

Dorsolateral prefrontal cortex response to negative tweets relates to executive functioning

Sarah M. Tashjian¹ and Adriana Galván^{1,2}

¹Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095, USA, and ²Brain Research Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA

Corresponding authors Sarah M. Tashjian and Adriana Galván, Department of Psychology, University of California, 1285 Franz Hall, Los Angeles, CA 90095-1563. E-mail: smtashjian@ucla.edu, agalvan@ucla.edu

Abstract

Cognitive performance can become impaired when a stimulus evokes an emotional response. Social media often elicits emotional reactions, but, despite social media's ubiquity, cognitive and neural consequences of exposure to negative online content are relatively unknown. Fifty-seven human adults (18–29 years; 38 female) who identified with at least one historically-marginalized group performed a novel 'Tweet Task'. While undergoing functional magnetic resonance imaging, participants completed a spatial reasoning task before and after reading a set of actual tweets. Participants were randomly assigned to read negative, discriminatory tweets from President Trump (Negative Condition) or neutral tweets (Neutral Condition). Participants in the Negative Condition reported worsening affect and demonstrated performance interference post-tweet compared to those in the Neutral Condition. Affect post-tweet was associated with parametric reductions in left dorsolateral prefrontal cortex, which predicted variance in performance beyond elicited negative affect. Performance effects were demonstrated on an unrelated spatial reasoning task suggesting that engaging with negative, emotionally-arousing content on social media can have deleterious effects on executive functioning in non-social domains.

Key words: affect; cognition; dlPFC; fMRI; social media

Social media is increasingly ubiquitous in daily life, but little research has identified neural and cognitive consequences of engaging with emotionally-charged information via social media. Over 126 million people use the social networking service Twitter daily (Twitter, 2019) and after the 2016 election Twitter became an official communication source for the White House. As political attitudes in the United States become more polarized, the potential for engaging with perceived negative content on social media increases. A New York Times analysis estimated over half of President Trump's 11000+ tweets since becoming President involved attacks, with 1421 of those 5889 attacks levied against minority groups and immigrants (Shear et al., 2019). This type of negative, emotionally-charged information can interfere with cognitive functioning given its potential to capture attention and reduce resources available for engaging in goal-directed behavior (Schmader et al., 2008; Schweizer et al., 2019; Brady et al., 2020). The current study investigated whether

consumption of negative, discriminatory social media content affects neural systems needed for executive functioning.

Emotion can have both beneficial and deleterious effects on cognition. For example, humans demonstrate enhanced memory for emotional events, but distraction due to emotional stimuli can interfere with goal-directed behavior. Key determinants are task relevance and valence, with task relevant stimuli generally enhancing performance and positive valence associated with less interference (Dolcos et al., 2011; Iordan and Dolcos, 2017). Much of this extant research explores emotional distraction using graphic, but generic, photographic stimuli or evolutionarily relevant threat. In this study, actual tweets independently rated as negatively-valenced and including discriminatory content were used to test influences of frequently encountered emotional stimuli. This approach improves ecological validity of the emotional distractor within a paradigm designed to measure cognitive interference. Additionally, the

Received: 30 March 2020; Revised: 18 June 2020; Accepted: 11 July 2020

© The Author(s) 2020. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

tweet stimuli used in this study were of personal relevance to the study participants, all of whom identified with at least one historically marginalized group. This element of the present work strengthens the application of study findings to theory development and practical debates about the potentially detrimental interaction between group identity and negative online content.

Emotion and cognition are complex and interrelated, exhibiting both shared and distinct neural resources. Models of affect-cognition interactions propose emotional interference is linked to increased activation in the brain's ventral affective system and decreased activation in the dorsal executive control system (Dolcos and McCarthy, 2006). Studies of emotional distraction demonstrate task-irrelevant emotional information activates the ventral system involved in emotional processing while disrupting activation in dorsal brain regions, specifically the dorsolateral prefrontal cortex (dlPFC), which is involved in maintenance of task-relevant information (Cromheeke and Mueller, 2014). Reduced activation in the executive system is in turn associated with impaired performance. Although the dlPFC is largely regarded as a region involved in cognition, it also plays a role in emotion regulation (Ochsner et al., 2012; Kohn et al., 2014). Exposure to information that elicits negative affect irrelevant to one's current goals generates regulatory attempts over the affective response. Regulatory attempts engage neural regions like the dlPFC, potentially interfering with the availability of these resources during future cognitive demands (Pessoa, 2009).

The current study paired a novel behavioral paradigm with functional magnetic resonance imaging (fMRI) to test the hypothesis that tweets inclusive of negative, discriminatory content would interfere with performance on a spatial reasoning task. All participants reported identifying with at least one historically marginalized identity group, increasing the personal relevance of the negative tweets. Deleterious effects of emotional interference may be particularly detrimental for marginalized individuals who are already disadvantaged due to systemic disparities and stigmatization. Tweets were independently rated for valence and participants were randomly assigned to two groups. We hypothesized participants assigned to read negatively-valenced, discriminatory tweets from President Trump (Negative Condition) would experience significantly worsening affect and impaired behavioral performance compared to those assigned to read neutral tweets (Neutral Condition). We hypothesized reduced activation in dorsal executive regions associated with worsening affect during tweet presentation would relate to worse subsequent performance on the task, in line with prior work on emotional interference.

