Received: 11 August 2021 | Revised: 4 November 2021 | Accepted: 13 December 2021

DOI: 10.1111/psyp.14000

PSYCHOPHYSIOLOGY SPR WILEY

ORIGINAL ARTICLE

fMRI replicability during emotional scene viewing: Functional regions and sample size

Nicola Sambuco

Center for the Study of Emotion and Attention, University of Florida, Gainesville, Florida, USA

Correspondence
Nicola Sambuco, Center for the Study of Emotion and Attention, University of Florida, PO Box 112766, Gainesville, FL, USA.
Email: nsambuco@ufl.edu

Funding information
National Institute of Mental Health, Grant/Award Number: MH094386 and MH098078

Abstract

Recent findings have questioned the replicability of functional magnetic resonance imaging (fMRI) in the study of affective processing, reporting low replicability of emotional enhancement during a face-matching task. However, poor replicability may instead reflect a lack of emotional engagement for face matching. In the current study, replicability of emotional enhancement was tested in a large ($N = 160$) sample when emotional engagement was assessed during pleasant, neutral, and unpleasant picture viewing, which reliably engages affective reactions in both the brain and the body. Replicability was computed using a subsampling technique, in which random sets of subjects of different sample sizes ($N = 20, 40, 60, 80$) were selected from the entire dataset, and replicability of emotional enhancement for peaks, clusters, and voxels were averaged across 500 permutations for each sample size. Consistent with previous findings, fMRI replicability increased with increasing sample size. On the other hand, even with relatively small samples, fMRI replicability for peaks, clusters, and voxels during emotional, compared to neutral, scene viewing was good to excellent. Importantly, replicability varied in different brain regions, with excellent replicability at both the cluster and peak level with an N of 40, at the most conservative threshold ($p < .001$), in the amygdala and the visual cortex. The data argue against general recommendations regarding sample size in fMRI studies of emotion, suggesting instead that degree of replicability depends on successful emotional engagement in task-related brain regions.

KEYWORDS

amygdala, emotion, fMRI, pictures, replicability, visual cortex

1 | INTRODUCTION

Affective dysfunction is common across psychiatric disorders, and fMRI studies of emotional processing, although expensive, promise to elucidate the contributing neural correlates. Recent studies, however, have questioned the clinical utility of functional brain imaging data of emotional processing, citing low replicability. For example, Turner et al. (2018) analyzed data from a face-matching task (Hariri et al., 2005; included in the Human Connectome Project), in which functional activity when matching emotional facial expressions to a target was compared to matching simple geometric shapes. Overall, emotional enhancement showed poor replicability for typical sample sizes in fMRI studies (e.g., $N = 16$ to 36), with much larger samples ($N > 100$ and higher) required to obtain somewhat replicable effects.

Although discriminating affective expressions may be important for social communication, numerous studies have reported low affective engagement in neural, physiological, and subjective reactions when viewing facial expressions of emotion, compared to when viewing naturalistic emotional scenes (e.g., Bradley et al., 2003; Britton et al., 2006; Wangelin et al., 2012), suggesting that poor replicability during face-matching may reflect a lack of emotional engagement. In the current study, following the methods of Turner et al. (2018), Turner et al. (2019) and Nee (2019), replicability of emotional processing in a large sample ($n = 160$) was assessed when emotional engagement was induced by viewing pleasant, neutral, and unpleasant pictures.

Multiple studies concur that viewing emotional (pleasant or unpleasant), compared to neutral, scenes prompt enhanced blood-oxygen-level-dependent (BOLD) activity in a variety of brain regions, including the amygdala, thalamus, inferotemporal cortex (ventral visual cortex), dorsal visual cortex, parietal cortex, inferior frontal gyrus (including part of the insula), lateral orbitofrontal cortex, and medial prefrontal cortex (e.g., Frank & Sabatinelli, 2014; Sabatinelli et al., 2011; Sambuco, Bradley, Herring, & Lang, 2020; Sambuco, Bradley, Herring, Hillbrandt, et al., 2020). Effects in these emotion perception regions are similar when pictures are presented rapidly (e.g., 167 ms; Junghöfer et al., 2006), or presented in an emotional (reward, threat of shock) or a neutral context (Padmala et al., 2017, 2019; Sambuco, Costa, et al., 2020; Somerville et al., 2013). Moreover, viewing emotional, compared to neutral, pictures elicits larger skin conductance responses, greater pupil dilation, more pronounced cardiac deceleration, greater alpha EEG reduction, and a larger late positive potential (LPP), compared to viewing neutral scenes (Bradley et al., 2001, 2017; Cuthbert et al., 2000; De Cesarei & Codispoti, 2011; Weinberg & Hajcak, 2010), with similar LPP modulation found when scenes are presented as briefly as 25-ms (unmasked) or repeated up to 60 times (Codispoti et al., 2006, 2009). Taken together, affective scene viewing reliably elicits a broad range of emotional reactions in both the brain and body.

