The neural code behind real-world recognition abilities

Simon Faghel-Soubeyrand[1,2], Meike Ramon[3], Eva Bamps [5], Matteo Zoia [6], Jessica Woodhams [2], Anne-Raphaelle Richoz[4], Roberto Caldara[4], Frédéric Gosselin*[1] & Ian Charest*[1,2]

1. Département de psychologie, Université de Montréal, Canada
2. School of Psychology, University of Birmingham, UK
3. Applied Face Cognition Lab, Institute of Psychology, University of Lausanne, Switzerland
4. Department of Psychology, University de Fribourg, Switzerland
5. Center for Contextual Psychiatry, Department of Neurosciences, KU Leuven, Belgium
6. Department for Biomedical Research, University of Bern, Switzerland

Corresponding authors:
Ian Charest,
i.charest@bham.ac.uk
Simon Faghel-Soubeyrand,
simon.faghel-soubeyrand@umontreal.ca

Word counts. Abstract: 92. Main text : 2292, Method : 4627

Summary

We recorded a large dataset of high-density electroencephalographic signals and used a combination of behavioural tests and machine learning to characterise the brain computations covarying with face recognition in individuals with extraordinary abilities. We show that individual face recognition ability can be accurately decoded from brain activity in an extended temporal interval for face and non-face objects. We demonstrate that this decoding is supported by perceptual and semantic brain computations.

Main

The ability to robustly recognise the faces of our colleagues, friends and family members is paramount to our success as social beings. Our brains complete this feat with apparent ease and speed, in a series of computations unfolding within tens of milliseconds in a wide brain network comprising the inferior occipital gyrus (OFA), the fusiform gyrus (FFA), the superior temporal sulcus (pSTS), and more anterior areas such as the anterior temporal lobe (ATL) [1–3]. Not all individuals, however, are equally competent at recognising faces in their surroundings. Developmental prosopagnosics show a great difficulty at this task despite an absence of brain injury [4]. In contrast, so-called “super-recognisers" exhibit remarkable face identification abilities, even after little exposure to a face several years in the past [5–7]. The specific nature of the neural processes responsible for these individual differences remains largely unknown. So far, studies have focused on univariate face-specific aspects of brain processing, such as the preferential responses to faces indexed electrophysiologically by the N170 component or by functional Magnetic Resonance Imaging (fMRI) localisers [8–14]. The distributed functional brain areas revealed by multivariate techniques suggests that the ability to recognise faces in the population likely involves brain areas beyond those previously revealed with univariate techniques [15]. Moreover, variations in ability across the population might also result from brain computations that do *not* differ between faces and objects [16]. In fact, characterising the computations implemented by the brain [17–19] is a challenging task, and this level of description is largely absent from previous literature exploring face recognition abilities. To tackle this challenge, here, we characterised the functional differences in brain activity predictive of face recognition ability with unprecedented levels of description, using a data-driven approach combining state-of-the-art computational tools and multivariate description of high-density electrophysiological (EEG) brain activity measured in a rare cohort of extraordinary face recognisers.

We recruited 16 super-recognisers, individuals better than the 98th percentile on a battery of face recognition tests [6] **Fig.1a**), and 17 typical-recognisers. We measured their brain potentials with high-density EEG while they performed a simple one-back task. The objects depicted in the stimuli belonged to multiple visual categories including face images of different sexes,

emotions, and identities, as well as images of man-made and non-face natural objects (e.g., a computer, a plant), animals (e.g., a giraffe, a monkey), and scenes (e.g., a city, a dining room) (**Fig. 1b**). Participants completed a total of more than 100,000 trials.

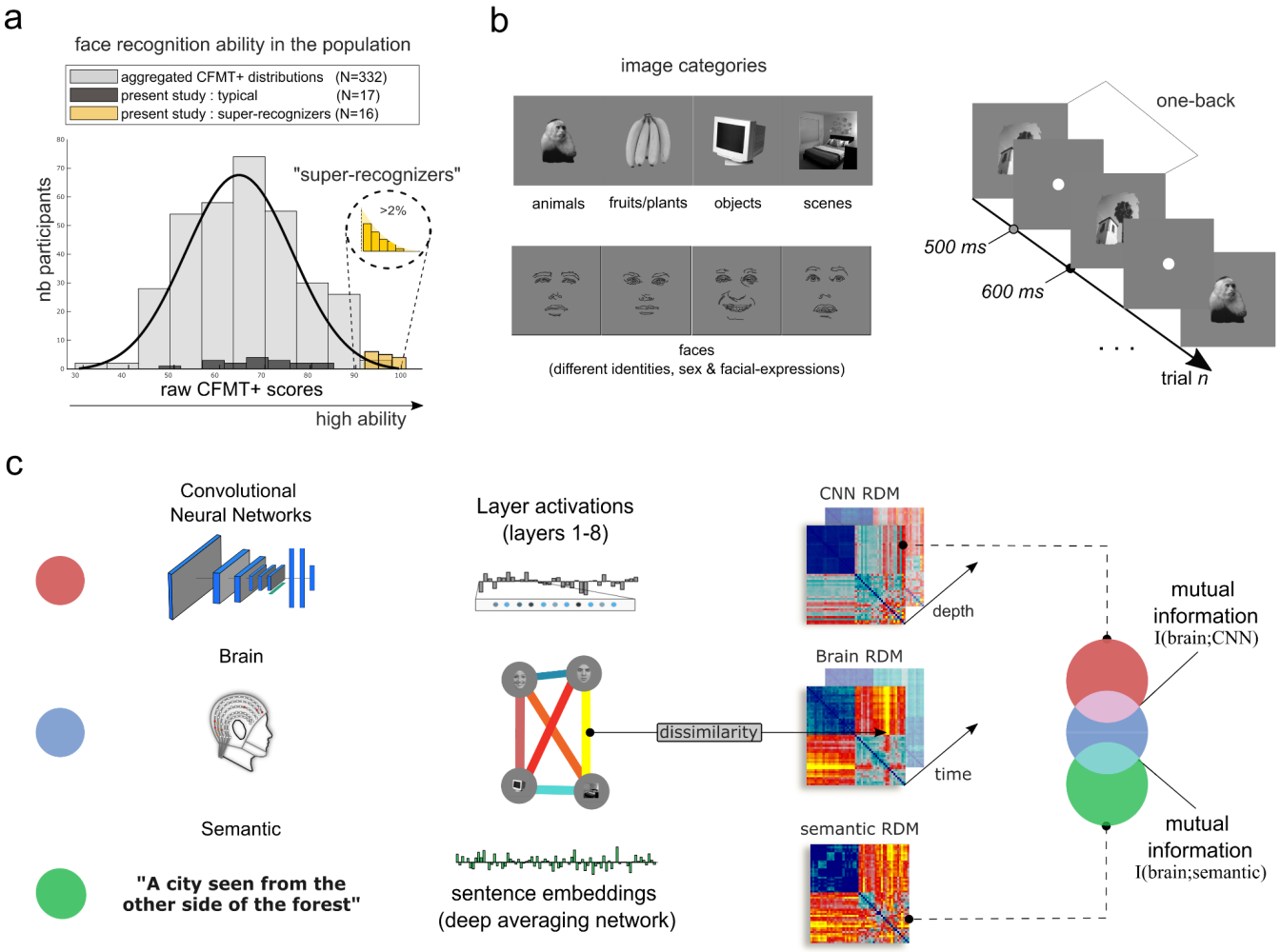


Figure 1. Experimental procedure, multivariate analysis and computational modelling of brain representations. a) The super-recogniser cohort consisted of 16 individuals in the top 1% of face recognition ability spectrum[6] (see online methods). An additional group of 17 neurotypical observers was recruited for this study (black bars). The Cambridge Face Memory Test long-form (CFMT+ [20]) scores of all these participants are depicted with those of 332 neurotypical observers from three independent studies [21–23]. **b)** Participants engaged in a one-back task while their brain activity was being recorded with high-density electroencephalography (EEG, 128 electrodes). The objects depicted in the stimuli belonged to various categories, such as faces, objects, and scenes. Note that drawings of faces are depicted here as an anonymised substitute to the experimental face stimuli presented to our participants. **c)** Representational dissimilarity matrices (RDM) were computed from convolutional neural networks (CNN) of vision[24,25], human brain activity, and a deep neural network of sentence semantics [26]. To characterise the CNN RDMs, we computed the pairwise similarity between unit activation patterns for each image independently in each CNN layer. Semantic RDMs were derived from human caption descriptions of the images transformed into semantic sentence embeddings. Brain RDMs were computed using cross-validated decoding performance between the EEG topographies from each pair

of stimuli at every 4 ms time-point. Mutual information [27] between the model RDMs and the brain RDMs was assessed, for every participant, at each 4 ms step from stimulus onset.

With this sizable and category-rich dataset, we first attempted to classify a participant as either a super- or a typical-recogniser based solely on their brain activity. More specifically, we trained Fisher linear discriminants to predict group-membership from single, 1-second trials of EEG patterns (in a moving searchlight of 5 neighbouring electrodes; see **Supplementary Fig.1** for a schematic of decoding analyses). We observed up to ~80% cross-validated decoding accuracies, peaking over right-lateralised electrodes. This is all the more impressive given that the noise ceiling imposed on our classification by the test-retest reliability of the CFMT+ [20], the gold-standard test we used to identify super-recognisers individuals, is ~93% correct (SD=2.28%; see online methods). Thus, the brains of individuals differing in their face recognition ability show important functional differences. We applied the same decoding procedure to each 4-ms interval of EEG recordings to reveal the time course of these functional differences. Group-membership predictions attained statistical significance (p<.001, permutation tests, **Fig. 2a**) from about 65 ms to at least 1 s after stimulus onset, peaking around 135 ms within the N170 window [28,29]. Notably, similar results were obtained following the presentation of both face *and* non-face visual stimuli (**Fig. 2a**; see also Supplementary Materials, *Behavioural results*), corroborating a critical prediction of domain-general accounts of face recognition [30–34].

A debate in individual differences research is whether the observed effects emerge from qualitative or quantitative changes in the supporting brain mechanisms [35,36]. The decoding results presented up to this point might give the impression that face recognition ability is supported by qualitatively different processes along the brain processing stream. However, these results were obtained with dichotomous classification models applied, by design, to the brains of individuals from a bimodal distribution of ability scores. To better assess the nature of the relationship between brain processing and ability in the general population, we thus performed a new decoding analysis on the typical-recognisers only, using a continuous regression model. Specifically, we used cross-validated fractional ridge regression [37] to predict *individual* CFMT+ face recognition ability scores from single-trial EEG data (see online methods). Overall, we replicated the previous decoding results: performance was above statistical threshold (p <.01, FDR-corrected) from about 80 ms to at least 1 s, peaking around 135 ms following stimulus onset, for both face and non-face stimuli (**Fig. 2b,** peak-r_{face} = .4149 at 133 ms, peak-$r_{non\text{-}face}$ = .4899 at 141 ms). This decoding of individual scores from EEG patterns suggests that face recognition ability is compatible with quantitative variations in brain mechanisms across individuals of the general population [21]. Altogether, our decoding analyses provide evidence for important, quantitative, temporally extended, and domain-general variations in the brain activity supporting face recognition abilities.