Materials and Methods

Participants

Fifty-seven human adults between the ages of 18 and 29 years ($M_{age} = 20.895 \pm 2.289$ years, 38 female) completed the study. Thirty percent of participants identified as Hispanic/Latinx, 21% Asian, 18% Caucasian, 14% Black, 10% multiple races/ethnicities, and 7% Middle Eastern. One additional participant completed the task behaviorally but did not complete the scan and was therefore excluded from analysis (female, 23 years). Two additional participants were excluded from analysis due to technical errors during the scan (2 females, $M_{age} = 22$ years). Participants were recruited via flyers. Thirty-six participants were previously enrolled in a laboratory study on post-election distress (see Tashjian & Galván, 2018). After receiving approval from the university's Institutional Review Board, participant eligibility for

this study was determined based on responses to three pre-screening questions including two questions about the personal effect of the 2016 U.S. presidential election and one question about identification with various marginalized identity groups. All participants reported feeling that the 2016 U.S. election result would personally affect them, rated affective response to the election as 5 or higher on a scale of 1 = 'no negative affective response' to 7 = 'an extremely negative affective response', and identified with at least one historically marginalized group in terms of gender identity, ethnicity, or sexual orientation. In addition to the prescreening questions, eligibility criteria included fluency in English, between the ages of 18 and 30 years, and right handedness. Exclusion criteria included no prior developmental, psychiatric, or neurological disorder; no psychotropic medication; not claustrophobic; and no metal in the body. Written consent was obtained and participants were compensated for their participation. This study complies with all relevant ethical regulations regarding human research participants. Participants were tested between December 2017 and April 2018.

Experimental paradigm

Participants were randomly assigned to one of two Conditions to complete the Tweet Task in the scanner. Twenty-nine participants ($M_{age} = 20.862$, $SD = 2.371$, 19 female) were assigned to the Negative Condition and 28 participants ($M_{age} = 20.929$, $SD = 2.243$, 19 female) were assigned to the Neutral Condition. Participants received spoken and written instructions and completed one mental rotation (MR) practice trial outside of the scanner before beginning the experimental session. During the Tweet Task (Figure 1), participants completed 5 task blocks in the following order: (1) MR trials ('pre-tweet MR trials', 30 trials), (2) affect ratings ('pre-ratings'), (3) tweet presentation (10 tweets), (4) affect ratings ('post-ratings'), and (5) MR trials ('post-tweet MR trials', 30 trials).

During pre- and post-tweet MR trials, participants were presented with pairs of objects, each consisting of seven to 11 cubes connected face-to-face to form a pipe-like object with four connected arms and two free ends (Ganis and Kievit, 2015). In each stimulus, one of the objects in the pair was rotated relative to the other with shading and depth cues. Two angle disparities were used: 100 and 150 degrees. Half of the shapes were matches, half were not matches. Stimuli were randomly presented. All shapes were white presented on a black background for 1100 ms (average RT from Ganis & Kievit in adults for 100 degree rotations = 3026 ms, 150 degree rotations = 3191 ms). When the shapes offset, participants were given a decision screen during which they were instructed to make a button press to indicate whether they thought the shapes were identical after mentally rotating one of them. No performance feedback was provided. Offset was based on button press to prevent missing data that might be incurred by pre-setting stimuli offset. Jittered interstimulus intervals (ISIs) displaying a white fixation cross on a black background were included ranging from 500-4000 ms, $M = 1466$ ms. ISIs served as implicit baseline.

During pre-affect ratings, participants rated their current affect for 9 states on a scale of 1 = 'not at all' to 4 = 'very much'. Six states indexed negative affect. Four states were taken from the Profile of Mood States subscales for anger ('angry', 'annoyed') and depression ('sad', 'hopeless') (McNair et al., 1971). Two states were added indexing disgust ('disgusted') and fear/anxiety ('worried') (Ekman, 1999). Three states indexed positive or neutral affect as a manipulation check: 'amused', 'indifferent', 'interested'. Affect presentation was randomized. After completing the tweet reading block, participants completed post-affect