The current study assesses the replicability of BOLD enhancement during emotional scene viewing by combining data from two recently published studies (Sambuco, Bradley, Herring, & Lang, 2020; Sambuco, Bradley, Herring, Hillbrandt, et al., 2020), in which a total of 160 subjects viewed emotional (pleasant or unpleasant) and neutral scenes. Replicability was computed using a subsampling technique, in which random sets of subjects of different sample sizes ($N = 20, 40, 60, 80$) were selected from the entire data set. Standard metrics of replicability utilized in previous fMRI studies (Nee, 2019; Turner et al., 2018) assessed functional activation when viewing emotional, compared to neutral, scenes for peaks, clusters, and voxels, averaged across 500 subsampling permutations for each sample size. Regions of interest were derived from meta-analytic findings from previous emotional perception studies identify reporting reliable BOLD enhancement in the amygdala and visual cortex (Sabatinelli et al., 2011), and additional cortical and subcortical regions in the occipital, temporal, and frontal lobes (Pessoa & Adolphs, 2010). Following recent studies (Bossier et al., 2020; Turner et al., 2018, 2019), mean replicability is computed across regions of interest and is expected to increase as the sample size increases. An implicit assumption underlying cross-region replicability is that replicability of BOLD activity is equivalent in different neural regions, and this was explicitly assessed here by additionally computing within-region replicability effects, with the amygdala and visual cortex expected to have the highest replicability among the regions activated during emotional scene viewing.

2 | METHOD

2.1 | Participants

A final sample of 160 participants (97 female; *M* age = 29.4, *SD* = 12.9) included 61 healthy adults (Sambuco, Bradley, Herring, Hillbrandt, et al., 2020), of whom 30 were students in Introductory Psychology courses at the University of Florida and 31 were participants from the community recruited from flyers and ads, together with 99 patients diagnosed with internalizing disorders (Sambuco, Bradley, Herring, & Lang, 2020).

Patients were selected from 162 individuals who were part of a larger investigation of affective dysfunction across the internalizing spectrum, which recruited men and women between the ages of 18 and 65 years subsequently diagnosed with a range of different internalizing disorders (Table S1). For all patients, a structured interview (Anxiety Disorder Interview Schedule IV; Brown et al., 1994) was administered by a licensed clinician to establish DSM-IV primary and comorbid diagnoses. Findings from previous analyses indicate that patients reporting severe psychopathology showed dampened emotional reactivity in the amygdala (Sambuco, Bradley, Herring, Hillbrandt, et al., 2020), and these patients were not included in the current replicability analyses. Instead, patients showing normal affective modulation in the amygdala were selected for inclusion ($N = 99$, Table S1; quintiles 3, 4, and 5 in Sambuco, Bradley, Herring, Hillbrandt, et al., 2020). Table S2 demonstrates that patients and healthy controls in the current sample showed similar emotional engagement in the amygdala and other functional regions included in the replicability analyses.

2.2 | Emotional visual processing: Rapid serial visual presentation

Pictures were 108 grayscale pictures that included 36 pleasant, 36 neutral, and 36 unpleasant pictures selected from the International Affective Picture System (IAPS; Lang et al., 2008), presented using rapid serial visual presentation (RSVP) in which 18 pictures of the same hedonic content (e.g., pleasant) were presented in a 6-s interval at the rate of 3 per second (i.e., 333 ms each). Each set of 18 scenes was presented three times, with all exemplars presented before the next repetition. A variable ITI (intertrial interval) of 9 or 12 s resulted in a total scan time of approximately five minutes. Twelve different orders were constructed that varied the serial position in which the pictures were presented across the study and the order of specific scenes within each 6-s stream. Additional details, including the IAPS catalog numbers, are described in Sambuco, Bradley, Herring, and Lang (2020).

2.3 | fMRI: Data collection and offline processing

Data were collected in a 3T Philips scanner with a 32-channel head coil. The scanning sequence began with acquiring a 160-slice sagittal scout set using a standard T1-weighted fast-field echo sequence. Functional volumes were 53 3.5-mm coronal slices acquired using a T2*-weighted echo-planar imaging sequence with a 3,000 ms TR, 30 ms TE, 90-degree flip angle, 72 × 72 acquisition matrix, and 180 mm FOV (2.5 × 2.5 in-plane voxel resolution). Offline, the functional data were slice-time adjusted, motion corrected, spatially smoothed (5.0 mm FWHM Gaussian kernel), and converted to percent BOLD signal using the Analysis of Functional Neuroimages software (AFNI, Cox, 1996). Percent BOLD signal change was spatially normalized to a Talairach template and resampled to 2.5-mm isotropic voxel size.

For each of the 160 participants, hemodynamic responses during scene viewing were estimated using a multiple regression model in which BOLD amplitude was convolved with a canonical gamma function using variables of emotional content (pleasant, neutral, and unpleasant) and motion (6). The statistical comparison of interest contrasted beta coefficients during emotional (pleasant and unpleasant), compared to neutral, picture viewing.

