nature portfolio

https://doi.org/10.1038/s41562-022-01434-3

Supplementary information

Relationship between nuclei-specific amygdala connectivity and mental health dimensions in humans

In the format provided by the authors and unedited

Supplementary Information

Relationship between nuclei-specific amygdala connectivity and mental health dimensions in humans

Miriam C Klein-Flügge, Daria EA Jensen, Yu Takagi, Luke Priestley, Lennart Verhagen, Stephen M Smith, Matthew FS Rushworth

Supplementary Results

Comparing the accuracy of nuclei-specific and whole-amygdala predictions

We tested whether parcellating the amygdala into subnuclei increased the accuracy of predictions of mental health dimensions. We repeated the robust regressions for the amygdala as a whole, i.e., using functional connectivity between the entire amygdala and the same set of 28 ROIs. As reported in the main text, the robust regression weights of all 28 edges derived from the 3T participants were applied to the 7T participants to obtain out-of-sample predictions. This produced significant predictions of negative emotions (one-sided Pearson's r(95)=0.2, p=0.026, CI=[0,.38]) and anger (Pearson's r(95)=0.18, p=0.037, CI=[-.01,.37]; **Fig6A**). However, predictions for all four behavioural dimensions were worse when considering whole-amygdala rather than nuclei-specific amygdala functional connectivity (LifeSat: nuclei: r=0.19, whole: r=0.17; NegEmot: nuclei: r=0.22, whole: r=0.20; Sleep: nuclei: r=0.05, whole: r=-0.09; Anger: nuclei: r=0.22, whole: r=0.18; compare **Figs 4E and 6A**). We note that – while the nuclei-specific predictions include seven times as many predictors as used in the whole amygdala control – the whole amygdala is made up of the exact same voxels as the seven nuclei together. If the subdivisions of our parcellation are not meaningful or if the fMRI signal in individual nuclei and therefore the derived functional connectivity values for individual nuclei are not robust, then separation into smaller subsets of voxels should merely increase the noise in our predictions. Alternatively, if there is value in considering amygdala nuclei separately, then using functional connectivity values from each nucleus might increase our prediction accuracy for predicting mental health dimensions. Nevertheless, to account for the number of predictors, we generated separate null distributions for the nuclei- and whole-amygdala versions to obtain directly comparable p-values (Methods). This slightly changed the precise p-values, but led to the same conclusions, suggesting negative emotions and anger could be predicted significantly (in relation to this this null distribution: p=0.026 and p=0.038, respectively). However, most importantly, the predictions of all four mental health dimensions still significantly benefited from considering the functional connectivity of separate amygdala nuclei as opposed to the amygdala as a whole (comparable p-values relative to appropriate null distribution from one-sided test assessing a positive relationship: LifeSat: nuclei: p=0.034, whole: p=0.051; NegEmot: nuclei: p=0.016, whole: p=0.026; Sleep: nuclei: p=0.310, whole: p=0.807; Anger: nuclei: p=0.015,

whole: p=0.038; **Fig 6B**). The same conclusion was reached when comparing the peak accuracy instead of the prediction achieved with the full set of edges (**Extended Data Fig 9A**). Peak accuracy using whole-amygdala edges occurred at n=16, 12, 3 and 9 for the four behavioural dimensions and fingerprints and scatterplots are shown for the earliest significant peak (**Fig 6C,D**) and global peak (**Extended Data Fig 9B**) for comparison with the nuclei version. However, even at the peak, prediction accuracies from whole-amygdala functional connectivity were worse than those achieved using nuclei-specific predictions reported above (correlation coefficients for whole-amygdala functional connectivity at the peak: r=0.185, 0.202, 0.127, 0.198; **Extended Data Fig 9**). Thus, despite their small size, considering the functional connectivity of individual subcortical nuclei benefitted predictions of mental health dimensions.

Supplementary Discussion

There is evidence that resting-state correlations reflect anatomical connectivity[1]. However, it is worth noting that, while rs-fMRI was used as a proxy for anatomical connectivity here, the patterns in resting-state functional connectivity we identify do not necessarily correspond to monosynaptic connections. Monosynaptic connections might dominate the positive functional connectivity values in **Figure 2B** which illustrates the mean functional connectivity of the amygdala nuclei with other ROIs. However, the negative functional connectivity observed between dlPFC regions and the amygdala likely reflects an indirect interaction mediated by another brain region given limited direct connections between the homologue of this region and the amygdala in macaques. Similarly, it is possible that the relations we identified between functional connectivity and mental health indices (**Figures 4-6**) may rely on a multi-component connection pathway or may involve connections between two amygdala nuclei.

We used an identical amygdala parcellation for all participants. This parcellation was based on our initial group dataset but could be replicated in additional 3T and 7T cohorts (**Extended Data Fig 4**). The amygdala is a relatively more preserved structure than cortex. There is individual variation in the presence, position, and organizational patterning of sulci and gyri in the cortex [2,3] but individual variability in amygdala nuclei has not been documented. Similarly, there is no evidence for major interindividual variation in other

subcortical nuclei, for example in the brainstem. Nevertheless, examining whether the shape or location of amygdala nuclei or other brain regions might systematically relate to mental health markers is an interesting avenue for future work. There is evidence for atypical neurodevelopment linked to transdiagnostic markers [4] and resting-state connectivity at rest (taken as a proxy for long-range brain connections) can predict the shape of task-related activations [5]. If the location or shape of a given nucleus or region of interest varies with a persons' mental health, then using a group mask which captures a given participant's true region more or less well could contribute to the effects we report. Investigating this question will require future work to examine individual parcellations which was not attempted here.