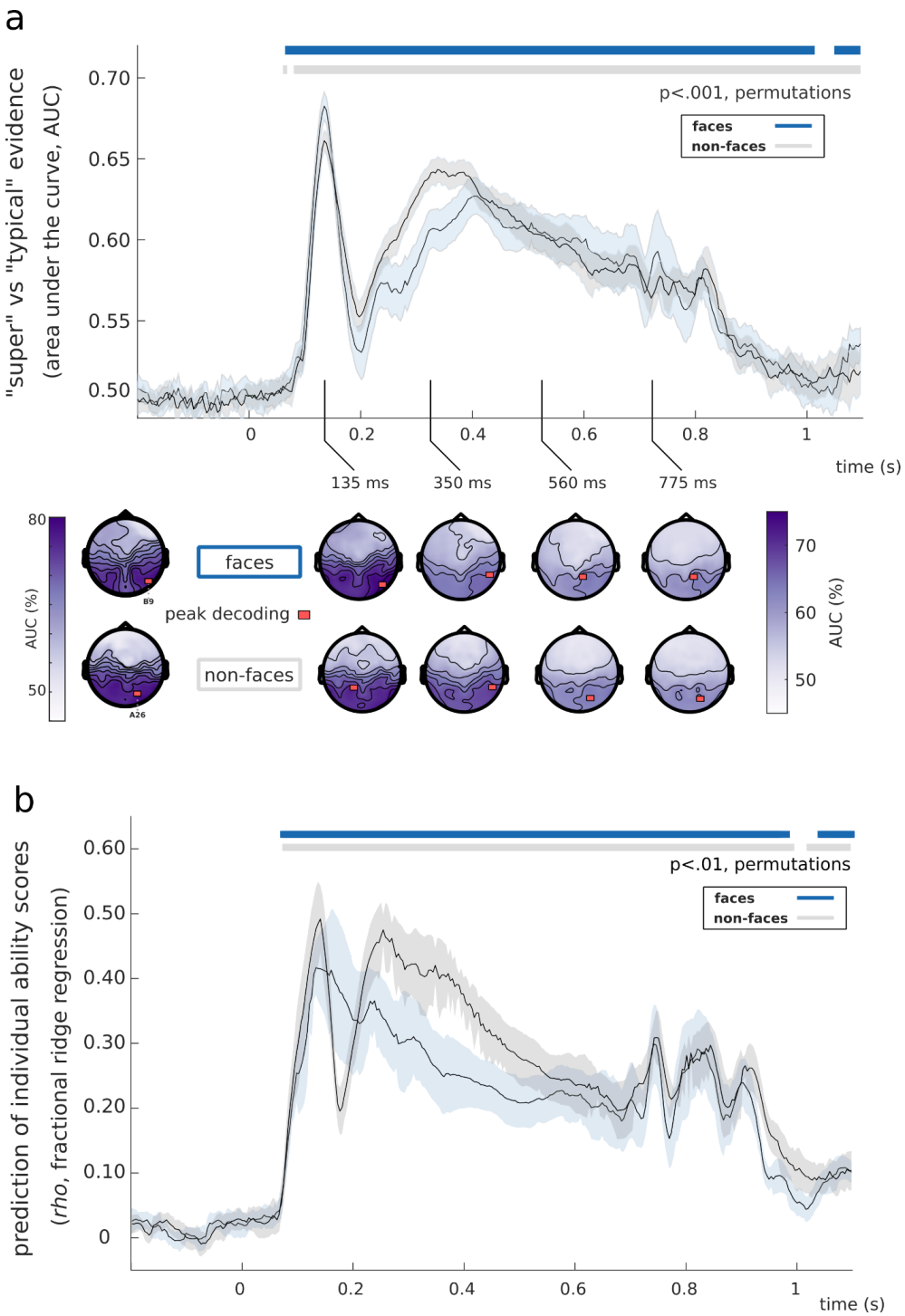


Figure 2. Decoding interindividual recognition ability variations from high-density EEG activity. **a**) Trial-by-trial group-membership predictions ("super" or "typical") were computed from high-density EEG patterns, for each 4-ms interval, while participants processed face (blue trace) or non-face stimuli (grey trace). Significant decoding performance occurred as early as 65 ms, peaked in the N170 window, and lasted up to at least 1 s after stimulus onset (p<.001). The topographies show results from a searchlight decoding analysis with classification performance attaining 78% accuracy around 135 ms over occipito-temporal electrodes. **b**) We decoded the typical-recognisers individual CFMT+ scores using a fractional ridge regression classifier. This yielded similar results with significant decoding as

early as 75 ms, peaking around the N170 time window (peak-r_{face} = .4149, peak-$r_{non\text{-}face}$ = .4899; p <.01, 1K permutations, 10 repetitions, FDR-corrected).

Temporal evidence, and domain-generality, however, are only partial indications of the level of brain computation [38,39]. To better characterise the specific visual brain computations covarying with face recognition ability, we compared, using representational similarity analysis [40–43], the brain representations of our participants to that of convolutional neural networks trained to categorise objects [24,25]. These CNNs process visual features of gradually higher complexity and abstraction along their layers [44], from low-level features (i.e., orientation, edges), to mid-level features (e.g., combinations of edges, contour, shape, texture) and high-level features (e.g., objects and object parts). The brain representations were characterised by computing representational dissimilarity matrices (RDMs) for every participant and for each 4-ms time interval. Brain RDMs were computed using the decoding performance of a linear discriminant model, where brain activity was decoded for every pair of stimulus conditions at a given time [45,46] see **Supplementary Fig. 2** for the group-average RDMs and time course of key categorical distinctions). The model representations were characterised by computing RDMs from the layers of the CNNs. Compared to typical participants, we found that the brain RDMs of super-recognisers showed larger mutual information [27] with the mid-layer RDMs of CNNs (e.g., layer 3, 4, 5 and 6 of AlexNet) between 133 and 165 ms (**Fig. 3a**, p<.05, cluster-test, see also **Supplementary Fig.2** for replication with VGG16). This indicates that the best face recognisers of the population have higher mid-level visual processing in the N170 window.

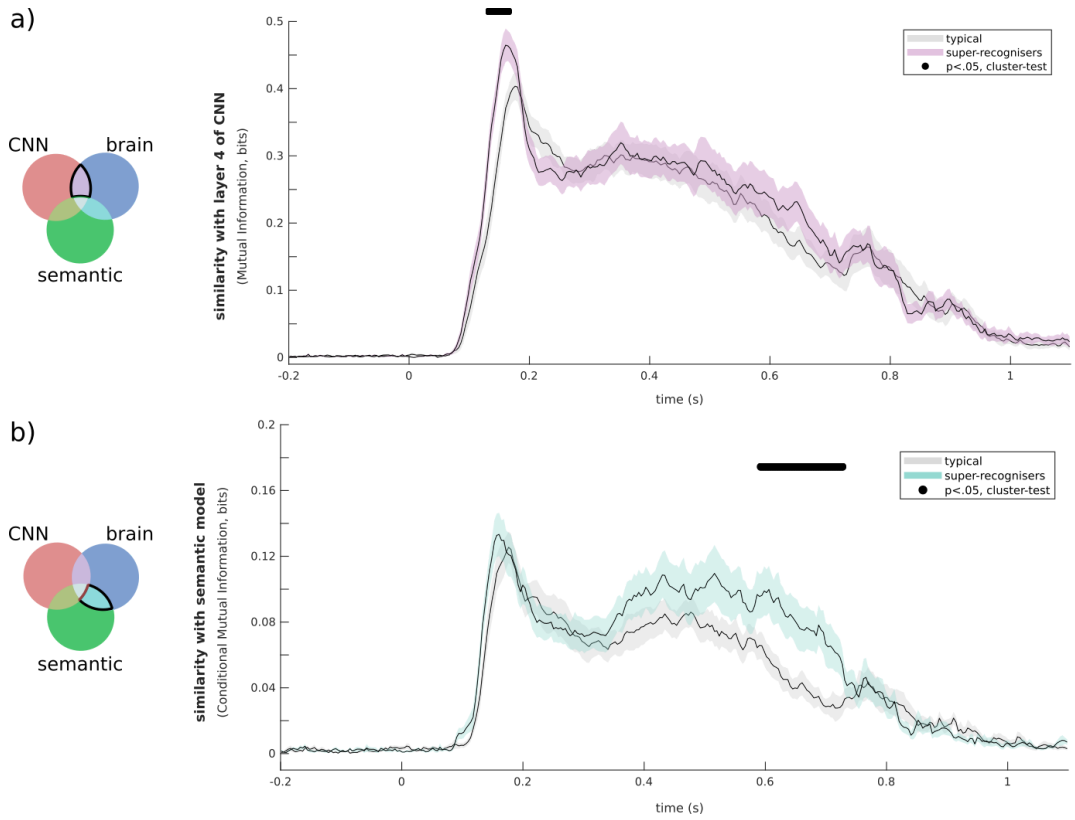


Figure 3. Comparison of super- and typical-recogniser brain representations with those of artificial neural networks of vision and sentence semantics. a) Mutual information between brain RDMs and AlexNet RDMs is shown for typical- (grey curve) and super-recognisers (pink curve). We found greater similarity with mid-level visual computations (layer 4 shown) in the brains of super-recognisers (black line indicates significant contrasts, p<.05, cluster-corrected) between 133 ms to 165 ms. **b**) Similarity with the semantic model, excluding shared mutual information between brain and AlexNet, differed for typical- and super-recognisers in a later time window centred around 650 ms (cyan curve; super > typical, p < .05, cluster-corrected). The shaded areas of all curves represent the standard error intervals of bootstrapped leave-one-out distributions.

The finding that face recognition ability could be decoded even 1 s after stimulus onset hints that brain computations beyond what is typically construed as pure visual processing also differ as a function of face recognition ability. To test this hypothesis, we asked five new participants to write captions describing the images presented during our experiment (e.g., “A city seen from the other side of a forest”), and used a deep averaging network (Google Universal Sentence Encoder, GUSE [26]) to transform these captions into embeddings (points in a semantic space). GUSE has been trained to predict semantic textual similarity from human judgments, and its embeddings have been shown to generalise to an array of other semantic judgement tasks [26]. We then compared the RDMs computed from these semantic embeddings to the brain RDMs of both typical- and super-recognisers. Importantly, this comparison excluded the information shared between the semantic and visual deep neural network models (see *Conditional Mutual Information* section; see also **Supplementary Fig.2** for additional

analyses on brain and DCNN mutual information excluding semantic information, which showed similar results as those in **Fig.3a**). We found larger mutual information with these caption-level computations in the brains of super- than in those of typical-recognisers between 598 and 727 ms (**Fig. 3b**, $p<.05$, cluster-test). This late time window contrasts with that uncovered with the visual computational model, observed much earlier within the N170. To our knowledge, this is the first time that processes beyond perception are associated with face recognition ability.