2.4 | Regions of interest (ROIs)

Functional regions demonstrating enhanced BOLD activity during emotional or pleasant scene viewing were determined using the data from the healthy participants. As described more completely in Sambuco, Bradley, Herring, Hillbrandt, et al. (2020), a whole-brain ANOVA, using a false discovery rate (FDR) of $p < .05$ with a minimum cluster size of 19 voxels (300 μl), computed using 2.5 mm^3 voxels, assessed the difference during emotional (pleasant and unpleasant) and neutral scene perception. Bilateral clusters (e.g., visual cortex, amygdala, inferior frontal gyrus) were averaged to provide a single estimate for each functional region listed in Tables 1 and S2. Figure 1 illustrates regions showing emotional enhancement during picture viewing, including the amygdala, the visual cortex, the inferior frontal gyrus (IFG), the medio-dorsal nucleus of the thalamus, and part of the posterior thalamus. The large functional cluster in the visual cortex was divided into two subclusters, one ventral (including inferotemporal cortex, inferior occipital cortex, and fusiform gyrus) and one dorsal (including mid and superior occipital cortex), to assess results from the face-matching paradigm in which only the ventral visual cortex demonstrated increased activation when viewing emotional faces compared to shapes (Elliott et al., 2020).

To identify regions showing valence-related activation, a whole-brain ANOVA, using a false discovery rate (FDR) of $p < .05$ with a minimum cluster size of 19 voxels (300 μl), computed using 2.5 mm^3 voxels, assessed the difference during pleasant and unpleasant scene perception (Sambuco, Bradley, Herring, Hillbrandt, et al., 2020). Replicating previous studies (Sabatinelli et al., 2007; Sescousse et al., 2013), two clusters demonstrated enhanced BOLD activity when viewing pleasant, compared to unpleasant, scenes, including the ventromedial prefrontal cortex (vmPFC) and the ventral striatum (Sambuco, Bradley, Herring, Hillbrandt, et al., 2020). Replicability analyses for this contrast are reported in the Supplementary material. Increased BOLD activity during unpleasant, compared to pleasant, scene viewing was not found in any region.

2.5 | Replicability analyses

Three separate indices of replicability were computed for peaks, clusters, and voxels (Nee, 2019; Turner et al., 2018, 2019), first averaged across all functional brain regions, replicating the methodology used in previous studies (Nee, 2019; Turner et al., 2018, 2019), and then assessed separately for each functional region. Permutation analyses randomly split the entire sample into two different independent groups of sample size 20, 40, 60, or 80, computing a replicability measure indexing emotional enhancement effects. Five hundred permutations were conducted for

Region	Cluster size #Voxels	Voxel-wise intensity (Pearson *r*) Sample size (*N*)			
Regions		20	40	60	80
Ventral visual ctx.	1587	0.38	0.55	0.65	0.71
Dorsal visual ctx.	1359	0.34	0.51	0.61	0.68
Parietal ctx.	746	0.19	0.33	0.44	0.51
Cerebellum C8	51	0.04	0.07	0.10	0.15
Supramarginal gyrus	142	0.21	0.34	0.45	0.54
Thalamus (posterior)	47	0.29	0.50	0.64	0.72
Temporal Pole	73	−0.01	0.01	0.07	0.11
Precentral gyr.	28	0.04	0.11	0.18	0.24
Amygdala	197	0.36	0.53	0.63	0.69
Thalamus (Dorsomedial)	160	0.14	0.24	0.34	0.43
Inferior frontal gyrus	2305	0.14	0.27	0.35	0.42
Pre-SMA	67	0.10	0.14	0.16	0.13
Mid cingulate cortex	184	0.03	0.07	0.09	0.11
Frontal pole	37	0.32	0.52	0.66	0.72

TABLE 1 Voxel-wise replicability, indexed by Person correlations of *t*-statistics (emotional vs. neutral scene viewing), for each functional region as a function of sample size ($N = 20, 40, 60, 80$) during emotional scene-viewing

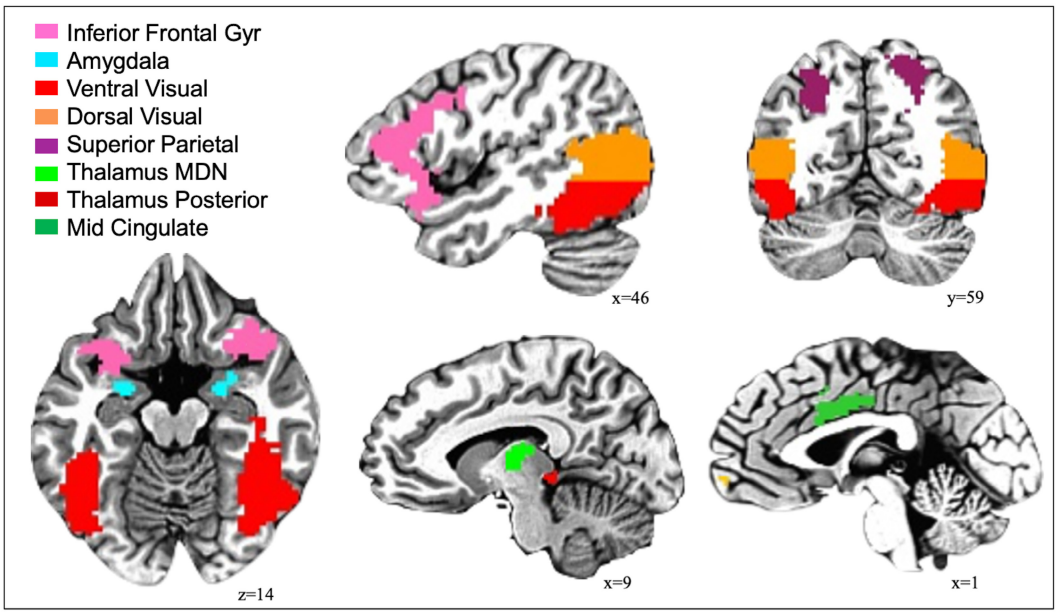

FIGURE 1 Regions showing enhanced functional brain activity when viewing emotional (pleasant or unpleasant), compared to neutral, scenes. Gyr, gyrus; MDN, medio-dorsal nucleus of the thalamus

each sample size, with the final replicability index averaged across all 500 permutations. For peak and cluster replicability indices, which relied on t-statistics comparing emotional to neutral scene viewing, the threshold was additionally varied from lenient to strict (i.e., $p < .05$, $p < .01$, $p < .001$), following recent studies (e.g., Turner et al., 2018).