Supplementary Methods

Participant selection

As described in the main text, participants were selected based on two criteria: the quality of the physiological variables acquired and their total DSM/ASR scores. The DSM/ASR scores were not used in any of the key analyses. Working on a subset of all 1206 HCP participants was necessary because one key aspect of the pre-processing was to correct rs-fMRI data for physiological noise, which particularly affects the key regions of this study such as the amygdala and brainstem. However, the quality of physiological data that has been acquired varies substantially across HCP participants. For example, only n=764 out of 1206 3T-HCP participants have recordings of physiological noise regressors for all four resting-state runs, and only n=636 have complete resting-state data, physiological recordings and behavioural scores (required for factor analysis, see below). A second dataset of the same size as the first (n=200) was selected for replication from the remaining n=436 3T-participants with complete data. It was not possible to match the variance in DSM/ASR scores to the first dataset and participants did not have quite as high-quality physiological noise recordings, but we still prioritized inclusion of participants at the upper and lower ends of the distribution of DSM scores (resulting DSM mean/variance in n=200: 3.96, 10.71; ASR mean/variance: 35.72, 481.67) and those with the largest number of runs with high-quality physiological noise recordings (severe problems in a maximum of 3 out of 8 = 4 cardiac + 4 respiratory traces). The resulting demographics in this second dataset of n=200 3T participants were mean age

28 ± .28; age range 22-36; 99 females, 101 males. A third dataset contained all 7T-HCP young adult participants not already included in either of the 3T datasets and with full resting-state and behavioural data, which left us with n=98 7T-HCP participants (mean age 29 ± .33; age range 23-36; 59 females, 39 males; DSM mean/variance: 3.43, 5.73; ASR mean/variance: 31.79, 253.43; **Supplementary Table 3**). Physiological noise recordings are not available in the 7T data, but the higher field strength significantly improves the tSNR in subcortical regions.

Validation of additional preprocessing

Additional pre-processing focussed on removal of physiological artefacts and led to gains in the temporal signal-to-noise ratio (tSNR) in amygdala and many of its projection targets and sources, e.g., medial temporal lobe areas, subgenual prefrontal cortex, and most prominently subcortical and brainstem structures (see Methods and **Extended Data Fig 1, B-C**). This pattern of tSNR changes is in line with prior work that shows physiological noise particularly affects brainstem and other subcortical regions because of their proximity to major vessels and pulsating fluid-filled spaces[6].

It is not possible to fully validate the additional pre-processing step for physiological noise correction without knowledge of the ground truth, which is not available in resting-state data. However, there is good reason to believe that our additional pre-processing helped recover meaningful signal in our data. First, because it resulted in tSNR changes in regions known to be most affected by pulsation artefacts (**Extended Data Fig 1B**), consistent with prior work. Second, because it improved a parcellation of the amygdala in two subsets of n=200 participants corrected for physiological noise, compared to the full dataset of all n=1206 3T-HCP participants not corrected for physiological noise artefacts (see below); the parcellation was solely evaluated based on known anatomical criteria (contiguous clusters; symmetry across hemispheres; similarity to post-mortem work).

Further details on MIGP for group connectome

MIGP is a computationally tractable method to approximate the group average time series using group-level PCA[7]. The two parameters specifying (a) the number of data-points kept on-line during the iterative computation of the average and (b) the cut-off describing the

number of principal components kept at the end were both set to 4800, corresponding to the number of data points in each individual's file.

Choosing the clustering depth

To evaluate the number of clusters, or in other words, the appropriate depth of the hierarchical clustering, we aimed for a balance between simplicity and detail, as well as anatomical plausibility. We carefully compared the location and size of the clusters obtained at increasing levels of depths to known anatomical subdivisions of the amygdala [8–11]. Given post-mortem work in humans and a large body of work in animals, we expected an anatomically plausible cluster solution to show good symmetry across hemispheres, i.e., to contain corresponding clusters across left and right hemispheres, and to involve spatially contiguous clusters. Another focus was on detail: for instance, it has been suggested that the two largest amygdala nuclei, basal and lateral nuclei, can be further split into several subdivisions [10]. We were also keen to identify the small central nucleus in both hemispheres, given its importance for connecting the amygdala to brainstem regions. The central nucleus split off at depth 10 and 12 of the hierarchical clustering algorithm in the left and right hemisphere respectively, at which point both hemispheres contained 7 clusters (AB/BM and CoN were still connected across hemispheres, so there were five uniquely left clusters, five uniquely right clusters and two clusters that contained both hemispheres, and thus depth 12). This clustering solution was also symmetrical across hemispheres. At the next depths from 13-15, AB/BM split between L and R hemispheres and the ventral part of the lateral nucleus split into two halves first in the right hemisphere and then in the left hemisphere. This was more detail than we would have anticipated or required for interpretation of further analyses. Throughout the results, we therefore focussed on the depth 12 cluster solution, which when merging corresponding clusters in both hemispheres yielded seven final clusters (**Figs 1B and 2A**). Other shallower and deeper clustering depths are shown in **Extended Data Fig 3**.

Replication of amygdala parcellation

Following the procedure described for the original sample of n=200 in the main text, we closely replicated the amygdala parcellation in two additional datasets (3T: n=200; 7T: n=98; **Extended Data Fig 4A**). However, a parcellation using the data from all n=1200 3T-HCP

participants, which had not been corrected for physiological noise, showed less symmetry and anatomical plausibility despite relying on more data. We used the parcellation generated from the first 3T dataset for all further analyses. We also quantified the similarity between the original and the two replication parcellations using two metrics: (a) the mean distance between cluster centroids and (b) the percentage overlap between parcellations (i.e., percentage of voxels with the same label). The overlap between the original parcellation and the 3T replication was 68% and the overlap between the original parcellation and the 7T replication was 54%. For further details and statistical analysis, see **Extended Data Fig 4B,C**).