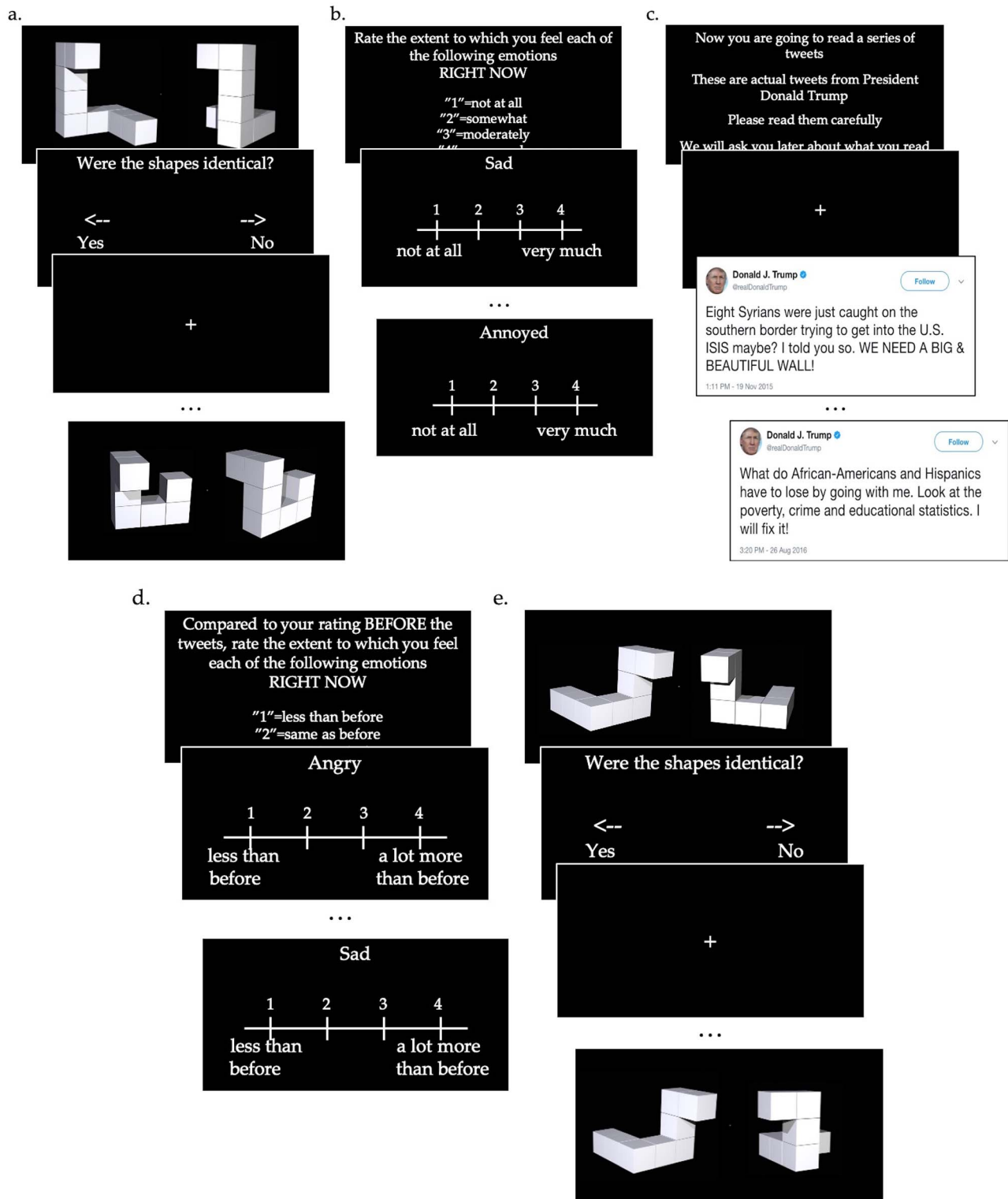


Fig. 1. Schematic of the Tweet Task. Participants completed 5 task blocks in the following order: (a) pre-tweet mental rotation (MR) trials (30 trials), (b) pre-affect ratings, (c) tweet presentation (10 tweets), (d) post-affect ratings, and (e) post-tweet MR trials (30 trials).

ratings of the same 9 states on a scale of 1 = 'less than before' to 4 = 'a lot more than before'. A comparative scale was used during post-rating to prevent pre-rating floor or ceiling effects from concealing affect change.

During the tweet reading block, participants either saw a set of actual tweets from President Trump independently rated as

negatively valenced and inclusive of discriminatory language (negative tweets) or a set of actual tweets related to general topics rated as neutrally valenced (neutral tweets). Neutral tweets were actual tweets but were shown to be sent from a single fake account called 'Stuff Facts' with a user image of an older white male to reduce disparate effects of visual presentation