Discussion

Using a combination of behavioural tests and computational modelling of brain activity, we show that face recognition ability covaries with a cascade of brain computations that extend not only beyond traditional category-specific neural markers, but also beyond purely visual processing itself. We achieved robust decoding of face recognition ability when examining electrophysiological responses to face *and* non-face stimuli, echoing pioneering MVPA studies having shown widely shared representations of face and objects in the brain, including in face-sensitive areas [15]. The decoding accuracies we reached when predicting variations in such a high-level perceptual ability from single 1-s EEG trials were impressive, showing promising research paths using a combination of EEG and MVPA to study individual differences in high-level vision [47]. Furthermore, by using Representational Similarity Analysis (RSA) to combine human electrophysiological data with deep learning models of vision and semantics, we were able to reveal the temporal and computational mechanisms linked with the ability to recognise faces, from visual to late semantic information processing in the brain. Doing so, we demonstrate for the first time a clear link between semantic computations and real-world perceptual ability in the population.

The specificity of face recognition remains highly debated [30–32]. Arguments on both sides arose with the first wave of brain imaging studies [3,28,33,48,49] and early neural models of face perception [50,51]. It has since then been largely informed by neuropsychological findings from individuals clinically and selectively impaired in face (i.e., either with brain-lesions or with congenital/developmental deficits, [52–56]) and object recognition [57]. On top of precisely describing the temporal aspects of brain variations behind face recognition ability in the *general* population, our decoding approach constitutes a straightforward test for the contribution of category-specific vs. domain-general processes in individuals with a higher ability to recognise faces. Our successful decoding of individual ability from the brain activity of typical- and super-recognisers engaging with non-face objects (e.g., cars, animals, natural scenes) suggests that the mechanisms underlying highly proficient ability give rise to enhanced neural representations that are not restricted to faces. Our findings thus indicate that variations in face-recognition ability reflect, *at least,* a set of features/processes common to face and non-face objects [16,50,58,59]. To be clear, we do not claim that face-specific brain processes are not contributing to face recognition ability in the population. Rather, our results

suggest that face-specific mechanisms alone do not fully account for face recognition ability variations. Given the limited spatial resolution of EEG, future decoding studies with high temporal and spatial resolution imaging [60] will be needed to assess the contribution of face-specific and domain-general visual processes.

We found that the N170 brain processing window not only encompassed the peak of the decoded evidence for individual ability, but that it also bounded the differences between typical and outstanding recognisers outlined by our modelling of brain representations with artificial models of vision. This approach specifically uncovered mid-level representations of these hierarchical models, which have been previously linked to processing in IT [44,61] as well as to mid and high-level feature representations such as combinations of edges and parts of objects [44,62], as the level of computations predicting recognition ability in the N170 window. Of course, having associated mid and high-level visual representations to the N170 does not mean that brain processing in this window is exclusively represented by computations at this level of abstraction. While we believe that further specifying the level and nature of processing in this important component offers promising avenues of inquiry [63], the findings presented here should be taken only as indications that this level of visual abstraction in the N170 is predictive of our individual ability to recognise faces.

Face recognition ability, however, was found to be associated with processes that extended beyond what is traditionally thought of as purely visual [17], i.e., extending those generally credited to ”core” face perception regions such as the OFA, FFA or pSTS [3,64,65]. To date, surprisingly few studies have explored how face recognition might be expressed across more diverse areas of the cortex (e.g., ATL and Medial Temporal Lobe [15,50,66–69]) in addition to traditional face-sensitive regions (e.g., OFA, FFA, pSTS). To our knowledge, none have explored the link between semantic brain representations and high-level recognition processes. Yet, recent studies using computational techniques have shown that semantic processes account for an important part of the representations in brain areas of both the perceptual and memory stream [18,70–72]. Here, we expanded our understanding of the nature of brain representations supporting real-world recognition ability. Our computational approach uncovered that higher face recognition ability not only enhances visual representations found around the face-sensitive N170, which is anatomically credited to the FFA [73], but that it also supports richer high-level semantic representations. Individuals within the top 1% of face cognition ability displayed brain representations that were more similar to an artificial model of semantics [26,70,71] in a late window around the P600 [74–76]. These findings fit well with recent modelling of human intracranial data showing that semantic representations occurring late after stimulus onset are associated with better long-term memory performance [71]. More generally, while the aforementioned computational studies have predicted brain activity with semantic models, our results are the first to highlight the importance of semantic representations as a support for crucial, real-world aspects of recognition behaviour in the general population. The association between semantic processes and recognition ability had

been suggested in previous models of face recognition [1,50,51,77], but to our knowledge it had never been shown empirically until now.

To conclude, our findings provide compelling evidence highlighting the dynamic nature of brain computations supporting real-world recognition ability. Our data-driven approach combining real-world behavioural tests and computational modelling of large-scale EEG recordings not only precisely characterised the extended temporal dynamics of brain representations related to individual face recognition ability variations, but also specified the nature of brain computations which are characteristic of the best face-recognisers of the general population. Doing so, we provide evidence for domain-general and systematic variations in brain representations across typical and outstanding face recognisers, and reveal a wide array of processes—extending to concept-level representations—contributing to this critical ability. We strongly believe that similar intersectional approaches will further contribute to our understanding of the brain-mind-behaviour link; revealing insights about how crucial human behaviour is embodied in the brain and how important cognitive mechanisms, such as perceptual and semantic processes, support it. From a practical point of view, creating even more robust decoding models, perhaps using brain activity from more realistic (e.g., dynamic) visual stimulation, could have broad applications. These applications could provide quick and accurate alternatives to the traditional standardised behavioural tests assessing face recognition ability, such as in the context of security settings that benefit from strong face processing skills among their personnel (e.g., police agencies, border patrol). Further applications could even include the construction of training components [78,79] designed to improve face recognition for individuals impaired in face recognition.

Methods

Participants

A total of 33 participants were recruited for this study. The first group consisted of 16 individuals with exceptional ability in face recognition — super-recognisers (see also the *Description of the super-recogniser cohort* section). The second group was composed of 17 neurotypical controls (20-37 years old, 11 female). These sample sizes were chosen according to the effect sizes described on previous multivariate object recognition studies [45,46,80] and exceeded those typically seen in studies on the behavioural correlates of superior face-recognition ability [6,7,21,78,81,82]. One participant (SR-1) from the super-recogniser group was rejected due to faulty EEG recording. This study was approved by the Ethics and Research Committee of the University of Birmingham, and informed (written) consent was obtained from all participants.

Description of the super-recogniser cohort

Super-recognisers represent a small part of the population considered as having extraordinary ability to recognise faces; such individuals are thus, by definition, difficult to recruit. The 16

individuals tested in the current study (30-44 years old, 10 female) were previously known super-recognisers. They self-reported having superior face recognition ability in their everyday life and this assessment was confirmed by extensive behavioural testing.

Eight of the 16 (SR-9 to SR-15) were identified according to the procedure described previously[7] using three challenging tests of face cognition, including two matching (i.e., perceptual) and one memory (i.e., recognition) task. The remaining eight completed an online test battery of 6 face cognition tasks hosted by Josh Davis at the University of Greenwich (unpublished). Common to these evaluations was Cambridge Face Memory Test long-form (CFMT+ [6]), which is the gold-standard in the field. All 16 super-recognisers scored at least 92 on the CFMT+, with an average of 95.31 (SD = 2.68). A score greater than 90 (i.e., 2 SD above the average score) is considered the norm to classify an individual as a super-recogniser [6,82,83] see also Ramon, 2021; Bobak et al., 2016). The minimal score of 92 obtained in the current sample corresponds to the 99th percentile of scores in face-recognition ability distribution, as indicated by CFMT+ scores of 332 typical participants from three independent studies [21–23] (M = 66.1702, SD = 11.9032, Median = 66, c.f. [6,83,84]). Both cohorts' behavioural test scores are provided in **Supplementary Table 1 and 2.**

Description of typical recognisers cohort

The 17 individuals referred to as ''typical recognisers'' tested in the current study (20-37 years old, 11 female) were recruited on the campuses of University of Fribourg and Birmingham University. None had a history of psychiatric diagnostic or neurological disorder, and all had normal or corrected to normal vision. Their CFMT+ scores ranged from 50 to 85 (mean and STD of 70.00 and 9.08), and fitted well the general spectrum ability scores (see **Fig.1a**; they did not differ from the N=332 population of ability scores (t(346) = 1.3065, *p*=0.1922).

Procedure

The Cambridge Face Memory Test +

All participants were administered a gold-standard face identification ability test, the CFMT long-form, or CFMT + [6]. In the CFMT (short-form) [20], participants are required to memorise a series of face identities, and to subsequently identify the newly learned faces amongst three faces. The long-form of CFMT (CFMT+) was created to increase the difficulty to better assess individual performance thresholds and includes an additional 50 trials, which are more difficult than those contained in the CMFT short form. As such, the CFMT+ includes a total of 102 trials of increasing difficulty.The last ones involve recognizing faces following viewpoint changes, lighting variations and the addition of visual noise. The duration of this test is about 15 minutes. EEG was not recorded while participants completed this test.

One-back task

Stimuli. During the one-back task, we presented participants with 49 images of faces, animals, plants, objects, and scenes. Face stimuli consisted of 24 face images from the Radboud Face dataset [85] varying in identity (8 identities), sex (8 male), and facial expressions (8 neutral faces,

8 expressing joy, 8 expressing fear). These faces were converted to 250x250 greyscale images. Each of these face images was aligned based on twenty landmarks (averaged to six mean coordinates for left and right eyes, left and right eyebrows, nose, and mouth) using Procrustes transformations (rotation, scaling and translation). Each face image was revealed through an ellipsoid mask that excluded non-facial cues such as hair. The non-face images consisted of 25 inanimate objects (e.g., a car, a computer screen, a flower, bananas), scenes (e.g., city landscape, a kitchen, a bedroom) and animals (e.g., a giraffe, a monkey, a puppy) that were sampled from the stimuli set of Kiani et al. (2007). These images were resized to match approximately the size of the face images, and converted to 250x250 greyscale images. Finally, the luminance profile of the resulting images was equalised across all images using the SHINE toolbox [86].