Following Turner et al. (2018), peak replicability assesses the extent to which the peak voxel activated during emotional, compared to neutral, scene viewing in one map (highest positive *t*-statistic) is active (above threshold, but not necessarily the identical peak) in the comparison map. For each permutation, a value of 1 is assigned if the peak location in each map is above threshold in the comparison map, a value of 0 is assigned if the peak location in each map is not above threshold in the comparison map, and a value of 0.5 is assigned if the peak location in one of the maps is above threshold in the comparison map.

Cluster replicability is computed on binarized active/inactive maps (based on the t-statistic comparing emotional to neutral scene viewing) with the Jaccard statistic computed for each pair of maps. Cluster replicability computes the level of overlap between the pairs of thresholded maps (within each ROI) identified at each replication, with the Jaccard index representing the ratio of the intersection of maps (which corresponds to the number of overlapping voxels in a conjunction map) divided by their union (representing the extent of activation of two maps outside the conjunction), and ranges from 0 (no overlap) to 1 (perfect overlap).

Voxel replicability computes the correlation between the t-statistics for each pair of maps at the individual voxel level. Following Turner et al. (2018), both positive and negative *t*-statistics are included in the computation, with the resulting Pearson correlation coefficient ranging from −1 (opposite *t*-statistics in comparison maps) to 1 (identical t-statistics in comparison maps).

The degree of replicability is reported using criteria for describing the goodness of the intraclass correlation coefficient (ICC; Cicchetti & Sparrow, 1981) with descriptors of poor (<0.4), fair (between 0.4 and 0.6), good (between 0.6 and 0.75), and excellent (>0.75) replicability.

3 | RESULTS

3.1 | Peak replicability

As expected, when averaged across all functional regions, peak replicability of emotional enhancement during scene-viewing increased with increasing sample size (Figure 2, top left). Overall, replicability of BOLD activity when viewing emotional, compared to neutral, scenes was higher for lenient ($p < .05$), compared to more conservative thresholds ($p < .01$ and $p < .001$), with excellent replicability (0.80) even with a relatively small sample ($N = 40$). Figure 2 (top right) illustrates that peak replicability varied across regions, with excellent replicability for emotional enhancement (close to 100%) in the amygdala, inferotemporal cortex, and dorsal visual cortex, even with a relatively small sample size of 20 at thresholds of either $p < .05$ or $p < .01$. For more conservative thresholds, replicability of peak activation reaches excellent levels (above 0.90) in the amygdala and visual cortex with a sample size of 40, while for other regions (e.g., thalamus, parietal cortex) good to excellent levels are reached at a sample size of 40 at the more lenient thresholds ($p < .05$, $p < .01$), with a somewhat larger sample size ($N = 60$ or 80) needed for fair to good replicability at a threshold of $p < .001$.

3.2 | Cluster replicability

When averaged across all functional regions, replicability of emotional enhancement for clusters increased with sample size, as illustrated in Figure 2 (bottom panel, left) and was higher at the most lenient threshold ($p < .05$) compared to more conservative thresholds ($p < .01$, $p < .001$). Overall, replicability was excellent with a sample size of 40, regardless of statistical threshold, in the amygdala, inferotemporal cortex, and dorsal visual cortex, and was also excellent with a sample size of 20 for $p < .05$. Other regions demonstrated more variability, with replicability varying from poor to good based on the sample size and the threshold.

3.3 | Voxel replicability

Across all functional regions, replicability of emotional enhancement for voxel-wise statistics increased with sample size, with voxel replicability (Pearson correlation coefficients) of 0.42, 0.60, 0.69, and 0.74, for *N*'s of 20, 40, 60, and 80, respectively. Voxel-wise replicability of emotional enhancement was highest in the amygdala, the inferotemporal cortex, dorsal visual cortex, posterior thalamus, and the frontal pole (Table 1), in which a sample size of 40 was sufficient to explain about 25% of the variance in the paired maps. In the remaining regions, a larger sample size was needed to achieve fair levels of replicability.

4 | DISCUSSION

Consistent with previous fMRI studies (Bossier et al., 2020; Nee, 2019; Turner et al., 2018, 2019), replicability of functional differences in BOLD activity increased with increasing sample size. However, rather than finding poor replicability of emotional enhancement, fMRI replicability of emotional enhancement during scene viewing was good (between 0.60 to 0.75) to excellent (above 0.75) for peaks, clusters, and voxels, even with relatively small samples (*N* of 40). Importantly, replicability varied in different brain regions, with those previously identified in meta-analyses as more reliably activated during emotional scene-viewing, including the amygdala and the visual cortex (Sabatinelli et al., 2011), prompting better fMRI replicability. Taken together, the data indicate that fMRI replicability of emotional engagement during affective picture-viewing is quite good, compared to the face-matching task, and confirm that fMRI replicability can vary with specific brain regions.