between Conditions (Figure S1). Prior to the study, 63 tweets posted by President Trump were independently rated for valence and arousal by 55 participants ($M_{age} = 25.255$, $SD = 2.797$, 28 females, 74.545% Caucasian) using Amazon's Mechanical Turk (MTurk). Tweets differed in length, but were selected based on valence ratings and were selected from actual tweets to preserve ecological validity of the stimuli (Table S1). Additionally, selected tweets were all understood independent of additional contextual information. Attention checks were included in the MTurk survey to ensure validity of ratings. Of the tweets rated on MTurk, 10 neutral and 10 negative tweets were selected for use in the Tweet Task based on ratings. Of the tweets designated for use, negative tweets were rated on average 1.900 (average range 1.491–2.636) and neutral tweets were rated 4.484 (average range 4.164–4.855) on a Likert scale of 1 = 'extremely negative' to 7 = 'extremely positive' with 4 = 'neutral'. See Supplemental Materials Table S1 for a list of tweets and valence ratings. Tweets were displayed during the task for 12 000 ms. Participants were instructed to 'read the tweets carefully' and that they would be 'asked about the tweets later'. Jittered ISIs displaying a white fixation cross on a black background were included ranging from 50–4000 ms, $M = 1950$ ms. ISIs served as implicit baseline. Outside of the scanner participants completed a short comprehension check during which they were asked to identify whether they previously read each of a set of 20 tweets (50% familiar, 50% unfamiliar; $M_{accuracy} = 97.969\%$, $SD_{accuracy} = 4.191\%$). Accuracy on the comprehension check did not differ by condition, $t(55) = 1.274$, $p = 0.208$, $M_{diff} = 0.014$, $SD_{diff} = 0.011$.

Questionnaires

Individual differences that may affect performance and affective response to negative tweets were assessed via questionnaire.

Participants ($n = 52$) reported on average news consumption, including via social media, per week over the last month on a scale of 1 = 'almost never', 3 = 'average 2–3x per week', 5 = 'more than 1x daily'. The question was framed with respect to news consumption rather than general social media use given conceptualization of Twitter communication from President Trump as a news source.

Prior work suggests shifts in motivation may affect or even explain performance effects (Vohs et al., 2012; Earle et al., 2015). To test whether there were differences by Condition (negative v. neutral tweets) in MR task motivation or perceived difficulty, participants ($n = 57$) answered six of follow-up questions after the scan. Participants were asked to answer four questions on a Likert scale of 1 = 'not at all', 4 = 'somewhat true', 7 = 'very true' as to how true each statement was for them. Likert scale questions were as follows: (1) I think I did pretty well at the MR task, compared to others; (2) After working on the MRs for a while, I felt competent; (3) I tried very hard on the MR task; (4) It was important for me to do well at the MR task. Participants also answered two questions about pre- and post-tweet perceptions: (5) I felt more tense doing the MR task [before] or [after] reading the tweets; (6) I felt more determined to do well on the MR task [before] or [after] reading the tweets.

fMRI data acquisition

Whole-brain fMRI data were acquired on a 3T Siemens Magnetom Prisma scanner: voxel size = 2.0 x 2.0 x 2.0 mm, TR = 1000 ms, TE = 37 ms, slice thickness = 2.0 mm, flip angle = 52 degrees, FOV = 208 mm, 60 interleaved slices, multi-band acceleration = 6x. AutoAlign was used for automated

positioning and alignment of anatomy-related slices using alignment perpendicular to the midsagittal plane and tilted along the corpus callosum contour. Structural images were acquired using a high-resolution, magnetization-prepared rapid-acquisition gradient echo (MPRAGE) sequence for registration: voxel size = 1.0 x 1.0 x 1.0 mm, TR = 1900 ms, TE = 2.26 ms, slice thickness = 1.0 mm, flip angle = 9 degrees, FOV = 250 mm, 176 slices.

Stimuli were presented using E-Prime Professional 2.0 and were projected onto a flat screen mounted in the scanner bore. Participants viewed the screen using a mirror mounted on a 32-channel head coil. Extensive head padding was used to minimize participant head motion and to enhance comfort. Participants made their responses with their right hand using a 4-finger-button response box.

fMRI data analyses

Preprocessing was conducted using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB Software Library). Preprocessing consisted of nonbrain removal using BET (Brain Extraction Tool), high-pass filtering (100-s cutoff), and spatial smoothing using a Gaussian kernel of FWHM 5 mm. The first three volumes were discarded to allow for image stabilization. Motion correction was performed with MCFLIRT (intra-modal motion correction tool) using 24 standard and extended regressors and additional individual spike regressors created using *fsl_motion_outliers* (frame displacement threshold = 75th percentile plus 1.5 times the interquartile range). Functional data were registered to subject-space MPRAGE images using boundary based registration (BBR; Greve and Fischl, 2009) and then to MNI 2.0 x 2.0 x 2.0 mm stereotaxic space with 12 degrees of freedom via FLIRT (FMRIB's Linear Image Registration Tool). FILM (FMRIB's Improved Linear Model) prewhitening was performed to estimate voxelwise autocorrelation and improve estimation efficiency.

To investigate main effects of the task, individual subject data using fixed-effects general linear models (GLMs) were defined with 6 multiple regressors: task instructions, MR stim presentation (pre and post), MR decision screen (pre), MR decision screen (post), affect rating screens (pre and post), tweet presentation. Using this GLM, whole-brain activation to tweets versus baseline and MR-post decisions versus baseline were identified.