Procedure. During the one-back task, we presented a succession of images to participants and asked them whether the current image was identical to the previous one (i.e., the one-back, see **Figure 1b**). This task was employed for two reasons. First, it ensured that participants focused their attention on the stimulus presented while their brain activity was being recorded. Second, this task involved processes similar to face or object recognition; our participants compared their representation of the presented image to another stored in working memory (i.e., the one-back image). A trial unravelled as follows: a white fixation point was presented on a homogeneously grey background (500 ms ± 50 ms jitter); a stimulus was randomly chosen among the 49 stimuli and presented on a homogeneously grey background (600 ms); the image disappeared, and a white fixation point appeared for another 500 ms, and so on. Participants had a maximum of 1100 ms to respond by pressing on the appropriate keyboard key after stimulus onset. This interval, as well as 200 ms before it, constituted the epoch of our EEG analyses. Feedback about the accuracy of responses was provided on a trial-by-trial basis. There was a probability of 0.1 that a trial was a one-back repetition. Each participant completed approximately 1600 trials of this one-back task in ~1.75 hour in each one of two EEG sessions. The two recording sessions were separated by at least one day and by a maximum of two weeks. In total, 105,600 one-back trials with concomitant EEG recording were completed by our forgiving participants.

Analysis

Analyses were performed using custom code written in MATLAB (MathWorks) and Python.

EEG recording and preprocessing

High-density electroencephalographic (EGG) data was continuously recorded at a sampling rate of 1024 Hz using a 128-channel BioSemi ActiveTwo headset (Biosemi B.V., Amsterdam, Netherlands). Electrodes' impedance was kept below 20 µV. Data for half of the super-recogniser participants (SR1-8) as well as 7 typical recognisers were collected at the University of Birmingham, while the other half (SR9-16) as well as 10 typical recognisers were collected at the University of Fribourg. Identical EEG setups were used for all participants. Data was preprocessed using FieldTrip [87] and in-house Matlab code: continuous raw data was first re-referenced relative to A1 (Oz), filtered with a band-pass filter [0.01-80 Hz], segmented into trial epochs from -200 ms to 1100 ms relative to image onset, and down-sampled at 256 Hz. These EEG recordings were made during the one-back task only, not during the other

behavioural tasks. This resulted in a total of 32.266 hours of epoched high-density EEG recordings.

Decoding individual ability from EEG activity

Our first goal was to examine if we could predict an individual's recognition ability based on their brain activity. We recorded a dataset of more than 100,000 trials of a one-back task while high-density EEG was being recorded, and used multivariate pattern analysis "decoding" to predict membership to either the "typical-" or to the "super-recogniser" group from individual trial EEG observations. To test claims that super-recognisers' enhanced visual representations are specific to the processing of faces [82,88], we trained model classifiers on EEG activity of participants viewing either i) face stimuli (varying in identity, facial-expression, and sex) or ii) non-face stimuli (i.e., including various objects, scenes, and animals; see **Figure 1a**). We used Fisher linear discriminant classifiers (5-fold cross-validation, 5 repetitions; [89]) for all decoding analyses of group-membership. The Area Under the Curve (AUC) was used to assess the accuracy of all decoding models used in the current study because, among other reasons, it is resistant to imbalanced groups. Decoding analyses were computed for each EEG session separately, and then averaged across sessions.

Decoding across time. A visualisation of the decoding analysis pipeline is shown in **Supplementary Fig.1**. Succinctly, a Fisher linear discriminant was trained to predict participants' group membership based on EEG activity across all trials of face or non-face condition, for each of the two sessions separately (~26,000 observations per condition, per session, 5-fold cross-validation, 5 repetitions [89]. This process was repeated for all successive 4-ms EEG intervals, starting from -200 ms and ending 1100 ms after stimulus onset, revealing the time course of decoding accuracy.. The resulting evidence indicates when super-recognisers can be categorised from brain activity when processing faces (blue) and non-face stimuli (grey), as shown in **Figure 2a**. We further isolated the brain activity resulting from our participant's discriminating between different stimulus types (e.g., face vs. face discrimination) by completing analyses in which we decoded group-membership from trials where participants specifically had to distinguish a face from another face (i.e., back-to-back face trials), as well as for trials where a non-face image needed to be distinguished from another non-face stimuli (i.e., back-to-back non-face, see **Supplementary figure 3**). These trials necessitated that our participant compare their representations of the presented image and the one stored in short-term memory. This showed similar findings, with one notable difference being that the face-face discrimination condition was the one that obtained peak decoding accuracy.

Searchlight analysis. We conducted a searchlight analysis in which we decoded group-membership from a subset of neighbouring channels (5 neighbours, ± 60 ms windows centred on 135 ms, 350 ms, 560 ms, and 775 ms) and repeated this process over all possible combinations of channels. Evidence for group membership in response to faces peaked over right hemisphere occipito-temporal electrodes (i.e., B9, B10, A30, A30 for time-windows centred on 135 ms, 350 ms, 560 ms, and 775 ms respectively, see **Figure 2a**), while it peaked in electrodes from both hemispheres for the non-face condition (i.e., A17, B10, A27, A27 for time-windows centred on 135 ms, 350 ms, 560 ms, and 775 ms respectively; **Figure 2a**). These results were also replicated in an additional analysis (see **Supplementary figure 2**,

lower panels) targeting trials where participants had to compare specific representations in memory (i.e., a face from another face). Finally, a searchlight analysis combining all EEG time points of a trial (5 neighbours, 0-1100 ms) yielded similar topographies as windows-specific searchlights, with accuracies of 77.3% for faces and 77.5% for non-face conditions (**Figure 2a**, left-most topographies).

Fractional ridge regression. We also attempted to predict individual face recognition ability scores (CFMT+) from EEG-patterns across time using fractional ridge regression models [37]. Ridge regression reduces the collinearity of regressors by penalising the L2-norm of the coefficients of linear regression using a ridge parameter, *k*. Fractional ridge regression is a systematic parameterization of ridge regression. In fractional ridge regression, hyperparameter *k* is chosen from the optimal fraction $\gamma = \|\beta^{\text{ordinary least-square}}\|_2 / \|\beta^{\text{ridge}}\|_2$, the ratio of the L2-norms of the unregularized and regularised coefficients. Here, we constructed ridge regression models that predicted the CFMT+ scores (a vector of size n_{trials} with values ranging from 50 to 100 and where each trial represented the CFMT+ score of a participant, see *Behavioural results* section in Supplementary material) from EEG patterns (a matrix of size n_{trials} x $n_{channels}$, with channels being the regressors). We chose 20 alphas ranging linearly from .001 to .99, training our models on a subset of 60% of the EEG patterns of our sample, validating on 30% and testing on the remaining 10%. This process was repeated 10 times. This approach yielded similar results (shown in **Figure 2b**) as those from the classifiers (shown in **Figure 2a**), with significant prediction of a participant's CFMT+ score as soon as 75 ms (face) and 80ms (non-face), peaking around the N170 (133 and 140 ms respectively for face and non-faces, with values at r_{face} = .4149, $r_{non\text{-}face}$ = .4899), and maintaining late after stimulus offset. Here, again, significance was assessed with permutation testing (see *Group comparisons and inferential statistics* section).

Note that all these decoding effects were present in two separate recording sessions with at least a day between them, although the session-average is shown in the figures.

Noise ceiling for group-membership decoding using CFMT+ reliability scores. To better interpret the magnitude of our group-membership decoding accuracies, we determined the maximum attainable accuracy when categorising a super-recogniser as such using the CFMT+ (i.e., the empirically imposed noise ceiling for our decoding group-membership analysis). To do so, we first simulated 1000 distributions of CFMT+ scores, each with N=32, and each having a Pearson correlation of .71±.01 with the real distribution of CFMT+ scores used in this study. The correlation coefficient of .71 was chosen according to the test-retest reliability the CFMT+ obtained in a previous study [90,91]. We then predicted super-recogniser participants from these distributions, averaged the predicted accuracy across simulated participants, and obtained the maximum accuracy from these simulated distributions. This process was repeated 10K times, creating a distribution of 10K simulated maximums with a mean of M=0.9310 and a standard deviation of SD=0.0228 that could be interpreted as the best attainable accuracy scores for prediction of group-membership (i.e., the noise ceiling).

Representational Similarity Analysis of brain and computational models

The differences in late representations found when decoding individual ability suggested that the brains of super-recognisers varied at several levels of computations. Super-recognisers'

profound ability could arise due to enhanced (i) high-level semantic representations (i.e., conceptual knowledge associated to faces), (ii) activation of feature-like visual representations (e.g., combinations of edges, curvatures such as those outlined in Long et al., 2018 [62]) or (iii) enhanced representations of generic object-like templates. In order to have a better understanding of the computational stages at which these individual differences emerge in the brain processing stream, we associated our participants' brain representations to those from the visual and semantic deep neural networks using Representational Similarity Analysis (RSA [40–43]). RSA captures the geometry between brain representations of different conditions (i.e., commonly different stimuli), and by doing so, generates a representation of a higher-level of abstraction than, say, voxels/topography changes [92]. Here, we employed RSA mainly because this level of abstraction confers one powerful advantage: the means to compare representations of distinct modalities (e.g., M/EEG & fMRI fusion [46]). This approach has indeed yielded interesting insights on the human brain by associating brain representations to that of computational models simulating object recognition [61,93].

Brain Representational Dissimilarity Matrices. We applied RSA to 4 ms intervals of high-density EEG patterns, and characterised the representational dynamics of super-recognisers and typical participants. Specifically, we obtained representational dissimilarity matrices (RDMs) for every participant and each of the two recording sessions by feeding EEG patterns (128 channels) to a linear discriminant whose task was to predict the identity of the stimulus shown, either stimulus s_i or stimulus s_j. By repeating this process for all possible pairwise comparisons of a single time frame (1176 for the 49 stimuli presented), we obtained an RDM. In this symmetrical matrix, larger values represent larger dissimilarity between pairs of stimuli. Cross-validated AUC served as classification accuracy/dissimilarity metric for each of these pairwise predictions. This process was repeated for all time frames (-200 to 1100 ms after stimulus onset), producing idiosyncratic time-resolved representational geometries [46,80]. The group-average time course of important categorical distinctions in the brains of our participants (e.g., face vs non-face images, face gender, and expressions, etc.), as well as the RDMs and 2D multidimensional scaling for specific points in time, are shown in **Supplementary figure 4**.