Naming of clusters

The labelling of clusters was largely based on the Atlas of the Human Brain [10] and a post-mortem parcellation at 7T with 100-150um resolution [11]. Similarities between our labelling and the labels of those two atlases as well as two common nomenclatures used in non-human primates are shown in **Supplementary Table 2**. **Supplementary Table 1** also summarizes the size of all clusters. The most dorsal, posterior and lateral nucleus (dark blue in **Figs 1B and 2A**) which was also the smallest in size (62 voxels across both hemispheres) perfectly matched in its size and position the central amygdaloid nucleus and was therefore labelled **Ce**. Judging from its position and size, it contained both medial and lateral divisions of the central nucleus (CeM and CeL) [10]. However, it is less clear whether it also contained the medial amygdaloid nucleus. The medial amygdaloid nucleus might have been part of this 'Ce' cluster or the adjacent cluster (middle blue in **Figs 1B and 2A**) which was positioned in a dorsal, posterior and medial location where the cortical amygdaloid nuclei are located (e.g. PCo=posterior cortical; ACoV and ACoD = anterior cortical, ventral & dorsal parts [10]; also referred to as CAT = cortico-amygdaloid transition area e.g. in [11]). We therefore labelled this adjacent cluster **CoN**, as an agglomeration of the cortical nuclei of the amygdala. It contained altogether 133 voxels across left and right hemispheres, and possibly comprised cortical nuclei as well as the medial nucleus. Ventral and anterior to the Ce and CoN nuclei, in a medial position within the amygdala (light blue in **Figs 1B and 2A**), we identified a portion of the basal amygdala which very likely contained Mai et al.'s ventral and dorsal basomedial (BMVM and BMDM), and probably also its basolateral intermediate and dorsal subdivision (BLI, BLD), and thus the majority of basomedial and basolateral aspects of the basal nucleus. We therefore refer to it

simply as the basal nucleus **B**. It contained 74 voxels across hemispheres and was adjacent to a slightly more medial subdivision of the basal nucleus which we refer to as auxiliary basal/basomedial (**AB/BM**, green in **Figs 1B and 2A**) and which contained 104 voxels across hemispheres. This cluster, AB/BM, based on its size and location, would have contained the ventromedial part of the basolateral nucleus (BLVM) as well as paralaminar subdivision of the basal nucleus (BLPL) in [10] and closely corresponded to what Saygin and colleagues [11] label AB as well. The remaining three clusters made up the lateral nucleus of the amygdala, namely its dorsal, intermediate and ventral portion (**LaD**, **LaI**, **LaV**, respectively, in red, yellow and dark red in **Figs 1B and 2A**). These clusters contained 84, 86 and 104 voxels, respectively. Because the ventral part LaV extended more medially, it likely contained parts or all of the basolateral intermediate (BLI) nucleus (or Ba/PL in [11]) and is therefore referred to as **LaV/BL**. Overall, there was good correspondence between the position and labelling of our nuclei, and those reported in [8–11] (**Supplementary Table 2**).

ROI selection

We had several *a priori* regions of interest which were informed by prior work, including anatomical work using tracers in macaque monkeys as well as work in humans with mental health disorders. Our ROI selection was thus entirely hypothesis-driven and followed two key criteria: (1) knowledge about a region's mono- or disynaptic connectivity with the amygdala and (2) knowledge about a region's importance in mental health/mood disorders. Most of our ROIs fulfilled both criteria. The majority of the ROIs we included are known to have monosynaptic connections with the amygdala (or in the case of some brainstem and dlPFC regions, di-synaptic connections that are made via the hypothalamus or medial PFC, respectively), and most regions with strong amygdala connectivity were included. We left out some regions with monosynaptic connections to the amygdala – for example anterior temporal lobe regions – because we did not have strong reasons to believe they were relevant for the mental health-related behaviours of interest in this study. Other notable omissions are the hypothalamus and hippocampus which we would have needed to parcellate and which it would not have made sense to include as whole structures. All our ROIs are illustrated in Fig **2C** and will be motivated one by one below. For an overview, all ROIs with

their respective sizes are listed in **Supplementary Table 1**. ROI selection was performed prior to and independent of subsequent analyses focusing on individual differences.

We included aspects of dlPFC despite it not having strong monosynaptic connections with the amygdala in monkeys, because of work involving neurostimulation to dlPFC, most commonly repetitive transcranial magnetic stimulation (rTMS), which has been shown to alleviate symptoms of mental health disorders, particularly depression [12–14] and because it has been implicated in the regulation of amygdala responses to threat [15]. The location of stimulation over dlPFC can be variable across studies but is most common over areas 9/46 and 46 and particularly effective when strong connectivity with dlPFC and area 25 is observed [16]. We therefore included all sub-clusters of areas 46 and 9/46 reported in HCP's multi-model parcellation version 1.0 which included 46 (316 vertices), 9-46d (379 vertices), a9-46v (147 vertices), and p9-46v (214 vertices).

