nicholas S, Kiss M (2011) The N170 component and its links to configural face processing: a rapid neural adaptation study. Brain Res 1376:76–87.
- Gobbini MI, Haxby JV (2006) Neural response to the visual familiarity of faces. Brain Res Bull 71:76–82.
- Gobbini MI, Haxby JV (2007) Neural systems for recognition of familiar faces. Neuropsychologia 45:32–41.
- Gobbini MI, Leibenluft E, Santiago N, Haxby JV (2004) Social and emotional attachment in the neural representation of faces. Neuroimage 22:1628– 1635.
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M (2013) MEG and EEG data analysis with MNE-Python. Front Neurosci 7:267.
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen MS (2014) MNE software for processing MEG and EEG data. Neuroimage 86:446–460.
- Grootswagers T, Wardle SG, Carlson TA (2017) Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data. J Cogn Neurosci 29:677– 697.
- Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233.
- Horton D, Wohl RR (1956) Mass communication and para-social interaction. Psychiatry 19:215–229.
- Jenkins R, White D, Van Montfort X, Burton AM (2011) Variability in photos of the same face. Cognition 121:313–323.
- Jeong SK, Xu Y (2016) Behaviorally relevant abstract object identity representation in the human parietal cortex. J Neurosci 36:1607–1619.
- Johnston RA, Edmonds AJ (2009) Familiar and unfamiliar face recognition: a review. Memory 17:577–596.
- Kaiser D, Nyga K (2020) Tracking cortical representations of facial attractiveness using time-resolved representational similarity analysis. Sci Rep 10:168521.
- Kosaka H, Omori M, Iidaka T, Murata T, Shimoyama T, Okada T, Sadato N, Yonekura Y, Wada Y (2003) Neural substrates participating in acquisition of facial familiarity: an fMRI study. Neuroimage 20:1734–1742.

- Kriegeskorte N (2008) Representational similarity analysis: connecting the branches of systems neuroscience. Front Syst Neurosci 2:4.
- Kriegeskorte N, Kievit RA (2013) Representational geometry: integrating cognition, computation, and the brain. Trends Cogn Sci 17:401–412.
- Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc Natl Acad Sci USA 104:20600–20605.
- Leube DT, Erb M, Grodd W, Bartels M, Kircher TT (2003) Successful episodic memory retrieval of newly learned faces activates a left fronto-parietal network. Cogn Brain Res 18:97–101.
- Leveroni CL, Seidenberg M, Mayer AR, Mead LA, Binder JR, Rao SM (2000) Neural systems underlying the recognition of familiar and newly learned faces. J Neurosci 20:878–886.
- Mandler G (1980) Recognizing: the judgment of previous occurrence. Psychol Rev 87:252–271.
- Natu V, O'Toole AJ (2011) The neural processing of familiar and unfamiliar faces: a review and synopsis. New York: Wiley/Blackwell.
- Nemrodov D, Niemeier M, Mok JN, Nestor A (2016) The time course of individual face recognition: a pattern analysis of ERP signals. Neuroimage 132:469–476.
- Nemrodov D, Niemeier M, Patel A, Nestor A (2018) The neural dynamics of facial identity processing: insights from EEG-based pattern analysis and image reconstruction. eNeuro 5:ENEURO.0358-17.2018.
- Nestor A, Plaut DC, Behrmann M (2011) Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. Proc Natl Acad Sci USA 108:9998–10003.
- Peterson MA, Rhodes G (2006) Perception of faces, objects, and scenes: Analytic and holistic processes. New York: Oxford University Press.
- Ramon M, Gobbini MI (2018) Familiarity matters: a review on prioritized processing of personally familiar faces. Vis Cogn 26:179–195.
- Ramon M, Vizioli L, Liu-Shuang J, Rossion B (2015) Neural microgenesis of personally familiar face recognition. Proc Natl Acad Sci USA 112:E4835– E4844.

Rapcsak SZ (2019) Face recognition. Curr Neurol Neurosci Rep 19:41.

- Renoult L, Irish M, Moscovitch M, Rugg MD (2019) From knowing to remembering: the semantic–episodic distinction. Trends Cogn Sci 23:1041–1057.
- Rugg MD, Yonelinas AP (2003) Human recognition memory: a cognitive neuroscience perspective. Trends Cogn Sci 7:313–319.
- Silson EH, Steel A, Kidder A, Gilmore AW, Baker CI (2019) Distinct subdivisions of human medial parietal cortex support recollection of people and places. Elife 8:e47391.
- Smith FW, Smith ML (2019) Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks. Neuroimage 195:261–271.
- Vida MD, Nestor A, Plaut DC, Behrmann M (2017) Spatiotemporal dynamics of similarity-based neural representations of facial identity. Proc Natl Acad Sci USA 114:388–393.
- Visconti di Oleggio Castello M, Gobbini MI (2015) Familiar face detection in 180ms. PLoS One 10:e0136548.
- 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 7:12237.
- Weibert K, Andrews TJ (2015) Activity in the right fusiform face area predicts the behavioural advantage for the perception of familiar faces. Neuropsychologia 75:588–596.
- Wiese H, Tüttenberg SC, Ingram BT, Chan CY, Gurbuz Z, Burton AM, Young AW (2019) A robust neural index of high face familiarity. Psychol Sci 30:261–272.
- Yonelinas AP (2002) The nature of recollection and familiarity: a review of 30 years of research. J Mem Lang 46:441–517.
- Young AW, Burton AM (2018) Are we face experts? Trends Cogn Sci 22:100–110.