The selected threshold for assessing emotional, compared to neutral, BOLD effects somewhat affected

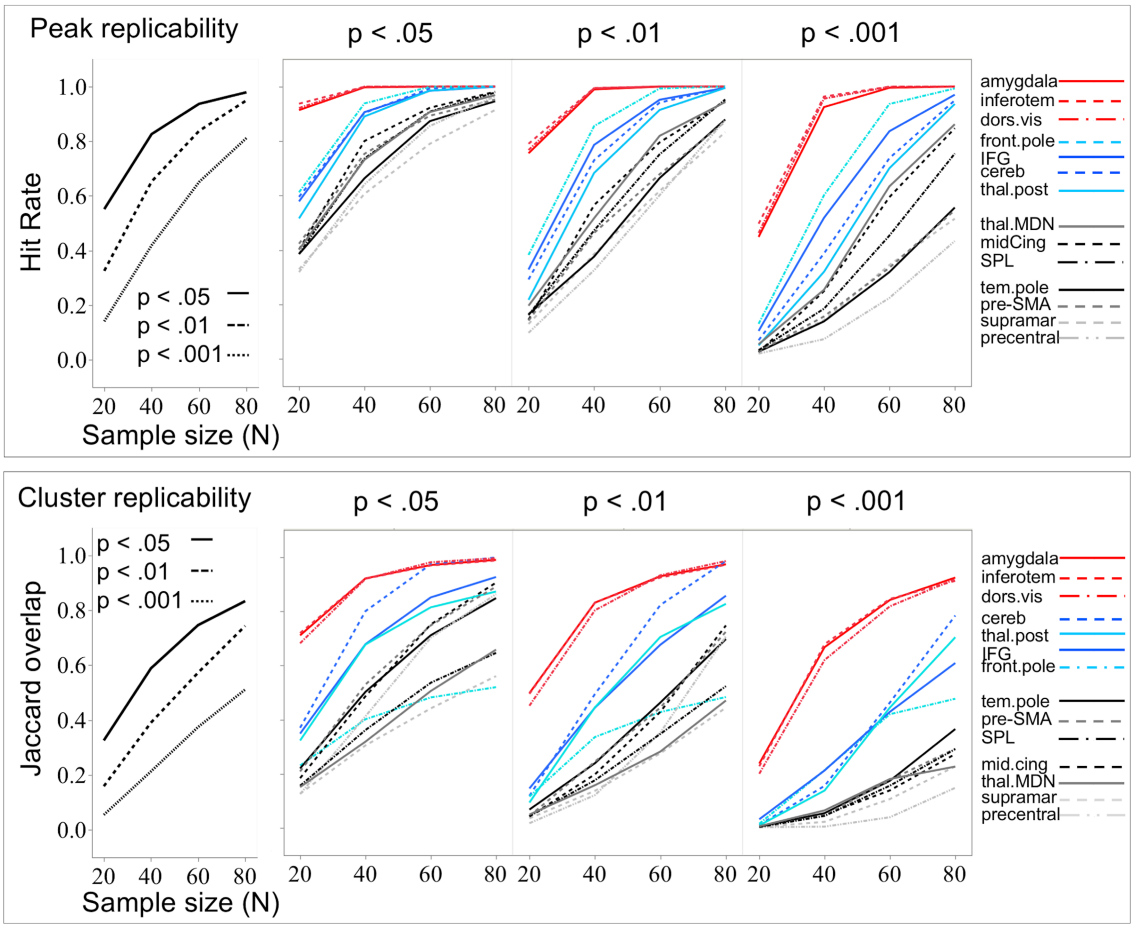

FIGURE 2 *Top panel.* Replicability of peak activation during emotional, compared to neutral, scene viewing, averaged across functional regions (left) and for specific regions thresholded at $p < .05$, .01 or .001. *Bottom Panel.* Replicability of cluster activation during emotional, compared to neutral, scene-viewing, averaged across functional regions (left) and for specific regions thresholded at $p < .05$, .01 or .001. cereb, cerebellum C8; thal.post, posterior thalamus; dors.vis, dorsal visual cortex; front.pole, frontal pole; IFG, inferior frontal gyrus (includes a portion of the anterior insula); inferotem, interotemporal cortex (IT); mid.cing, mid cingulate cortex; precentral, precentral gyrus; SPL, superior parietal lobule; supramar, supramarginal gyrus; tem.pole, temporal pole; thal.DMN, medio-dorsal nucleus of the thalamus

replicability, with better replicability for peaks and clusters using more lenient ($p < .05$) compared to more stringent thresholds ($p < .01$, $p < .001$). This, and similar findings in previous replicability studies (Nee, 2019; Turner et al., 2018, 2019), may be surprising as more liberal thresholds are typically expected to inflate false positives (Eklund et al., 2016; Woo et al., 2014) and decrease replicability. However, if analyses are restricted to regions reliably activated during a specific task (e.g., during emotional, compared neutral, scene viewing), the likelihood of false positives is greatly reduced. Lenient thresholds may be more deleterious for replicability in analyses assessing activation across the entire brain (particularly with small sample sizes) but, as demonstrated here and elsewhere, better replicability can result if critical functional regions are targeted (Cremers et al., 2017; Thirion et al., 2007).