Additional analyses were conducted to estimate neural habituation, defined as greater response decrement over the course of stimuli presentation. To estimate neural change during tweet presentation, a parametric regressor was added modeling the linear effect of presentation order with decreasing weights from early to late presentation. This linear modulated regressor was orthogonalized with respect to the lower order regressor representing average activation during tweet presentation (Mumford et al., 2015). Using this GLM, parametric decreases in neural activation versus baseline were identified. To identify affect-induced performance deficits specific to neural change, covariations between fMRI signal and individual differences in post-tweet affect changes were assessed. Both standard and parametric regressors were modeled with a canonical (double-gamma) hemodynamic response function for a duration from stimulus onset to stimulus offset. Temporal derivatives were included as covariates of no interest for all regressors, allowing a better fit for the whole model and reducing unexplained noise. Group-level analyses were performed using the FMRIB Local Analysis of Mixed Effects (FLAME-1) module in FSL (Beckmann et al., 2003), $Z > 3.0$, FWE-corrected $p < 0.05$. Outliers were

de-weighted in the multisubject statistics using mixture modeling (Woolrich, 2008).

Statistical analyses

Data analyses were performed using R statistical software (version 3.6.1) using lme4 (Bates et al., 2015; version 1.1.21). P-values below 0.050 were regarded as statistically significant and p-values between 0.050 and 0.100 (inclusive) were regarded as marginally significant. Data were analyzed using a multi-level modeling framework because the data consisted of repeated measures (trials) nested within individuals. Additionally, likelihood ratio tests indicated evidence of individual effects on MR-post performance, Accuracy $\chi^2(1, N = 57) = 84.799, p < 0.001$, RT $\chi^2(1, N = 57) = 351.910, p < 0.001$. Accuracy was coded as 0 = incorrect response, 1 = correct response and RT is in ms. Random intercepts were included to account for within-person individual differences. Significance testing was conducted using Satterthwaite approximations (Kuznetsova et al., 2017). Effect sizes reported as R^2 were calculated using the r.squaredGLMM() function in MuMIn (Bartoń; version 1.43.6) and are reported as conditional effects of variance explained by the entire model and using the delta method for binomial models. Effect sizes reported as d were calculated as Cohen's d using pooled standard deviation.

Data availability

Data and materials can be accessed at Open Science Framework (OSF, <https://osf.io/sfh48/>), including raw behavioral data and unthresholded statistical maps for Figures 4-6.

Results

MR performance

On average, participants completed 20.263 out of 30 pre-tweet MR trials correctly (range 9–29, $SD = 4.422$) and 21.421 out of 30 post-tweet MR trials correctly (range 11–30, $SD = 4.773$). Reaction time during pre-tweet MR decisions averaged 1.189 sec, range 0.360–2.573, $SD = 0.500$, and 1.097 sec, range 0.361–2.863, $SD = 0.486$ during post-tweet MR decisions.

Age was not significantly associated with pre- or post-tweet MR accuracy, $B = 0.067, SE = 0.042, z = 1.621, R^2 = 0.071, p = 0.105$, and $B = 0.005, SE = 0.050, z = 0.094, R^2 = 0.101, p = 0.925$, respectively, nor was it associated with pre- or post-tweet MR RT, $B = -.006, SE = 0.029, \beta = -.0206, R^2 = 0.258, p = 0.838$, and $B = -.035, SE = 0.028, \beta = -1.256, R^2 = 0.256, p = 0.214$, respectively. Sex was not significantly associated with pre- or post-tweet MR accuracy, $B = -.282, SE = 0.198, z = -1.424, R^2 = 0.070, p = 0.154$, and $B = -.326, SE = 0.241, z = -1.353, R^2 = 0.102, p = 0.176$, respectively, nor was it associated with pre- or post-tweet MR RT, $B = 0.096, SE = 0.141, \beta = .680, R^2 = 0.258, p = 0.500$ and $B = 0.050, SE = 0.138, \beta = .366, R^2 = 0.256, p = 0.716$, respectively.

Condition (negative v. neutral tweets) was not directly associated with post-tweet MR accuracy, $B = -.027, SE = 0.229, z = -.117, R^2 = 0.101, p = 0.907$, or post-tweet MR RT, $B = -.083, SE = 0.129, \beta = -.639, R^2 = 0.256, p = 0.525$, nor was it associated with pre-tweet MR accuracy, $B = -.130, SE = 0.189, z = -.685, R^2 = 0.071, p = 0.493$, or pre-tweet MR RT, $B = -.003, SE = 0.134, \beta = -.024, R^2 = 0.258, p = 0.981$. Condition interacted with trial order, $R^2 = 0.263$ (Table 1, Figure 2), such that individuals in the Neutral Condition demonstrated post-tweet MR RT improvements as the trials progressed, simple effect $B = -.013, SE = 0.004, \beta = -3.439, R^2 = 0.210, p < 0.001$, but

Table 1. Mixed effects regression model of Condition interacting with trial order to predict reaction time during mental rotation post-test trials

Fixed Effects	B	SE	β	p
Intercept	1.339	.105	12.813	<.001
Condition	-.320	.147	-2.182	.032
Trial Order	-.013	.003	-4.074	<.001
Interaction	.015	.004	3.445	<.001
Random Effects	Variance	SD		
Intercept	.217	.466		
Residual	.631	.795		

Note: $N = 57$, trials = 1710.

those in the Negative Condition showed no post-tweet trial-related improvements, simple effect $B = 0.002, SE = 0.002, \beta = .984, R^2 = 0.356, p = 0.325$. This effect held controlling for pre-tweet MR RT and post-tweet MR accuracy. No significant interaction between Condition and trial order was found for pre-tweet MR RT, $B = -.006, SE = 0.005, \beta = -1.348, R^2 = 0.272, p = 0.178$.