Visual Deep Convolutional Neural Networks RDMs. We used AlexNet [24], a well-known deep convolutional neural network with hierarchical architecture, as a model to approximate the visual computations along the ventral stream [44]. AlexNet is a feedforward model consisting of 5 convolutional layers—conv1 (96 x 55 x 55), conv2 (256 x 27 x 27), conv3 (384 x 13 x 13), and conv5 (256 x 13 x 13)—as well as three fully-connected layers—fc6 (4096 x 1), fc7 (4096 x 1), and fc8 (1000 x 1) which uses a softmax function to transform the features into class probability. AlexNet's architecture represents visual features of gradually higher complexity and abstraction along each of its 8 layers [44], from low-level features (i.e., contrast, edge), to mid-level features (e.g., combinations of edges, contour, shape, texture) and high-level features (e.g., object parts & entire objects). Long et al. (2018), for example, have shown that representations of mid-level features, which contribute to a large portion of the organisation of the ventral stream, peak with the processing of mid-level layer 4 of AlexNet. Such mid-level features were also recently associated with conscious access to visual information [94]. Another reason for using this model is because of its relative simplicity compared to other more recent DCNNs (which can attain >100 layers [95]). We fed the 49 images presented to our participants to this pre-trained model, and extracted each layer's neural activation. Colour images were used to better parallel the natural processing of the model, which was trained on colour

images. We used these activations as representational patterns that we compared between each pair of images (Pearson correlation), and repeated this process for all pairs of images and each of the 8 layers of the model, creating a total of 8 RDMs (the time course of mutual information between brain and DCNN is shown for layer 4 of AlexNet in **Figure 3a**). We also replicated the findings shown in **Figure 3a** using another well-known DCNN, VGG-16 (see **Supplementary figure 2b**). VGG-16 RDMs were computed in an identical manner as those of AlexNet, targeting the activation of 5 convolutional layers and 3 fully connected layers. The last layer of each convolution "block", e.g., "conv5_3" for conv5, was used to compute these activations.

Semantic Deep Averaging Neural Network RDM. To characterise semantic representations, we first asked participants to provide a precise sentence description of each of the stimuli we presented to our participants during the EEG one-back task. These sentence captions were collected in five independent human observers online, using the Meadows platform (www.meadows-research.com). We obtained written descriptions (e.g., "a city seen from the other side of the forest", see Fig.1d) for every participant, which we fed as input to Google's universal sentence encoder (GUSE [26]) that produces an output vector of 512 dimensions which we refer to as a sentence embedding. We then computed the dissimilarity (cosine distance) between sentence embeddings of every possible pair of sentences (1176 for the 49 stimuli presented), resulting in an RDM for each participant that captures high-level semantic information from human sentences. This RDM is referred to as a semantic RDM.

Association of representations of the brain with computational models

We associated our participants' brain RDMs to those from the visual and semantic deep neural networks (described in the previous section) using Gaussian Copula Mutual Information [27]. Mutual Information is an entropy-based metric that measures the statistical dependence between two variables. It does not make assumptions on the nature of their relationship, and is thus sensitive to both linear and nonlinear relationships. Gaussian copula Mutual Information (GCMI) is a conservative and robust method of measuring MI that uses Gaussian-copula rank normalisation prior to computation of MI. It can be efficiently applied to neuroimaging data.

Mutual information of brain and computational models. We measured the shared representations between brain and computational modes by computing, for each participant, the MI between brain RDMs emerging from each point in time (-200 to 1100 ms after stimulus onset) and the RDM either from the visual DCNNs or the semantic model. For the shared information between brain and DCNN, this created 8 ($n_{participants}$ x n_{time}) matrix, one for each layer, which indicated the contribution of visual features of varying complexity in the brain of our participants. The time course of this information for both groups is shown in **figure 3a** for AlexNet 4th layer. For the shared information between brain and semantic model, this created a single ($n_{participants}$ x n_{time}) matrix which indicated the contribution of high-level semantic relationships of human sentences to the brain representation of our participants.

Conditional mutual information. Conditional Mutual Information (CMI) measures the statistical dependence between two variables X and Y, removing the effect from a third variable Z. It is similar to partial correlation, but is a more powerful approach considering that it removes statistical relationships of any form, whereas partial correlation only removes the

linear relationships emerging from the third variable on X and Y. Here, we used CMI to partial out the effect of specific computations in the brain, i.e., either visual or semantic relationships with brain representations. Indeed, computations of a visual DCNN and semantic models could be shared to some extent, thereby blurring their respective role in our participants' brain representations.

First, to control for possible contributions of semantic computations in the shared representations between brain and DCNNs RDMs, we computed the CMI between brain RDMs and the DCNNs' RDMs, removing the effect from the semantic model RDM. We repeated this for every point in time and every layer of the DCNN, creating a ($n_{participants}$ x n_{time}) matrix for each layer that indicated a contribution of visual features in the brain of our participants without any shared contribution from semantic representations. The time courses of this information, noted CMI(brain;DCNN | semantic), are shown in **Supplementary figure 2b** for each group. Although the magnitude of the MI was clearly reduced compared to MI(brain; DCNN), indicating shared variation between visual and semantic processing (see **Supplementary figure 2a** for a comparison), group contrasts indicated similar effects as the one obtained without controlling for semantic RDMs (see next section on *Group comparison and inferential statistics*).

Second, to control for possible contributions of visual computations in the shared representations between brain and semantic RDMs, we computed the CMI between brain RDMs and the semantic RDM, removing the effect from the RDM of the last layer of AlexNet (see Dwivedi et al., 2021 for a demonstration that these layers are the most task-specific). We repeated this for every point in time, creating a ($n_{participants}$ x n_{time}) matrix that indicated a contribution of semantic representations in the brain of our participants without any contribution of visual categorical representations. The time courses of this information, noted CMI(brain;semantic | DCNN), is shown in **Figure 3b** for each group.

Group comparison and inferential statistics

Comparison of MI time courses. To examine whether face recognition ability impacted specific brain computations, we compared the time course of MI (see the *Association of representations of the brain with computational models* section) between the super and typical recogniser groups using independent samples t-tests and the Monte Carlo method for simulation of distribution, as implemented in the Fieldtrip Toolbox [87]. Family-wise errors were controlled for using cluster-based corrections, with maximum cluster size as cluster-level statistic and an arbitrary *critical-p* value for cluster statistic of .07 for the comparison of MI (brain; semantic) and .01 for the comparison of MI (brain; DCNN) time courses. Control analyses (with CMI (brain; semantic | DCNN) and CMI (brain; DCNN | semantic), see **Supplementary figure 2**) were made in an identical manner.

Time course of group-membership decoding. Significance was assessed using non-parametric permutation testing. Specifically, we constructed an empiric null distribution of decoding values by training the linear classifier to identify shuffled group-membership labels (i.e., a vector of size ~50,000) from experimental EEG patterns (128 channels x ~50,000 trials). This process was repeated for each time point and each one of the two sessions. Considering that the average of the decoding time course from the two sessions were shown, we then

averaged the resulting decoding values of both sessions (i.e., the null hypothesis being that the *average* of decoding values could not be dissociable than the ones obtained from the session-averaged distributions trained from shuffled labels), and repeated this process 1000 times. We then compared the real, experimental decoding value at each time point to its corresponding null distribution, and rejected the null hypothesis if the decoding value was greater than the prescribed critical value at a p <.001 level.

***Time course of individual ability decoding using ridge regression*.** Significance was again assessed using non-parametric permutation testing. Specifically, we constructed null distributions of ridge values by constructing ridge regression models that predicted the randomly shuffled CFMT+ scores (a vector of size n_{trials} where each trial represented the CFMT+ score of a randomly chosen participant) from EEG patterns (a matrix of size n_{trials} x $n_{channels}$, with channels being the regressors). We repeated this process (further described in the *Fractional ridge regression* section) for 1000 permutations. We finally compared the real, experimental regression value at each time point to its corresponding null distribution, and rejected the null hypothesis if the decoding value was greater than the prescribed critical value at a p <.01 level.

Data availability

The data that support the findings of this study will be available online upon publication of the manuscript.

Code availability

The MATLAB and Python codes that support the findings of this study will be available online upon publication of the manuscript.

Acknowledgements

We thank Prof. Josh P. Davis for sharing behavioural scores of super-recognisers and establishing first contact to the UK-based Super-Recognizers reported here. Funding for this project was supported by an ERC Starting Grant [ERC-StG-759432] to I.C, an ERSC-IAA grant to J.W., I.C. and S.F.S., by a Swiss National Science Foundation PRIMA (Promoting Women in Academia) grant [PR00P1_179872] to MR, and by NSERC and IVADO graduate scholarships to S.F.S. We also thank Mick Neville, from Super-Recognisers Ltd., who helped us to get in contact with some of our super-recognizer participants.

Author contributions
(CRediT standardised author statement)

S.F-S. : Conceptualisation, methodology, software, formal analysis, investigation, data curation, writing - original draft, visualisation, supervision, project administration, funding acquisition. **M.R. :** Investigation, resources, project administration, funding acquisition,

supervision, writing - review and editing. **E.B.:** investigation, project administration. **M.Z.**: investigation. **J.W.:** funding acquisition, writing - review and editing. **A-R.R.**: Investigation. **R.C.:** Resources. **F.G.:** Methodology, writing - original draft, supervision, resources. **I.C.**: Supervision, Methodology, software, resources, formal analysis, writing - original draft, project administration, funding acquisition.

Competing interests

The authors declare no competing interests.

References

1. Duchaine, B. & Yovel, G. *Annu Rev Vis Sci* **1**, 393–416 (2015).
2. Grill-Spector, K., Weiner, K.S., Kay, K. & Gomez, J. *Annu Rev Vis Sci* **3**, 167–196 (2017).
3. Kanwisher, N., McDermott, J. & Chun, M.M. *J. Neurosci.* **17**, 4302–4311 (1997).
4. Susilo, T. & Duchaine, B. *Current Opinion in Neurobiology* **23**, 423–429 (2013).
5. Noyes, E., Phillips, P.J. & O'Toole, A.J. *Face processing: Systems, disorders and cultural differences* 173–201 (2017).
6. Russell, R., Duchaine, B. & Nakayama, K. *Psychon. Bull. Rev.* **16**, 252–257 (2009).
7. Ramon, M. *Neuropsychologia* 107809 (2021).doi:10.1016/j.neuropsychologia.2021.107809
8. Herzmann, G., Kunina, O., Sommer, W. & Wilhelm, O. *J. Cogn. Neurosci.* **22**, 571–589 (2010).
9. Huang, L. et al. *Front. Hum. Neurosci.* **8**, 483 (2014).
10. Kaltwasser, L., Hildebrandt, A., Recio, G., Wilhelm, O. & Sommer, W. *Cogn. Affect. Behav. Neurosci.* **14**, 861–878 (2014).
11. Lohse, M. et al. *J. Neurosci.* **36**, 3821–3828 (2016).
12. Nowparast Rostami, H., Sommer, W., Zhou, C., Wilhelm, O. & Hildebrandt, A. *Cortex* **95**, 192–210 (2017).
13. Rossion, B., Retter, T.L. & Liu-Shuang, J. *Eur. J. Neurosci.* **52**, 4283–4344 (2020).
14. Elbich, D.B. & Scherf, S. *Neuroimage* **147**, 409–422 (2017).
15. Haxby, J.V. et al. *Science* **293**, 2425–2430 (2001).
16. Kay, K.N. & Yeatman, J.D. *Elife* **6**, (2017).
17. Harel, A., Kravitz, D. & Baker, C.I. *Front. Hum. Neurosci.* **7**, 885 (2013).
18. Dwivedi, K., Bonner, M.F., Cichy, R.M. & Roig, G. *PLoS Comput. Biol.* **17**, e1009267 (2021).
19. di Oleggio Castello, M.V., Haxby, J.V. & Ida Gobbini, M. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
20. Duchaine, B. & Nakayama, K. *Neuropsychologia* (2006).at <https://www.sciencedirect.com/science/article/pii/S0028393205002496>
21. Tardif, J. et al. *Psychol. Sci.* **30**, 300–308 (2019).
22. Fysh, M.C., Stacchi, L. & Ramon, M. *R Soc Open Sci* **7**, 200233 (2020).
23. Faghel-Soubeyrand, S. et al. *J. Vis.* **19**, 136d–136d (2019).
24. Krizhevsky, A., Sutskever, I. & Hinton, G.E. *Advances in Neural Information Processing Systems 25* 1097–1105 (2012).at