In addition, the replicability crisis in fMRI is sometimes attributed to the over-inclusion of small clusters in fMRI analyses, with the assumption that larger clusters will have better replicability. In the current study, however, differences in replicability were found that did not covary with region size. For example, good to excellent replicability was observed in both large (e.g., dorsal and ventral visual cortex, 1587 or 1359 voxels, respectively) and small regions (e.g., amygdala, ~200 voxels), whereas poorer replicability was observed in the larger inferior frontal gyrus (2305 voxels). Taken together, these findings demonstrate that fMRI replicability is not simply a function of region size. Instead, replicability of BOLD effects was best in regions identified in a classic model of emotional visual processing, developed in non-human primates (Amaral & Price, 1984; Amaral et al., 1992; Freese & Amaral, 2005; Spiegler & Mishkin, 1981) and validated in humans (Frank et al., 2019; Morris et al., 1998; Sabatinelli et al., 2005, 2009, 2014; Vuilleumier et al., 2004) which involves recurrent activation between the amygdala and inferotemporal (part of the ventral visual cortex). Because functional activity in the amygdala is transdiagnostically

modulated in psychopathology (McTeague et al., 2020), the current study suggests that fMRI studies of emotional scene viewing could be useful in investigating emotional dysfunction across the diagnostic spectrum.

While peak and cluster level indices of replicability greatly exceeded the poor to fair levels of replicability reported when matching facial expressions of emotion, voxel replicability was somewhat more similar to that reported by Turner et al. (2018). Voxel level replicability, which assesses the covariation between t-statistics in each voxel of different maps, is the least informative index, however, as functional data are typically reported by peak location and cluster size. For these two replicability indices (i.e., peaks and clusters), replicability of emotional scene perception was far superior to face-matching, with the least informative index (voxel level) more equivalent.

In the current study, data from healthy individuals were combined with data from a set of anxiety patients selected to show equivalent enhanced functional activity in the amygdala (Sambuco, Bradley, Herring, & Lang, 2020). One possibility is that mixing the two samples might negatively impact replicability due to participant variability. Nonetheless, excellent levels of replicability were found for emotional (vs. neutral) scene viewing in the amygdala and visual cortex even with a sample size of 40 at the most conservative threshold ($p < .001$), which, if anything, may be underestimated here.

Large scale analyses with hundreds or even thousands of participants, such as those using data from the Human Connectome Project, the IAGEN consortium, the Adolescent Brain Cognitive Development (ABCD) project, or UK Biobank Brain Imaging, provide a unique opportunity to investigate the neural correlates of emotion, cognition, and mental health. The current data, however, suggest that fMRI replicability will also depend, to a large extent, on the specific emotional (or cognitive) tasks included in existing databases. Whereas fMRI replicability when matching facial expressions of emotion was very poor even with large samples, emotional scene-viewing produced excellent replicability with fairly small samples sizes.

More generally, rather than determining paradigm-free parameters dictating ideal sample sizes (or other parameters, such as trial number), fMRI replicability is probably best determined based on empirical tests (i.e., replicability analyses) in the context of specific tasks, which include assessment of different functionally relevant brain regions. The current data demonstrate that during rapid serial visual presentation of emotional scenes, a brief scan time (~5 min) is sufficient to produce excellent replicability of emotional, compared to neutral, BOLD differences for peaks, clusters, and voxels, even with relatively small sample sizes (i.e., 20 to 40), suggesting that, in the study of emotion, replicability depends on successful emotional engagement that can vary in different task-related brain regions.

ACKNOWLEDGEMENT

Thanks to Margaret Bradley for comments and input regarding data reporting and manuscript preparation. The data were collected, in part, by research supported by NIMH grants MH094386 and MH098078 to Peter Lang.

AUTHOR CONTRIBUTIONS

Nicola Sambuco: Conceptualization; Data curation; Formal analysis; Methodology; Software; Validation; Visualization; Writing – original draft; Writing – review & editing.

ORCID

Nicola Sambuco https://orcid.org/0000-0001-8941-1049

REFERENCES

Amaral, D. G., & Price, J. L. (1984). Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *The Journal of Comparative Neurology*, *230*, 465–496. https://doi.org/10.1002/cne.902300402

Amaral, D. G., Price, J. L., Pitkanen, A., & Carmichael, S. T. (1992). Anatomical organization of the primate amygdaloid complex. In J. P. Aggleton (Ed.), *The amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction* (pp. 1–66). Wiley-Liss.