Affective response to tweets

Affect change scores, measured with post-tweet responses, for the six negative affect states reported post-tweets were highly correlated, $r_s > 0.610$ ($M = 0.728$), $p_s < 0.001$. Dimension reduction analyses indicated all states loaded on a single component (77.591% of variance, Varimax rotation, component loadings > 0.822 with no cross loadings), thus an average composite change score was calculated for the six affect states. Analyses of each individual negative affect state are provided in Supplemental Materials.

Age was not significantly associated with negative affect change, $r(57) = -.122, p = 0.367$. Negative affect change did not significantly differ by sex, $t(55) = 1.098, d = 0.311, p = 0.277$, $M_{males} = 2.430, SD_{males} = 0.930, M_{females} = 2.132, SD_{females} = 0.983$.

Negative affect change significantly differed by Condition such that those in the Negative Condition reported significantly worsening affect post-tweet compared with the Neutral Condition, $t(55) = -10.453, p < 0.001, M_{negative} = 3.000, SD_{negative} = 0.639, M_{neutral} = 1.4345, SD_{neutral} = 0.476$ (Figure 3a). Individuals in the Negative Condition reported significantly higher worsening affect for all six negative affect states compared to those in the Neutral Condition (Supplemental Materials Table S2). There were no differences in pre-tweet negative affect by Condition, $t(55) = 0.670, d = 0.178, p = 0.506, M_{negative} = 1.310, SD_{negative} = 0.295, M_{neutral} = 1.375, SD_{neutral} = 0.425$. Negative affect change interacted with trial order to predict post-tweet MR RT, $B = 0.005, SE = 0.002, \beta = 2.173, R^2 = 0.264, p = 0.030$ (Figure 3b), such that reduced levels of negative affect post-tweet was associated with improved RT over the course of the task, simple effect $B = -.012, SE = 0.005, \beta = -2.640, R^2 = 0.178, p = 0.009$, but increased levels of negative affect was associated with no RT improvements over the course of the post-test, simple effect $B = 0.002, SE = 0.003, \beta = .760, R^2 = 0.400, p = 0.448$. This effect held controlling for Condition. RT effects were driven by increases in anger, annoyance, and disgust (Supplemental Materials Table S3). There was no significant interaction for post-tweet MR accuracy, $B = -.0005, SE = 0.006, z = -.075, R^2 = 0.101, p = 0.940$.

Positive affect (amused, interested) change scores were significantly correlated, $r = 0.366, p = 0.005$; pre-tweet positive affect correlations, $r = 0.645, p < 0.001$. An average score was created for pre- and post-tweet positive affect scores. There were no differences in pre-tweet positive affect (average