<http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>

25. Simonyan, K. & Zisserman, A. *arXiv [cs.CV]* (2014).at <http://arxiv.org/abs/1409.1556>
26. Cer, D. et al. *arXiv [cs.CL]* (2018).at <http://arxiv.org/abs/1803.11175>
27. Ince, R.A.A. et al. *Hum. Brain Mapp.* **38**, 1541–1573 (2017).
28. Bentin, S., Allison, T., Puce, A., Perez, E. & McCarthy, G. *J. Cogn. Neurosci.* **8**, 551–565 (1996).
29. Rossion, B. & Jacques, C. *The Oxford handbook of event-related potential components.* **641**, 115–141 (2012).
30. Geskin, J. & Behrmann, M. *Cogn. Neuropsychol.* **35**, 4–54 (2018).
31. Behrmann, M. & Avidan, G. *Trends Cogn. Sci.* **9**, 180–187 (2005).
32. Garrido, L., Duchaine, B. & DeGutis, J. *Cogn. Neuropsychol.* **35**, 55–58 (2018).
33. Kanwisher, N. *Nat. Neurosci.* **3**, 759–763 (2000).
34. Grill-Spector, K., Knouf, N. & Kanwisher, N. *Nat. Neurosci.* **7**, 555–562 (2004).
35. Barton, J.J.S. & Corrow, S.L. *Neuropsychologia* **89**, 119–124 (2016).
36. Bobak, A.K., Parris, B.A., Gregory, N.J., Bennetts, R.J. & Bate, S. *Quarterly Journal of Experimental Psychology* **70**, 201–217 (2017).
37. Rokem, A. & Kay, K. *Gigascience* **9**, (2020).
38. McDermott, J., Schiller, P.H. & Gallant, J.L. *Proceedings of the* (2002).at <https://www.pnas.org/content/99/3/1645.short>
39. Lamme, V.A. & Roelfsema, P.R. *Trends Neurosci.* **23**, 571–579 (2000).
40. Kriegeskorte, N., Mur, M. & Bandettini, P. *Front. Syst. Neurosci.* **2**, 4 (2008).
41. Kriegeskorte, N. et al. *Neuron* **60**, 1126–1141 (2008).
42. Kriegeskorte, N. & Kievit, R.A. *Trends Cogn. Sci.* **17**, 401–412 (8/2013).
43. Charest, I., Kievit, R.A., Schmitz, T.W., Deca, D. & Kriegeskorte, N. *Proceedings of the National Academy of Sciences* **111**, 14565–14570 (2014).
44. Güçlü, U. & van Gerven, M.A.J. *J. Neurosci.* **35**, 10005–10014 (2015).
45. Carlson, T.A., Tovar, D.A., Alink, A. & Kriegeskorte, N. *J. Vis.* **13**, 1–1 (2013).
46. Cichy, R.M., Pantazis, D. & Oliva, A. *Nat. Neurosci.* **17**, 455–462 (2014).
47. Dubois, J. & Adolphs, R. *Trends Cogn. Sci.* **20**, 425–443 (2016).
48. Gauthier, I., Skudlarski, P., Gore, J.C. & Anderson, A.W. *Nat. Neurosci.* **3**, 191–197 (2000).
49. Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H. & Livingstone, M.S. *Science* **311**, 670–674 (2006).
50. Haxby, J.V., Hoffman, E.A. & Gobbini, M.I. *Trends Cogn. Sci.* **4**, 223–233 (2000).
51. Bruce, V. & Young, A. *Br. J. Psychol.* **77 (Pt 3)**, 305–327 (1986).
52. Busigny, T., Graf, M., Mayer, E. & Rossion, B. *Neuropsychologia* **48**, 2051–2067 (2010).
53. Dricot, L., Sorger, B., Schiltz, C., Goebel, R. & Rossion, B. *Neuroimage* **40**, 318–332 (2008).
54. Rossion, B. *J. Neuropsychol.* (2018).doi:10.1111/jnp.12162
55. Avidan, G. et al. *Cereb. Cortex* **24**, 1565–1578 (2014).
56. Duchaine, B.C. & Nakayama, K. *Curr. Opin. Neurobiol.* **16**, 166–173 (2006).
57. Moscovitch, M., Winocur, G. & Behrmann, M. *J. Cogn. Neurosci.* **9**, 555–604 (1997).
58. McGugin, R.W., Gatenby, J.C., Gore, J.C. & Gauthier, I. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17063–17068 (2012).
59. Behrmann, M. & Plaut, D.C. *Trends Cogn. Sci.* **17**, 210–219 (2013).
60. Fan, X., Wang, F., Shao, H., Zhang, P. & He, S. *eLife* **9**, (2020).
61. Khaligh-Razavi, S.-M. & Kriegeskorte, N. *PLoS Comput. Biol.* **10**, e1003915 (2014).

62. Long, B., Yu, C.-P. & Konkle, T. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E9015–E9024 (2018).
63. Ince, R.A.A. et al. *Cerebral Cortex* **26**, 4123–4135 (2016).
64. Hoffman, E.A. & Haxby, J.V. *Nat. Neurosci.* **3**, 80–84 (2000).
65. Gauthier, I. et al. *J. Cogn. Neurosci.* **12**, 495–504 (2000).
66. Ramon, M., Vizioli, L., Liu-Shuang, J. & Rossion, B. *Proceedings of the National Academy of Sciences* **112**, E4835–E4844 (2015).
67. Collins, J.A. & Olson, I.R. *Neuropsychologia* **61**, 65–79 (2014).
68. Jonas, J. et al. *Cortex* **72**, 140–155 (2015).
69. Cohen, A.L. et al. *Brain* **142**, 3975–3990 (2019).
70. Popham, S.F. et al. *Nature Neuroscience* **24**, 1628–1636 (2021).
71. Liu, J. et al. *Sci Adv* **7**, eabg9715 (2021).
72. Fernandino, L., Tong, J.-Q., Conant, L.L., Humphries, C.J. & Binder, J.R. *Proc. Natl. Acad. Sci. U. S. A.* **119**, (2022).
73. Gao, C., Conte, S., Richards, J.E., Xie, W. & Hanayik, T. *Psychophysiology* **56**, e13336 (2019).
74. van Herten, M., Kolk, H.H.J. & Chwilla, D.J. *Brain Res. Cogn. Brain Res.* **22**, 241–255 (2005).
75. Shen, W., Fiori-Duharcourt, N. & Isel, F. *Neuroreport* **27**, 548–558 (2016).
76. Eimer, M., Gosling, A. & Duchaine, B. *Brain* **135**, 542–554 (2012).
77. Burton, A.M., Young, A.W., Bruce, V., Johnston, R.A. & Ellis, A.W. *Cognition* **39**, 129–166 (1991).
78. Faghel-Soubeyrand, S., Dupuis-Roy, N. & Gosselin, F. *J. Exp. Psychol. Gen.* **148**, 1834–1841 (2019).
79. Adolphs, R. et al. *Nature* **433**, 68–72 (2005).
80. Hebart, M.N., Bankson, B.B., Harel, A., Baker, C.I. & Cichy, R.M. *eLife* **7**, (2018).
81. Shah, P., Gaule, A., Sowden, S., Bird, G. & Cook, R. *Royal Society Open Science* **2**, 140343 (2015).
82. Bobak, A.K., Bennetts, R.J., Parris, B.A., Jansari, A. & Bate, S. *Cortex* **82**, 48–62 (2016).
83. Davis, J.P., Lander, K., Evans, R. & Jansari, A. *Appl. Cogn. Psychol.* **30**, 827–840 (2016).
84. Bobak, A.K., Pampoulov, P. & Bate, S. *Front. Psychol.* **7**, 1378 (2016).
85. Langner, O. et al. *Cognition and Emotion* **24**, 1377–1388 (2010).
86. Willenbockel, V. et al. *Behav. Res. Methods* **42**, 671–684 (2010).
87. Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J.-M. *Comput. Intell. Neurosci.* **2011**, 156869 (2011).
88. Belanova, E., Davis, J.P. & Thompson, T. *Cortex* **108**, 92–111 (2018).
89. Treder, M.S. *Front. Neurosci.* **14**, 289 (2020).
90. Murray, E. & Bate, S. *R Soc Open Sci* **7**, 200884 (2020).
91. Arrington, M., Elbich, D., Dai, J., Duchaine, B. & Suzanne Scherf, K. *Behavior Research Methods* (2022).doi:10.3758/s13428-022-01805-8
92. Kriegeskorte, N. & Diedrichsen, J. *Annu. Rev. Neurosci.* **42**, 407–432 (2019).
93. Spoerer, C.J., Kietzmann, T.C., Mehrer, J., Charest, I. & Kriegeskorte, N. *PLoS Comput. Biol.* **16**, e1008215 (2020).
94. Lindh, D., Sligte, I.G., Assecondi, S., Shapiro, K.L. & Charest, I. *Nature Communications* **10**, (2019).
95. He, K., Zhang, X., Ren, S. & Sun, J. *Proceedings of the IEEE conference on computer vision and pattern recognition* 770–778 (2016).at <http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html>

96. Kaneshiro, B., Perreau Guimaraes, M., Kim, H.-S., Norcia, A.M. & Suppes, P. *PLoS One* **10**, e0135697 (2015).
97. Dobs, K., Isik, L., Pantazis, D. & Kanwisher, N. *Nat. Commun.* **10**, 1258 (2019).