Bossier, H., Roels, S. P., Seurinck, R., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Quinlan, E. B., Desrivières, S., Flor, H., Grigis, A., Garavan, H., Gowland, P., Heinz, A., Ittermann, B., Martinot, J.-L., Artiges, E., Nees, F., Orfanos, D. P., Poustka, L., … Moerkerke, B. (2020). The empirical replicability of task-based fMRI as a function of sample size. *NeuroImage*, *212*, e116601. https://doi.org/10.1016/j.neuroimage.2020.116601

Bradley, M. M., Codispoti, M., Cuthbert, B. M., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion*, *1*, 276–296. https://doi.org/10.1037/1528-3542.1.3.276

Bradley, M. M., Sabatinelli, D., Lang, P. J., Fitzsimmons, J. R., King, W., & Desai, P. (2003). Activation of the visual cortex in motivated attention. *Behavioral Neuroscience*, *117*(2), 369–380. https://doi.org/10.1037/0735-7044.117.2.369

Bradley, M. M., Sapigao, R. G., & Lang, P. J. (2017). Sympathetic ANS modulation of pupil diameter in emotional scene perception: Effects of hedonic content, brightness, and contrast. *Psychophysiology*, *54*(10), 1419–1435. https://doi.org/10.1111/psyp.12890

Britton, J. C., Taylor, S. F., Sudheimer, K. D., & Liberzon, I. (2006). Facial expressions and complex IAPS pictures: Common and differential networks. *NeuroImage*, *31*(2), 906–919. https://doi.org/10.1016/j.neuroimage.2005.12.050

Brown, T. A., DiNardo, P. A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV (ADIS-IV)*. Psychological Corporation.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, *86*, 127–137.

Codispoti, M., Ferrari, V., & Bradley, M. M. (2006). Repetitive picture processing: Autonomic and cortical correlates. *Brain Research*, *1068*(1), 213–220. https://doi.org/10.1016/j.brainres.2005.11.009

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*(4), 731–738. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Cremers, H. R., Wager, T. D., & Yarkoni, T. (2017). The relation between statistical power and inference in fMRI. *PLoS One*, *12*(11), e0184923. https://doi.org/10.1371/journal.pone.0184923

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*, 95–111. https://doi.org/10.1016/S0301-0511(99)00044-7

De Cesarei, A., & Codispoti, M. (2011). Affective modulation of the LPP and α-ERD during picture viewing. *Psychophysiology*, *48*(10), 1397–1404. https://doi.org/10.1111/j.1469-8986.2011.01204.x

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905. https://doi.org/10.1073/pnas.1602413113

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., Sison, M. L., Moffitt, T. E., Caspi, A., & Hariri, A. R. (2020). What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychological Science*, *31*(7), 792–806. https://doi.org/10.1177/0956797620916786

Frank, D. W., Costa, V. D., Averbeck, B. B., & Sabatinelli, D. (2019). Directional interconnectivity of the human amygdala, fusiform gyrus, and orbitofrontal cortex in emotional scene perception. *Journal of Neurophysiology*, *122*(4), 1530–1537. https://doi.org/10.1152/jn.00780.2018

Frank, D. W., & Sabatinelli, D. (2014). Human thalamic and amygdala modulation in emotional scene perception. *Brain Research*, *1587*, 69–76. https://doi.org/10.1016/j.brainres.2014.08.061

Freese, J. L., & Amaral, D. G. (2005). The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. *The Journal of Comparative Neurology*, *486*, 295–317. https://doi.org/10.1002/cne.20520

Hariri, A. R., Drabant, E. M., Munoz, K. E., Kolachana, B. S., Mattay, V. S., Egan, M. F., & Weinberger, D. R. (2005). A susceptibility gene for affective disorders and the response of the human amygdala. *Archives of General Psychiatry*, *62*(2), 146–152. https://doi.org/10.1001/archpsyc.62.2.146

Junghöfer, M., Sabatinelli, D., Bradley, M. M., Schupp, H. T., Elbert, T. R., & Lang, P. J. (2006). Fleeting images: Rapid affect discrimination in the visual cortex. *NeuroReport*, *17*(2), 225. https://doi.org/10.1097/01.wnr.0000198437.59883.bb

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8. University of Florida.

McTeague, L. M., Rosenberg, B. M., Lopez, J. W., Carreon, D. M., Huemer, J., Jiang, Y., Chick, C. F., Eickhoff, S. B., & Etkin, A. (2020). Identification of common neural circuit disruptions in emotional processing across psychiatric disorders. *American Journal of Psychiatry*, *177*(5), 411–421. https://doi.org/10.1176/appi.ajp.2019.18111271

Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, *121*, 47–57. https://doi.org/10.1093/brain/121.1.47

Nee, D. E. (2019). fMRI replicability depends upon sufficient individual-level data. *Communications Biology*, *2*(1), 130. https://doi.org/10.1038/s42003-019-0378-6

Padmala, S., Sambuco, N., & Pessoa, L. (2019). Interactions between reward motivation and emotional processing. *Progress in Brain Research*, *247*, 1–21. https://doi.org/10.1016/bs.pbr.2019.03.023

Padmala, S., Sirbu, M., & Pessoa, L. (2017). Potential reward reduces the adverse impact of negative distractor stimuli. *Social Cognitive and Affective Neuroscience*, *12*(9), nsx067. https://doi.org/10.1093/scan/nsx067

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a "low road" to "many roads" of evaluating biological significance. *Nature Reviews Neuroscience*, *11*, 773–783. https://doi.org/10.1038/nrn2920