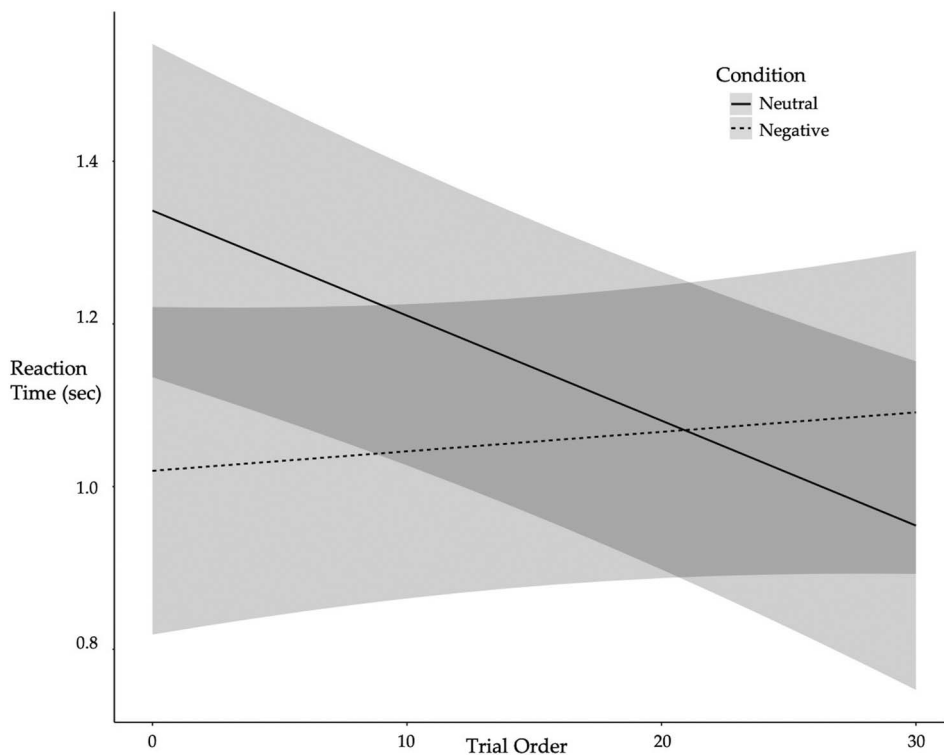


Fig. 2. Individuals in the Neutral Condition showed improvement in RT as post-tweet MR trials progressed whereas individuals in the Negative Condition showed no improvement over the course of the MR post-test. Lines depict predicted values (marginal effects) for the regression model with 95% confidence interval bands shown in gray. Condition: 0 = neutral, 1 = negative. $N = 57$.

amused, interested) or neutral affect (indifferent) by Condition, $t_{\text{positiveaff}}(55) = -.773$, $d_{\text{positiveaff}} = 0.205$, $p_{\text{positiveaff}} = 0.443$, $M_{\text{negative}} = 2.603$, $SD_{\text{negative}} = 0.795$, $M_{\text{neutral}} = 2.446$, $SD_{\text{neutral}} = 0.737$; $t_{\text{neutralaff}}(55) = -1.662$, $d_{\text{neutralaff}} = 0.440$, $p_{\text{neutralaff}} = 0.102$, $M_{\text{negative}} = 2.586$, $SD_{\text{negative}} = 0.983$, $M_{\text{neutral}} = 2.179$, $SD_{\text{neutral}} = 0.863$. Post-tweet neutral affect did not significantly differ by condition, $t(55) = 1.500$, $d = 0.397$, $p = 0.139$, $M_{\text{negative}} = 1.483$, $SD_{\text{negative}} = 0.688$, $M_{\text{neutral}} = 1.786$, $SD_{\text{neutral}} = 0.833$. Post-tweet positive affect significantly differed by condition such that those in the neutral condition reported increases in positive affect after tweet presentation, $t(55) = 4.958$, $d = 1.317$, $p < 0.001$, $M_{\text{negative}} = 1.913$, $SD_{\text{negative}} = 0.682$, $M_{\text{neutral}} = 2.750$, $SD_{\text{neutral}} = 0.585$.

Post-hoc analyses were conducted assessing whether frequency of news consumption were associated with affect change. On average, participants reported a score of 2.846 for news consumption, $SD = 1.211$, range = 1–5. News consumption did not significantly differ by Condition, $t(50) = 1.387$, $p = 0.172$. News consumption was not associated with affect change in either Condition, Negative $B = 0.121$, $SE = 0.108$, $t = 1.119$, $R^2 = 0.050$, $p = 0.274$; Neutral $B = -.018$, $SE = 0.084$, $t = -.212$, $R^2 = 0.002$, $p = 0.834$.

fMRI main effects

Compared to baseline, tweets elicited greater activation in distributed regions previously identified as relevant for emotional response and regulation (bilateral amygdala, orbitofrontal cortex), memory encoding (bilateral hippocampus, middle and superior temporal gyrus), and language processing (inferior frontal gyrus, precentral gyrus) (Figure 4a, Table 2). Tweets elicited deactivation in regions of the executive control network and default mode network compared to baseline (Figure 4b, Table 2).

Mental rotation decisions post-tweet elicited greater activation compared to baseline in regions associated with working memory and attention (dlPFC, anterior cingulate cortex, bilateral insula, middle frontal gyrus) and spatial cognition (precentral and postcentral gyrus, parietal lobule) (Figure 5, Table 2).

Neural habituation

To identify neural change during tweets, linear changes in activation across trials were assessed. Trials were parametrically modulated with decreasing weights from earliest to latest trials. For the contrast of tweets versus baseline, regions of the dorsal executive system demonstrated greater linear decreases over the course of tweet presentation (Figure 6a, Table 2). Worsening negative affect pre- to post-tweet was associated with targeted increases in the left dlPFC habituation (Figure 6b, Table 2). dlPFC habituation was associated with worsening affect for each negative affect state (Supplemental Materials Table S4).

Performance effects of neural change associated with affect changes were investigated for several reasons including that there was a range of affective response in each Condition and given hypotheses regarding the effect of emotional load on cognitive performance. dlPFC activation change significantly moderated the association between trial order and RT, $R^2 = 0.270$ (Table 3, Figure 7), such that those with greater habituation demonstrated no RT improvements, simple effect $B = 0.001$, $SE = 0.002$, $\beta = .421$, $R^2 = 0.417$, $p = 0.674$, whereas those with less activation change showed RT improvements over the course of the post-test trials, simple effect $B = -.010$, $SE = 0.004$, $\beta = -2.858$, $R^2 = 0.208$, $p = 0.004$. Analyses controlled for affect change and Condition.