Supplementary material

Behavioural results

All participants' face recognition ability was assessed using the Cambridge Face Memory Test long-form (CFMT+[20]). Scores on the CFMT+ ranged from 50 to 85 in the typical recognisers group (M=70.00; stdev=9.09), and from 92 to 100 in the experimental super-recogniser group (M=95.38, stdev=2.68; difference between groups : $t(31)=10.6958$, p<.00001 see Figure 1a). During the main experimental task, participants completed a one-back task (Figure 1b). Accuracy was significantly greater for the SRs (M_{SRs}=.93, stdev = .054) than for the controls (M_{ctrls}=.83, stdev=.094; $t(31)=3.7911$, p = 6.50e-04) for both face (M_{SRs}= .9260, stdev=.0471; M_{ctrls}=.8144, stdev=.1066; $t(31)=3.8440$, p = 5.6242e-04) and non-face images (M_{SRs}=.9456, stdev=.0651; M_{ctrls}=.8532, stdev=.0929; $t(31)=3.2855$, p =.0025). We observed no significant differences in response times between the two groups (p > .3). Furthermore, accuracy in the one-back task was positively correlated with scores on the CFMT+ (r = .68, p<.001; RT was marginally associated with CFMT+, r = .37, p =.04), suggesting that this effect does not emerge from different lapse rates in our participants.

Supplementary table 1. Birmingham

subject-ID	CFMT+	GFMT	Face Array	LASIE Match	Black and White	Super-rec	Longterm	one-back faces	one-back non-faces
SR-1 (JP1)	93	40	38	83	34	13	9	.9328	0.977
SR-2 (EG1)	95	40	37	88	33	12	9	.8500	0.7903
SR-3 (JKW)	97	40	32	84	33	12	8	.9325	0.9718
SR-4 (EM1)	93	40	38	92	34	13	8	.9726	0.9835
SR-5 (SM1)	98	40	31	91	33	14	9	.9326	0.9787
SR-6 (KH1)	100	40	39	82	40	14	10	.9823	0.9817
SR-7 (ES1)	98	40	40	92	33	13	8	.9319	0.9647
SR-8 (DE1)	96	38	40	90	39	12	10	.9362	0.9787
Max_score	102	40	40	100	40	14	10	1	1

Supplementary table 2. Fribourg (see also [7])

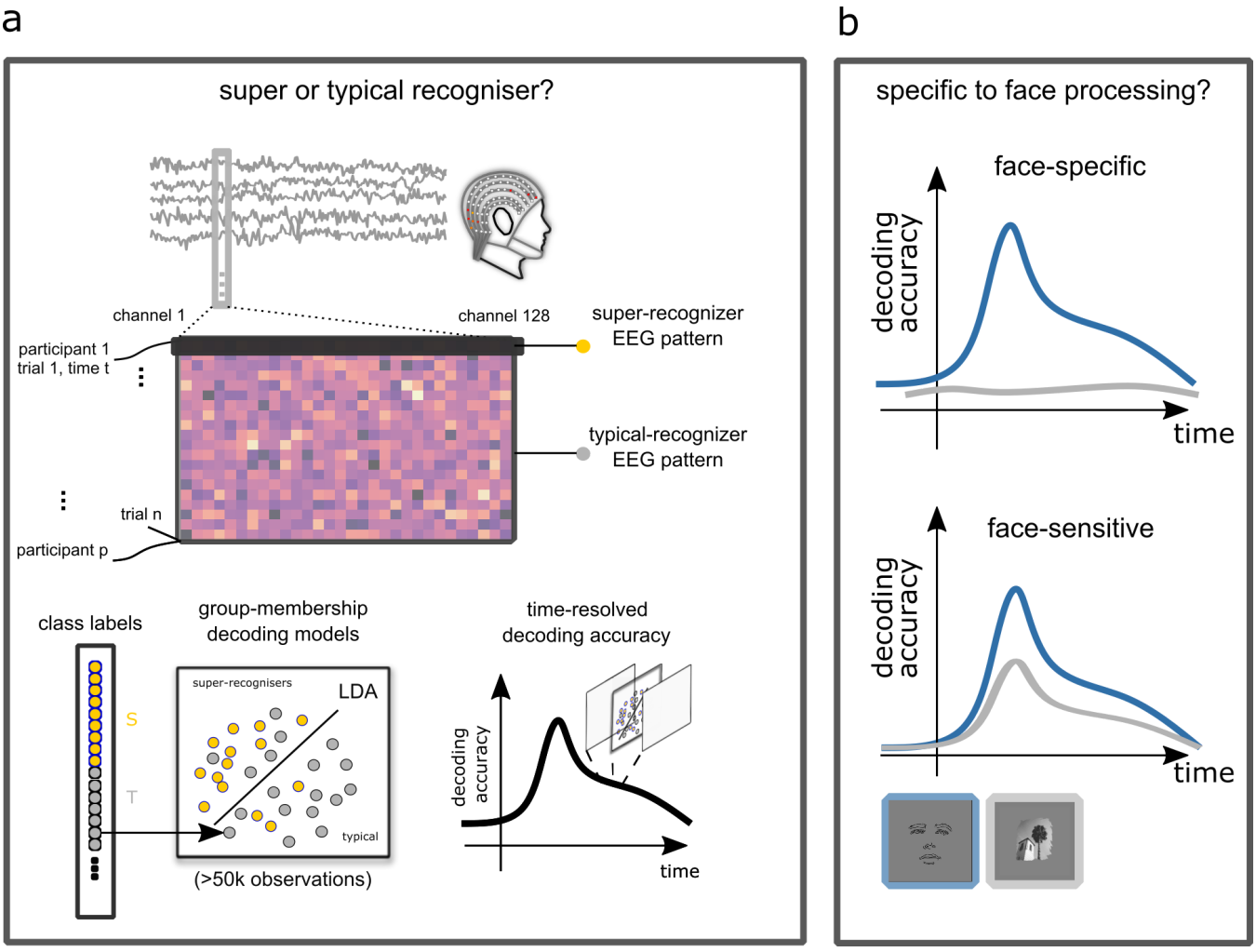
subject-ID	CFMT+	FICST score	YBT long raw score	one-back faces	one-back non-faces
SR-9 (RF1)	92	0	29	.8132	0.8384
SR-10 (M1)	99	0	17	.9409	0.9619
SR-11 (UC1)	92	0	20	.9635	0.9808
SR-12 (MW1)	93	1	20	.8581	0.8212
SR-13 (NC1)	92	7	17	.9530	0.985
SR-14 (MB2)	96	0	18	.9604	0.9886
SR-15 (AH1)	94	3	16	.9112	0.9624
SR-16 (FW1)	97	1	22	.9792	0.9643
Max_score	102	0	35	1	

Univariate analyses of EEG associated with individual ability

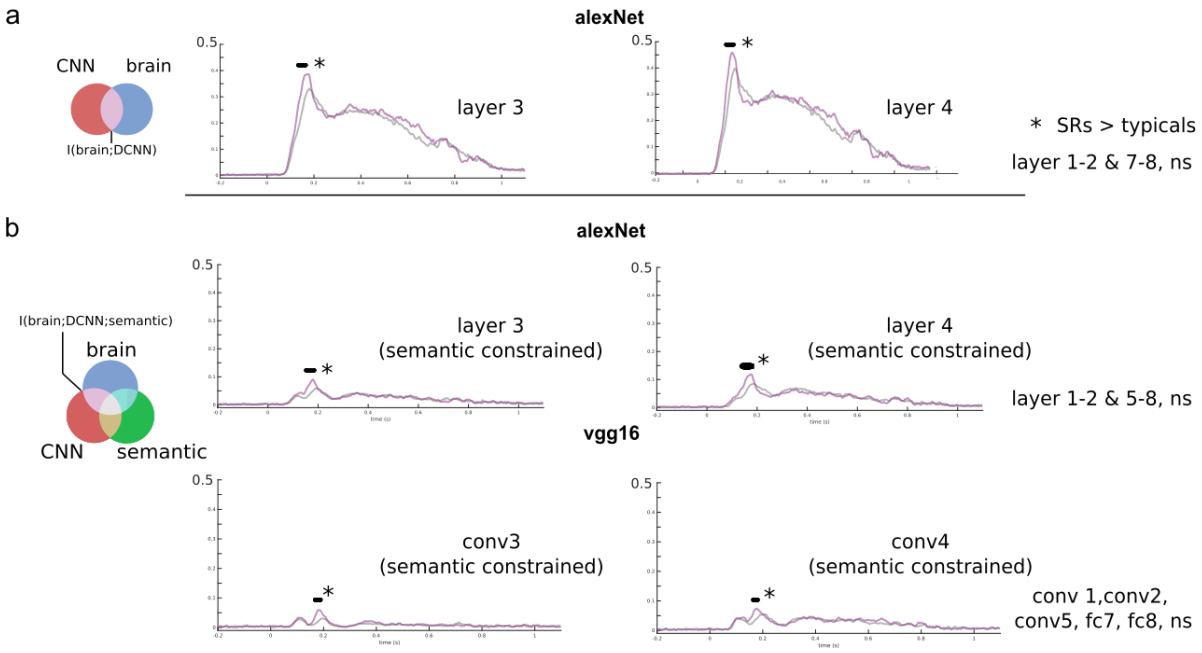
All univariate analyses were completed *post-hoc.*

N170 amplitude and latency. To better compare our findings with more traditional face-ERP studies, we also computed tests on the N170 time window of the ERPs of both groups. We extracted peak ERP amplitude and latency for every participant in a window corresponding to the N170 component (computed as the minimum amplitude within 110-200 ms at electrodes [B6 – B7 – B8 – A28] on the right hemisphere and [A9 – A10 – A11 – A15] on the left hemisphere), for each of the 2 presentation conditions (face, non-face), and tested the group and conditions effects with an ANOVA. The univariate effects found indicate that the peak decoding results observed at 135 ms (for super vs typical recognizers) emerged mainly from SRs having earlier N170 ($F_{group}(60,1) = 19.23$, $p<.0001$; $F_{conditions}(60,1) = 5.86$, $p = .0185$; $F_{interaction}(60,1) = 1$, $p = .32$) and more ample N170 effects compared to typical recognisers ($F_{group}(60,1) = 13.75$, $p=.0005$; $F_{conditions}(60,1) = 33.78$, $p < .0001$; $F_{interaction}(60,1)=0.32$, $p > .50$). Importantly, an absence of interaction effects were found for both peak amplitude and latency. This further indicates that brain processing differences in the N170 window are not specific to face stimuli, i.e. they extend beyond faces in super-recognisers.