On the medial and orbital surface, amygdala connectivity gradually changes along an anterior-posterior axis, with strongest connectivity posteriorly closest to the corpus callosum [17,18]. This also mimics the transition between agranular and dysgranular/granular cortex, and unimodal to transmodal connectivity [19,20]. We included all agranular regions in the medial and orbital prefrontal cortex; all the likely homologues of these areas have strong monosynaptic connectivity with the amygdala in monkeys. This included areas 32, 25, 24 and the most posterior part of OFC. We also included granular area 9m, adjacent do area 32 in medial frontal cortex, and frontal operculum, which has also been highlighted in tracer studies for its connections with the amygdala. As above, we took the parcels obtained from HCP's multi-model parcellation version 1.0 [21] which are labelled areas 25 (54 vertices), a24 (89 vertices), p24 (66 vertices), a24pr (75 vertices), s32 (55 vertices), p32 (122 vertices), d32 (147 vertices), a32pr (163 vertices), p32pr (190 vertices), 9m (408 vertices) and pOFC (83 vertices). Frontal operculum contains FOP1-FOP5 (with 61, 101, 83, 240 and 193 voxels, respectively). Apart from their strong connectivity with the amygdala many of these regions have indeed been implicated in mood disorders and social cognition. For example, PET work shows abnormal metabolism in subgenual PFC, including area 25, ventral 24 and possibly 32 in depression[22], deep brain stimulation in subgenual regions or adjacent white matter fibres can alleviate symptoms of depression [23–25], and negative biases in decision-making can be induced by stimulating pregenual cortex (rostral area 24 and dorsal area 32; [26]). In addition, several

studies have also highlighted the importance of peri-/pregenual cortex in social cognition [27,28]. In summary, we included four cortical regions on the lateral, 10 cortical regions on the medial, one cortical region on the orbital surface, and five frontal opercular regions, and thus a total of 20 cortical ROIs.

Subcortically, our major focus was on the key nuclei associated with different neurotransmitter systems because of their importance for mental well-being. We included the substantia nigra (SN), which contains many dopaminergic neurons and the nucleus accumbens (NAc), an area receiving strong dopaminergic innervation [29]. DA has been implicated in mental health disorders; for example, Parkinson's disease, which is characterized by a loss of DA neurons in SN, leads to depression in a large percentage of patients (~35%; [30]). But DA also plays a key role in reward-learning and sleep regulation. Striatal dysfunction has, for instance, been associated with sleep disturbances and a subset of NAc core neurons was found to regulate slow-wave sleep [31,32]. The SN mask was taken from the NITRC Atlas of the basal ganglia [33] and contained 134 voxels. The NAc was taken from the Harvard Subcortical Atlas and contained 188 voxels.

The bed nucleus of the stria terminalis (BNST) was included because of its role in mediating the long-term effects of anxiety and responses to stress [34]. It is also sometimes considered part of the extended amygdala. The BNST mask was obtained from [35] and contained 45 voxels.

Two regions with opposing functionality within the periaqueductal grey (PAG) were included because of their importance in regulating autonomic arousal: ventrolateral PAG (vlPAG) which mediates rest- and digest-related behaviour and dorsal PAG (dPAG) which mediates fight and flight responses. The masks for these regions were taken from [36]. The dPAG was the summation of Faull et al.'s dorsomedial (dm) and dorsolateral (dl) aspects of PAG; vlPAG contained 43 voxels and dPAG 45 voxels.

The role of serotonin and of selective serotonin reuptake inhibitors (SSRI) in the pathology and treatment of mental health disorders is well known. The raphe nuclei are the most important source of serotonin in the brain. Masks for dorsal and median raphe nuclei were taken from the Harvard Ascending Arousal Network Atlas [37]. The dorsal raphe nucleus (RN_DR) contained 23 voxels, and the median raphe nucleus (RN_MR) contained 8 voxels. Finally, locus coeruleus (LC), the main site of noradrenaline production was defined based on

[38] and contained 20 voxels. Probabilistic masks were binarized first, including all voxels with probability >.25, in other words, voxels that had a larger than 25% chance of being within the given region (NAc, SN). Binary files and all masks we received in binary format (BNST, PAG, LC, RN) were subsampled to 2mm. While most regions were then simply binarized again using any voxels >.25 in subsampled space, we carefully and visually inspected all subcortical ROIs individually at 2mm to try and achieve the best tracing of their shape, volume and position when compared with a higher-resolution versions from published atlases. However, this meant that we did not apply the exact same threshold for each ROI; we deviated from our standardized procedure in three cases to improve the masks: nucleus accumbens, dorsal and median raphe. In these three cases, slightly different thresholds provided more anatomically plausible ROIs. NAc thresholded at .25 would have yielded an unusually large ROI, so a threshold of >.35 was applied in the second step; for the raphe nuclei, thresholds were adjusted manually to maximise anatomical plausibility (>.6 and >.69N for dorsal and median, respectively). Similarly, in some cases, a few voxels were added or removed manually from the ROI to achieve better tracing of the original higher resolution mask. Importantly, all of these mask optimization steps were performed based on anatomical criteria and prior to any subsequent analyses. All ROIs are shown in Figure 2C and published as part of the OSF repository https://doi.org/10.17605/OSF.IO/EGM2R.