Sabatinelli, D., Bradley, M. M., Fitzsimmons, J. R., & Lang, P. J. (2005). Parallel amygdala and inferotemporal activation reflect emotional intensity and fear relevance. *NeuroImage*, *24*(4), 1265–1270. https://doi.org/10.1016/j.neuroimage.2004.12.015

Sabatinelli, D., Bradley, M. M., Lang, P. J., Costa, V. D., & Versace, F. (2007). Pleasure rather than salience activates human nucleus accumbens and medial prefrontal cortex. *Journal of Neurophysiology*, *98*(3), 1374–1379. https://doi.org/10.1152/jn.00230.2007

Sabatinelli, D., Fortune, E. E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W. T., Beck, S., & Jeffries, J. (2011). Emotional perception: Meta-analyses of face and natural scene processing. *NeuroImage*, *54*(3), 2524–2533. https://doi.org/10.1016/j.neuroimage.2010.10.011

Sabatinelli, D., Frank, D. W., Wanger, T. J., Dhamala, M., Adhikari, B. M., & Li, X. (2014). The timing and directional connectivity of human frontoparietal and ventral visual attention networks in emotional scene perception. *Neuroscience*, *277*, 229–238. https://doi.org/10.1016/j.neuroscience.2014.07.005

Sabatinelli, D., Lang, P. J., Bradley, M. M., Costa, V. D., & Keil, A. (2009). The timing of emotional discrimination in human amygdala and ventral visual cortex. *The Journal of Neuroscience*, *29*(47), 14864–14868. https://doi.org/10.1523/jneurosci.3278-09.2009

Sambuco, N., Bradley, M., Herring, D., Hillbrandt, K., & Lang, P. J. (2020). Transdiagnostic trauma severity in anxiety and mood disorders: Functional brain activity during emotional scene processing. *Psychophysiology*, *57*(1), 1–12. https://doi.org/10.1111/psyp.13349

Sambuco, N., Bradley, M. M., Herring, D. R., & Lang, P. J. (2020). Common circuit or paradigm shift? The functional brain in emotional scene perception and emotional

imagery. *Psychophysiology*, *57*(4), 1–14. https://doi.org/10.1111/psyp.13522

Sambuco, N., Costa, V. D., Lang, P. J., & Bradley, M. M. (2020). Aversive perception in a threat context: Separate and independent neural activation. *Biological Psychology*, *154*, e107926. https://doi.org/10.1016/j.biopsycho.2020.107926

Sescousse, G., Caldú, X., Segura, B., & Dreher, J.-C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *37*(4), 681–696. https://doi.org/10.1016/j.neubiorev.2013.02.002

Somerville, L. H., Wagner, D. D., Wig, G. S., Moran, J. M., Whalen, P. J., & Kelley, W. M. (2013). Interactions between transient and sustained neural signals support the generation and regulation of anxious emotion. *Cerebral Cortex*, *23*(1), 49–60. https://doi.org/10.1093/cercor/bhr373

Spiegler, B. J., & Mishkin, M. (1981). Evidence for the sequential participation of inferior temporal cortex and amygdala in the acquisition of stimulus-reward associations. *Behavioral Brain Research*, *3*, 303–317. https://doi.org/10.1016/0166-4328(81)90002-4

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., & Poline, J.-B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *NeuroImage*, *35*(1), 105–120. https://doi.org/10.1016/j.neuroimage.2006.11.054

Turner, B. O., Paul, E. J., Miller, M. B., & Barbey, A. K. (2018). Small sample sizes reduce the replicability of task-based fMRI studies. *Communications Biology*, *1*(1), 62. https://doi.org/10.1038/s42003-018-0073-z

Turner, B. O., Santander, T., Paul, E. J., Barbey, A. K., & Miller, M. B. (2019). Reply to: fMRI replicability depends upon sufficient individual-level data. *Communications Biology*, *2*(1), 129. https://doi.org/10.1038/s42003-019-0379-5

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, *7*(11), 1271–1278. https://doi.org/10.1038/nn1341

Wangelin, B. C., Bradley, M. M., Kastner, A., & Lang, P. J. (2012). Affective engagement for facial expressions and emotional scenes: The influence of social anxiety. *Biological Psychology*, *91*(1), 103–110. https://doi.org/10.1016/j.biopsycho.2012.05.002

Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: The time-course of neural activity elicited by specific picture content. *Emotion*, *10*(6), 767–782. https://doi.org/10.1037/a0020242

Woo, C. W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. *NeuroImage*, *91*, 412–419. https://doi.org/10.1016/j.neuroimage.2013.12.058

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

FIGURE S1 Replicability of peak activation (left) and cluster activation (right) during pleasant, compared to unpleasant, scene viewing in the ventromedial prefrontal cortex (vmPFC) and the ventral striatum thresholded at $p < .05$, .01 or .001

TABLE S1 Demographic and diagnostic information of anxiety/mood disorder sample

TABLE S2 Characteristics of the regions showing enhanced BOLD activity during emotional, compared to neutral, scene viewing including the number of voxels, BOLD % change for emotional and neutral contents, and ANOVAs assessing the effects of Content (emotional, neutral), Group (healthy adults, patients) and their interaction

How to cite this article: Sambuco, N. (2022). fMRI replicability during emotional scene viewing: Functional regions and sample size. *Psychophysiology*, 59, e14000. https://doi.org/10.1111/psyp.14000