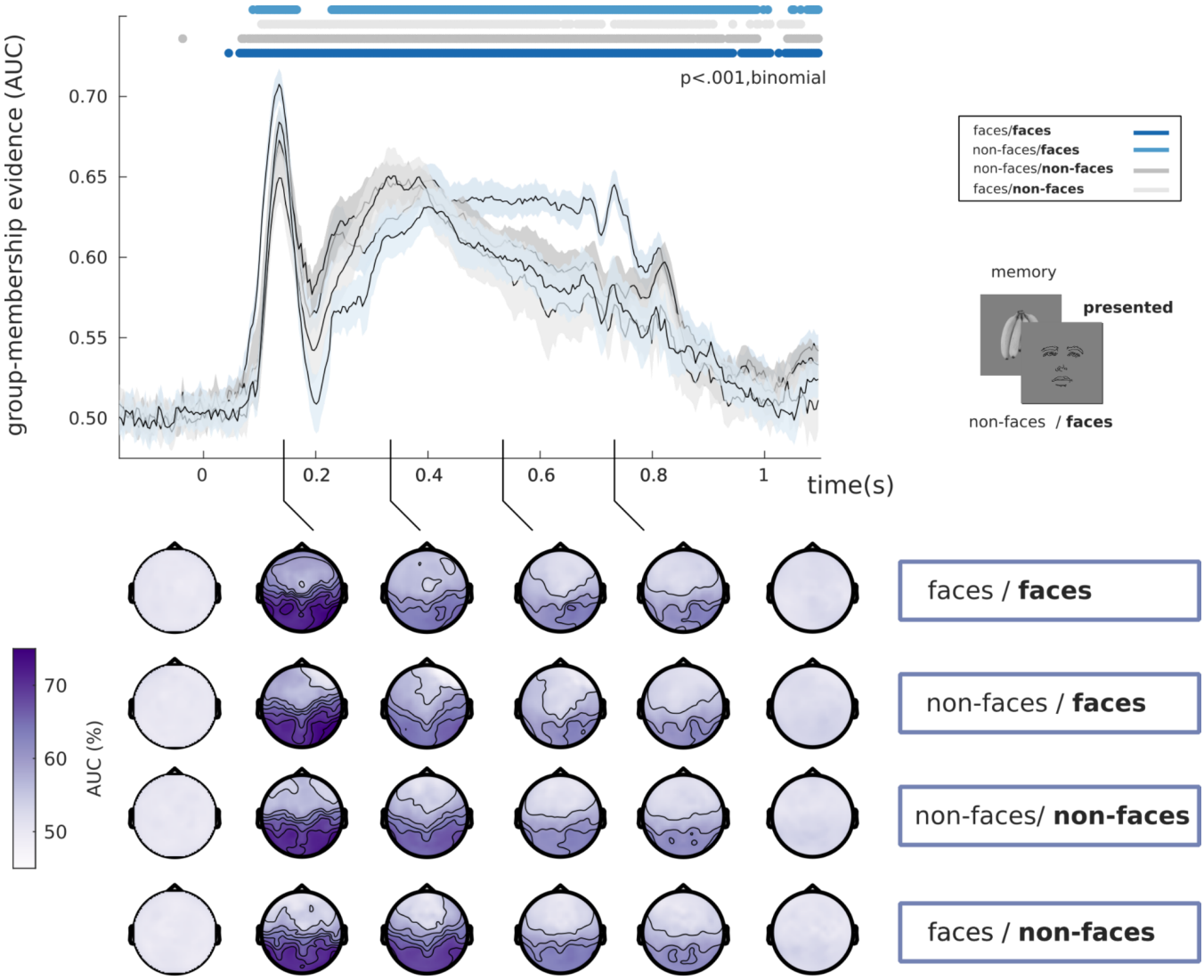
Lateralisation. Univariate analyses further revealed a lateralisation effect that was hinted at by those shown with decoding searchlights analyses (described in the method section). Indeed, the more right-lateralized the N170 was in an individual (computed as the minimum amplitude within 110-200 ms at electrodes [B6 – B7 – B8 – A28] on the right hemisphere and [A9 – A10 – A11 – A15] on the left hemisphere), the higher score he/she had at the CFMT+ ($r(31) = -.44$, $p = .01$). This finding was also present on group contrasts : super-recognisers showed stronger N170 amplitude asymmetry to the right hemisphere compared to typical recognisers ($t(30) = -2.8542$, $p = .01$). The N170 was not lateralised for non-face stimuli ($p > .25$).



Supplementary figure 1. Visualisation of EEG decoding analyses. (a) We probed the brain dynamics behind the outstanding visual ability of super-recognisers using multivariate pattern analysis applied to more than 100,000 observations of the recorded time-resolved EEG patterns. Specifically, EEG-patterns (128 channel topographies) were fed to a linear classifier whose task was to predict whether it belonged to the super or typical recogniser group. A classifier for each 4 ms time step was trained, validated and tested to reveal the time course of group-membership decoding accuracy. (b) An identical process was repeated for face and non-face conditions, which permitted us to test the hypothesis that the brain representations underlying face recognition ability are face-specific (i.e. no evidence for face recognition ability from non-face brain processing, first panel) or face sensitive (i.e. presence of evidence for face recognition ability from non-face brain processing). Note that a face drawing is depicted here as an anonymised substitute to the experimental face stimuli presented to our participants.

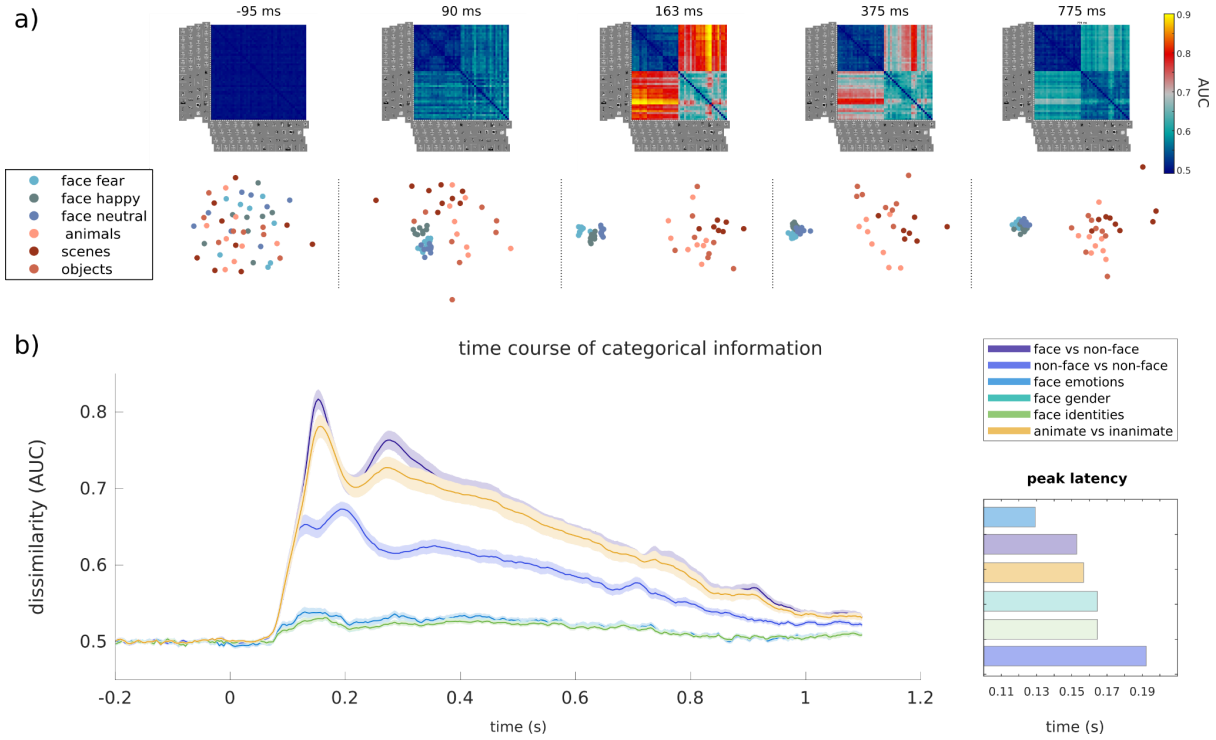


Supplementary figure 2. Time course of similarity between brain RDMs and DCNNs RDMs, constrained on semantic model RDMs. a) Mutual Information between RDMs computed from layer 3 and 4 of AlexNet and brain RDMs at every time point was derived, and showed significant contrasts between super and typical recognisers around 130-160 ms. b) To ensure that these effects could not emerge from shared semantic/DCNNs computations, we computed the MI between DCNN and the brain, constrained on the Mutual Information from the semantic model RDM. Although it reduced the magnitude of shared representations (indicating similarly shared information between semantic, brain and DCNN RDMs), this showed significant effects with AlexNet's layer 3 and 4 (b, first row) within the same time window. An identical analysis pipeline was completed with another DCNN, VGG-16[24,25] (further described in the *Visual Deep Convolutional Neural Networks RDMs* section). Here, again, we found a significant effect between super and typical recognisers within the same time window, and within the same layers (conv3 and conv4).



Supplementary figure 3. To isolate the brain activity resulting from the discrimination between different stimuli types (e.g. face-to-face discrimination), we also computed the time course of evidence for group membership from trials where participants specifically had to distinguish a face from another face (i.e. back-to-back face trials), as well as for trials where a non-face image needed to be distinguished from another non-face stimuli (i.e. back-to-back non-face). Group-membership multivariate pattern analysis were computed in an identical manner as the ones shown on Fig.2, but from topographies where participants had to detect changes from within category stimuli (i.e. face vs. face) and between category stimuli (i.e. face vs. non-face stimulus). This showed similar timecourses as previously described for face and non-face conditions, although in this case the evidence in the face-vs-face attained higher values compared to other conditions in late time points (~400-750ms). This suggests larger brain processing differences for face representations (compared to non-face representations) in

super-recognisers Note that a face drawing is depicted here as an anonymised substitute to the experimental face stimuli presented to our participants.



Supplementary figure 4. EEG representational geometry dynamics. a) Representational Similarity Analysis (RSA) was applied to time-resolved EEG patterns, using decoding AUC as dissimilarity measure between pairs of images (5 fold cross-validation, 5 repetitions) to create Representational Dissimilarity Matrices (RDMs). Multidimensional scaling was employed to visualise these high-dimensional brain representations on a 2D plane, which showed clear distinctions between various categories (e.g. face clusters, scenes clusters, animal clusters, etc.). b) We revealed categorical information unfolding in time by averaging dissimilarities from different pairs of comparisons between representations (e.g. faces vs non-face objects) and averaged across participants. The peak latencies for different categorical distinctions in the brain of our participants are shown on the rightmost panel. Brain representations for the distinction of face vs. non-face objects (a hallmark of the N170 [29]) dominated all other categorical distinctions [45,96], and peaked at 153ms (about 40 ms sooner than information discriminating within non-face objects, which peaked at 192 ms). For within-face categories, we found that evidence for face emotions peaked sooner (around 130 ms) than face gender and face identity (both peaking at 165 ms; see also [97] for similar results using MEG).