Included behavioural scores

We included 33 behaviours composed of

(1) Measures from the NIH Toolbox Emotion Battery (www.nihtoolbox.org) [39,40](total: 17); each item was administered on a 5-point scale with options ranging from “not at all” to “very much”. In all cases a-d below, scores <40 are considered low and scores >60 are considered high.
    a. Six measures from the Negative Affect toolbox (Anger Affect: obtained using computer-adaptive testing (CAT), Anger Hostility: obtained from a questionnaire with 5 items, Anger Aggression: also 5 items, Fear Affect: CAT, Fear Somatic: 6 items, Sadness: CAT);
    b. Three measures from the Psychological Well-Being toolbox (Life Satisfaction, Mean Purpose, Positive Affect) – all obtained using CAT

c. Six measures from the Social Relationships toolbox (Friendship, Loneliness, Perceived Hostility, Perceived Rejection, Emotional Support, Instrumental Support); loneliness obtained from a questionnaire containing 5 items, all others from questionnaires containing 8 items.
d. Two measures from the Stress and Self Efficacy toolbox (Perceived Stress: 10 items, Self-Efficacy: CAT)

(2) Measures from the Pittsburgh Sleep Questionnaire [41] (total 9) composed of minutes to fall asleep (past month); hours of sleep per night (past month); sleep trouble: can't go to sleep within 30 minutes; sleep trouble: wake-up in middle of night or early morning; sleep trouble: had bad dreams; overall sleep quality; how often taken sleep medicine; how often trouble staying awake during the day; how often trouble keeping up enthusiasm during the day. All of these were rated on a scale from 0-9.

(3) Measures from the five-factor model [42] a) neuroticism; b) extroversion/introversion; c) agreeableness; d) openness; and e) conscientiousness [43]. HCP data collection administered the 60-item version of the Costa and McRae Neuroticism/Extroversion/Openness Five Factor Inventory (NEO-FFI), which has good reliability and validity [43]. This measure was obtained as part of the Penn Computerized Cognitive Battery [44].

(4) Measures from the Penn Emotion Recognition Test, again obtained as part of the Penn Emotion Recognition Test. During this test, participants are presented with 40 faces and need to identify the emotion of the face from the five options happy, sad, angry, scared and no feeling. There are eight faces in each category. We included a) the number of Correct Anger Identifications (ER40ANG) ranging from 0-8 and b) the number of Correct Fear Identifications (ER40FEAR) ranging from 0-8.

Justification and replication of factor analysis

A factor analysis, rather than principal component analysis (PCA), is especially appropriate when the aim, as in the current study, is data reduction with maximal interpretability that is useful for establishing latent causes as opposed to data reduction into components explaining maximal variance[45–48]. A PCA would provide a data reduction of the behavioural scores that

aims to obtain orthogonal principal components which explain maximal variance. By contrast, rather than merely reducing the data, here we were interested in a model of the latent variables that might have generated the behavioural scores and which explains interrelationships among them, taking into account their unique and shared variance and offering maximal interpretability [45–47]. A factor analysis therefore was the most appropriate choice of method.

The key factor analysis used throughout the study was performed on n=200 3T participants, but the number of factors (k=4) was extracted from an initial inspection of n=100 participants. Importantly, the same four factors replicated in our full first dataset of n=200 3T-HCP participants and inspection of a potential fifth factor showed lack of interpretability and would have introduced a high correlation between two of the factors (r=.5; compared to highest correlation in our set of four: .35). Moreover, the factor analysis replicated to several other datasets (**Extended Data Fig 7**), most importantly to the full set of all n=1206 3T HCP participants (Pearson's correlation between factor loadings n=200 vs n=1206 3T participants: r=.94, p=1.5e-16, r=.93, p=1e-14, r=.97, p=9.6e-10, r=.9, p=1.3e-12), but also the n=400 3T and the n=98 7T participants included for neural analyses (**Extended Data Fig 7**). Note that a factor analysis performed after removing the two Penn Emotion Recognition scores (which did not contribute much to any of the four factors) resulted in identical factor loadings (all Pearson's r=1; p<1e=35).

Null distribution for smallest and peak prediction

To establish whether (a) the size of the smallest significant network and (b) Pearson's r at the peak prediction across all 196 models were significant, we generated two null distributions using permutation. As before, predictions of the 7T behavioural scores were based on increasing number of edges, using the order and weight of edges extracted from robust regressions on 3T participants, but this procedure was now repeated n=10,000 times using shuffled behavioural scores. For each iteration, we determined the size of the smallest network that reached significance (r>0.168) for a given behavioural dimension and averaged this number across the four behaviours to build the first null distribution (if no significant predictions were achieved for a given iteration and behaviour, we used a score of 197, i.e. the maximum number of edges plus 1). In each iteration, we also determined the maximum

Pearson's r achieved for each behaviour and averaged these four maximum r values across behaviours to build the null distribution. This ensured in both cases that the number of tests (n=196) and the number of behavioural dimensions (n=4) was accounted for. One p-value was then extracted from each of these two null distributions.

To examine if our predictions of 7T mental health dimensions benefited from using the top edges identified in the 3T data, i.e., to establish whether the order of inclusion of the edges mattered, we generated a further null distribution: we included the same number of edges (from 1 to 196), but this time they were randomly chosen from the full set of 196 rather than sorted in order of importance based on their absolute regression coefficient. This procedure was repeated 10,000 times at each step between 1 and 196. The resulting null distribution (black line in **Fig5A**) shows, at each step, how much adding any connection out of the original set of 196 is expected to help the prediction. We report the percentage of models out of the first 70 and all 196 where the 3T-informed order of edges produced better predictions than when edges were included in random order (i.e., the difference between the coloured bars and the black null distribution shown in Fig5A).

Controlling for amygdala parcellation

To show that parcellating the amygdala yielded improvements in prediction accuracy, we also repeated the regression procedure with only the connections from our ROIs to the entire amygdala instead of all individual nuclei (a total of 28 possible predictors). Again, robust regression coefficients obtained from fitting the 3T participants were applied to the 7T participants' functional connectivity values to predict the 7T participants' behavioural scores (**Fig 6A**). However, instead of using 196 FC predictors, predictions were derived from only 28 FC values. Unlike in other situations, however, where a larger number of predictors is expected to perform better, here, the predictors captured the same information in both cases, namely functional connectivity with the total of all amygdala voxels. Subdividing the amygdala into smaller groups of voxels could have two potential effects: if subdivisions are meaningless or BOLD measurements in smaller sets of voxels too noisy, predictions generated from individual nuclei functional connectivity might be worse than those obtained from more robust whole-amygdala functional connectivity, despite including more predictors; alternatively, if amygdala subdivisions into nuclei are meaningful, this might increase the

401 prediction accuracy of the nuclei-version over and above the whole-amygdala version for
402 predicting mental health dimensions.

403 Nevertheless, to make predictions obtained from whole vs nuclei-specific amygdala
404 functional connectivity more comparable, we generated null distributions in both cases. We
405 shuffled the order of participants' behavioural scores in both 3T and 7T datasets and repeated
406 the above procedure of predicting 7T-behavioural scores from 3T-weights 10,000 times. The
407 p-values extracted from the respective null distributions account for the number of
408 predictors, allowing a direct comparison (**Fig 6B**). Thus, this test established if parcellating the
409 amygdala into nuclei helped us improve the accuracy of our predictions.

410

Supplementary Table 1: Vertex/voxel size of *a priori* ROIs and amygdala nuclei

ROI	**Size**
Subcortex	**Voxels**
SN	134
NAc	188
BNST	45
vlPAG	43
dPAG	45
RN_DR	23
RN_MR	8
LC	20
Cortex	**Vertices**
25	54
a24	89
p24	66
a24pr	75
s32	55
p32	122
d32	147
a32pr	163
p32pr	190
9m	408
pOFC	83
FOP1	61
FOP2	101
FOP3	83
FOP4	240
FOP5	193
46	316
9-46d	379
a9-47v	147
p9-46v	214
Amygdala nuclei	**Voxels**
Ce	62
CoN	133
BaL	74
AB/BM	104
LaI	84
LaD	86
LaV/BL	104

Supplementary Table 2: Correspondence of nuclei labels across different laboratories and atlases

This study (human)	**Mai et al. (human)**	**Saygin et al (human)**	**Allen Brain Institute (34yo-human)**	**Fudge et al (macaque)**	**Barbas et al (macaque)**
CoN	PCo/ACoV/ACoD	CAT/Co	CoPv/CoPd	PAC/Me	Co (AAA/Me)
Ce	CeM/CeL	Ce/Me	CEm/CEl	CeN	Ce
B	BLI/BLD/BMVM/BMDM	Ba	BLD/BMD/BLI/BMV	Bi/Bmc	BL
AB/BM	BM/BLVM/BLPL	AB	AHA/BMV	ABmc/ABpc	BM
LaD	LaD (L/A/M)	La	LaD	L	La
LaI	LaI	La	LaI	L	La
LaV/BL	LaV/BLI	La/Ba/PL	LaV/BLVl	L/Bpc	La/BL

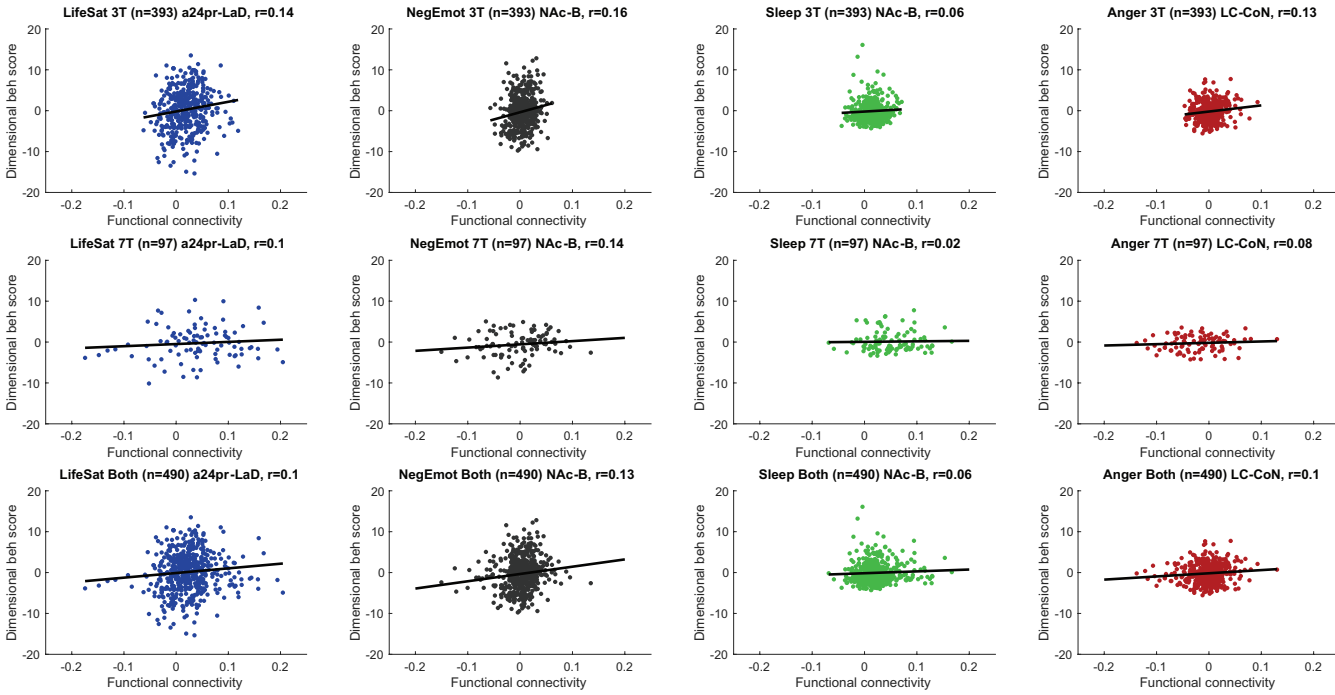
Supplementary Table 3, Demographics of participant cohorts

	Original n=200 3T	Replication n=200 3T	Replication n=98 7T	All n=1206
Mean age	28.95	28.28	29.42	28.84
Standard error age	0.26	0.28	0.33	0.11
Max age	36	36	36	37
Min age	22	22	23	22
Number of females	108	99	59	656
Number of males	92	101	39	550
DSM - mean	4.46	3.96	3.43	4.25
ASR mean	37.90	35.72	31.79	37.43
DSM SD	16.64	10.70	5.73	12.24
ASR SD	669.08	481.67	253.43	523.82
Employment status *	1.53	1.50	1.57	1.52
Income	4.92	5.01	5.06	5.00
Education	14.98	15.09	14.89	14.86
Cognitive status (MMSE score)	29.1	29.03	28.98	28.98
Fluid intelligence PMAT ****	16.56	16.76	17.45	16.65

* not working = 0, part-time employment = 1; full-time employment = 2
** Total household income: <$10,000 = 1,10K-19,999 = 2, 20K-29,999 = 3,30K-39,999 = 4, 40K-49,999 = 5, 50K-74,999 = 6, 75K-99,999 = 7, >=100,000 = 8
*** Years of education completed: <11 = 11; 12; 13; 14; 15; 16; 17+ = 17
**** Penn Progressive Matrices: Number of Correct Responses

Supplementary Figure 1

Relationship between single edge and behavioural scores illustrates properties of 3T and 7T datasets



Supplementary Figure 1, Properties of 3T and 7T datasets We used the maximum number of participants from the full 3T HCP cohort for which high quality physiological noise recordings were available, and all remaining non-overlapping 7T participants. The resulting data had a better spread in behavioural dimensions in the 3T dataset (variance on the y axis, 1st row), but a better spread in functional connectivity in the 7T dataset (variance on the x axis, probably due to improved signals in subcortical regions; 2nd row). Individual connections relevant for each behaviour were chosen purely for illustration of the scatterplots here.

Supplementary References

1. O'Reilly, J. X. *et al.* Causal effect of disconnection lesions on interhemispheric functional connectivity in rhesus monkeys. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13982–13987 (2013).
2. Kubik, S. & Ono, M. Atlas of the Cerebral Sulci by Stefan Kubik, M. Ono | Waterstones. https://www.waterstones.com/book/atlas-of-the-cerebral-sulci/stefan-kubik/m-ono/9783137321019 (1990).
3. Petrides, M. Atlas of the Morphology of the Human Cerebral Cortex on the Average MNI Brain - 1st Edition. https://www.elsevier.com/books/atlas-of-the-morphology-of-the-human-cerebral-cortex-on-the-average-mni-brain/petrides/978-0-12-800932-1 (2018).
4. Parkes, L. *et al.* Transdiagnostic dimensions of psychopathology explain individuals' unique deviations from normative neurodevelopment in brain structure. *Transl. Psychiatry* **11**, 232 (2021).
5. Tavor, I. *et al.* Task-free MRI predicts individual differences in brain activity during task performance. *Science* **352**, 216–220 (2016).
6. Harrison, O. K., Guell, X., Klein-Flügge, M. C. & Barry, R. L. Structural and resting state functional connectivity beyond the cortex. *NeuroImage* **240**, 118379 (2021).
7. Smith, S. M., Hyvärinen, A., Varoquaux, G., Miller, K. L. & Beckmann, C. F. Group-PCA for very large fMRI datasets. *Neuroimage* **101**, 738–749 (2014).
8. Fudge, J. L., Kunishio, K., Walsh, P., Richard, C. & Haber, S. N. Amygdaloid projections to ventromedial striatal subterritories in the primate. *Neuroscience* **110**, 257–275 (2002).

9. Zikopoulos, B., John, Y. J., García-Cabezas, M. Á., Bunce, J. G. & Barbas, H. The intercalated nuclear complex of the primate amygdala. *Neuroscience* **330**, 267–290 (2016).
10. Mai, J. K., Majtanik, M. & Paxinos, G. *Atlas of the Human Brain*. (Academic Press, 2015).
11. Saygin, Z. M. *et al.* High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: manual segmentation to automatic atlas. *NeuroImage* **155**, 370–382 (2017).
12. Dichter, G. S., Gibbs, D. & Smoski, M. J. A systematic review of relations between resting-state functional-MRI and treatment response in major depressive disorder. *J. Affect. Disord.* **172**, 8–17 (2015).
13. Salomons, T. V. *et al.* Resting-state cortico-thalamic-striatal connectivity predicts response to dorsomedial prefrontal rTMS in major depressive disorder. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **39**, 488–498 (2014).
14. Furtado, C. P. *et al.* An investigation of medial temporal lobe changes and cognition following antidepressant response: a prospective rTMS study. *Brain Stimulat.* **6**, 346–354 (2013).
15. Ironside, M. *et al.* Effect of Prefrontal Cortex Stimulation on Regulation of Amygdala Response to Threat in Individuals With Trait Anxiety: A Randomized Clinical Trial. *JAMA Psychiatry* (2018) doi:10.1001/jamapsychiatry.2018.2172.
16. Fox, M. D., Buckner, R. L., White, M. P., Greicius, M. D. & Pascual-Leone, A. Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. *Biol. Psychiatry* **72**, 595–603 (2012).

17. Aggleton, J. P., Burton, M. J. & Passingham, R. E. Cortical and subcortical afferents to the amygdala of the rhesus monkey (Macaca mulatta). *Brain Res.* **190**, 347–368 (1980).
18. Amaral, D. G., Price, J. L., Pitkanen, A. & Carmichael, S. T. Anatomical organization of the primate amygdaloid complex. in *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction* 1–66 (New York: Wiley-Liss, 1992).
19. Huntenburg, J. M., Bazin, P.-L. & Margulies, D. S. Large-Scale Gradients in Human Cortical Organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).
20. Margulies, D. S. *et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12574–12579 (2016).
21. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
22. Drevets, W. C. *et al.* Subgenual prefrontal cortex abnormalities in mood disorders. *Nature* **386**, 824–827 (1997).
23. Dandekar, M. P., Fenoy, A. J., Carvalho, A. F., Soares, J. C. & Quevedo, J. Deep brain stimulation for treatment-resistant depression: an integrative review of preclinical and clinical findings and translational implications. *Mol. Psychiatry* **23**, 1094–1112 (2018).
24. Riva-Posse, P. *et al.* A connectomic approach for subcallosal cingulate deep brain stimulation surgery: prospective targeting in treatment-resistant depression. *Mol. Psychiatry* **23**, 843–849 (2018).

25. Holtzheimer, P. E. *et al.* Subcallosal cingulate deep brain stimulation for treatment-resistant depression: a multisite, randomised, sham-controlled trial. *Lancet Psychiatry* **4**, 839–849 (2017).
26. Amemori, K. & Graybiel, A. M. Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat. Neurosci.* **15**, 776–785 (2012).
27. Apps, M. A. J., Rushworth, M. F. S. & Chang, S. W. C. The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron* **90**, 692–707 (2016).
28. Wittmann, M. K., Lockwood, P. L. & Rushworth, M. F. S. Neural Mechanisms of Social Cognition in Primates. *Annu. Rev. Neurosci.* **41**, 99–118 (2018).
29. Averbeck, B. B. & Costa, V. D. Motivational neural circuits underlying reinforcement learning. *Nat. Neurosci.* **20**, 505–512 (2017).
30. Aarsland, D., Påhlhagen, S., Ballard, C. G., Ehrt, U. & Svenningsson, P. Depression in Parkinson disease--epidemiology, mechanisms and management. *Nat. Rev. Neurol.* **8**, 35–47 (2011).
31. Brown, R. E., Basheer, R., McKenna, J. T., Strecker, R. E. & McCarley, R. W. Control of Sleep and Wakefulness. *Physiol. Rev.* **92**, 1087–1187 (2012).
32. Oishi, Y. *et al.* Slow-wave sleep is controlled by a subset of nucleus accumbens core neurons in mice. *Nat. Commun.* **8**, 1–12 (2017).
33. Keuken, M. C. *et al.* Quantifying inter-individual anatomical variability in the subcortex using 7 T structural MRI. *NeuroImage* **94**, 40–46 (2014).
34. Lebow, M. A. & Chen, A. Overshadowed by the amygdala: the bed nucleus of the stria terminalis emerges as key to psychiatric disorders. *Mol. Psychiatry* **21**, 450–463 (2016).

35. Folloni, D. *et al.* Dichotomous organization of amygdala/temporal-prefrontal bundles in both humans and monkeys. *eLife* **8**, e47175 (2019).
36. Faull, O. K., Jenkinson, M., Ezra, M. & Pattinson, K. T. Conditioned respiratory threat in the subdivisions of the human periaqueductal gray. *eLife* **5**, (2016).
37. Edlow, B. L. *et al.* Neuroanatomic connectivity of the human ascending arousal system critical to consciousness and its disorders. *J. Neuropathol. Exp. Neurol.* **71**, 531–546 (2012).
38. Betts, M. J., Cardenas-Blanco, A., Kanowski, M., Jessen, F. & Düzel, E. In vivo MRI assessment of the human locus coeruleus along its rostrocaudal extent in young and older adults. *NeuroImage* **163**, 150–159 (2017).
39. Gershon, R. C. *et al.* NIH Toolbox for Assessment of Neurological and Behavioral Function. *Neurology* **80**, S2–S6 (2013).
40. Salsman, J. M. *et al.* Emotion assessment using the NIH Toolbox. *Neurology* **80**, S76–S86 (2013).
41. Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R. & Kupfer, D. J. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* **28**, 193–213 (1989).
42. Heine, S. J. & Buchtel, E. E. Personality: the universal and the culturally specific. *Annu. Rev. Psychol.* **60**, 369–394 (2009).
43. McCrae, R. R. & Costa Jr., P. T. A contemplated revision of the NEO Five-Factor Inventory. *Personal. Individ. Differ.* **36**, 587–596 (2004).
44. Gur, R. C. *et al.* A cognitive neuroscience based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *J. Neurosci. Methods* **187**, 254–262 (2010).

45. Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A. & Daw, N. D. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* **5**, e11305 (2016).
46. Rouault, M., Seow, T., Gillan, C. M. & Fleming, S. M. Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biol. Psychiatry* **84**, 443–451 (2018).
47. Moutoussis, M. *et al.* Decision-making ability, psychopathology, and brain connectivity. *Neuron* (2021) doi:10.1016/j.neuron.2021.04.019.
48. Scholl, J., Trier, H. A., Rushworth, M. F. S. & Kolling, N. The effect of apathy and compulsivity on planning and stopping in sequential decision-making. *PLOS Biol.* **20**, e3001566 (2022).