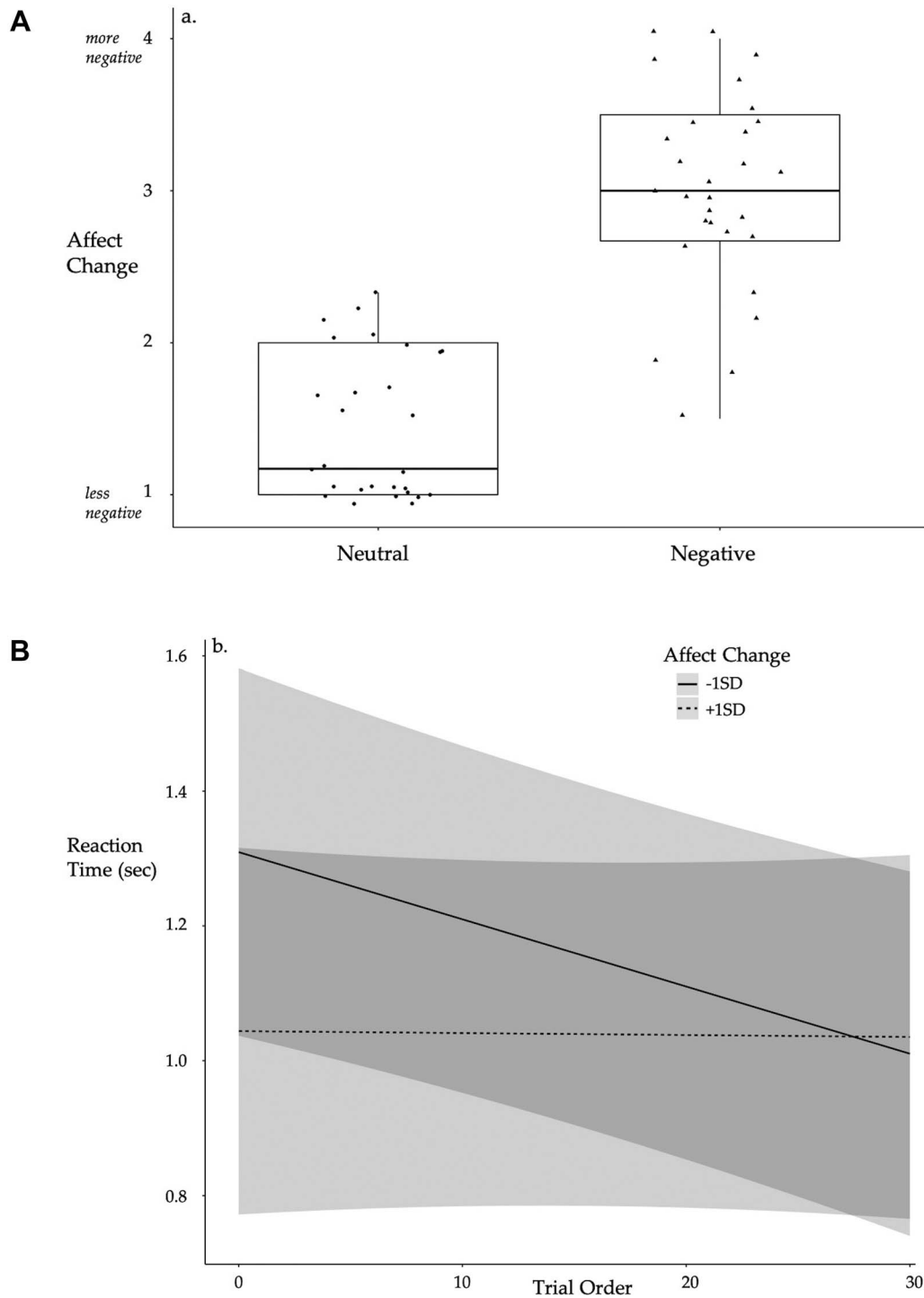


Fig. 3. (a) Significant differences in affect change by condition such that individuals in the Negative Condition reported worsening affect after reading tweets. Boxplots depict median line with interquartile range box (25% to 75%) and minimum/maximum lines. Individual points are provided by group. (b) Individuals reporting improved affect demonstrated improved RT as post-tweet MR trials progressed whereas individuals reporting worsening affect showed no improvement over the course of the MR post-test, controlling for Condition. Lines depict predicted values (marginal effects) for the regression model with 95% confidence interval bands shown in gray. Condition: 0 = neutral, 1 = negative. Affect change: higher scores = worsening affect. $N = 57$.

Descriptive statistics for the post-task questionnaire are provided in Supplemental Materials Table S5. There were no significant differences by Condition on MR task motivation or perceptions of competence. Participants did not differ in the extent to which they felt: (Q1) they did well on the task, $t(55) = 0.561$,

$p = 0.577$; (2) they were competent on the task, $t(55) = 0.052$, $p = 0.959$; (3) they tried hard on the task, $t(55) = -0.327$, $p = 0.745$; or it was important to do well on the task, $t(55) = -0.896$, $p = 0.374$. Participants also did not significantly differ by Condition in the extent to which they felt more determined to do well pre- or

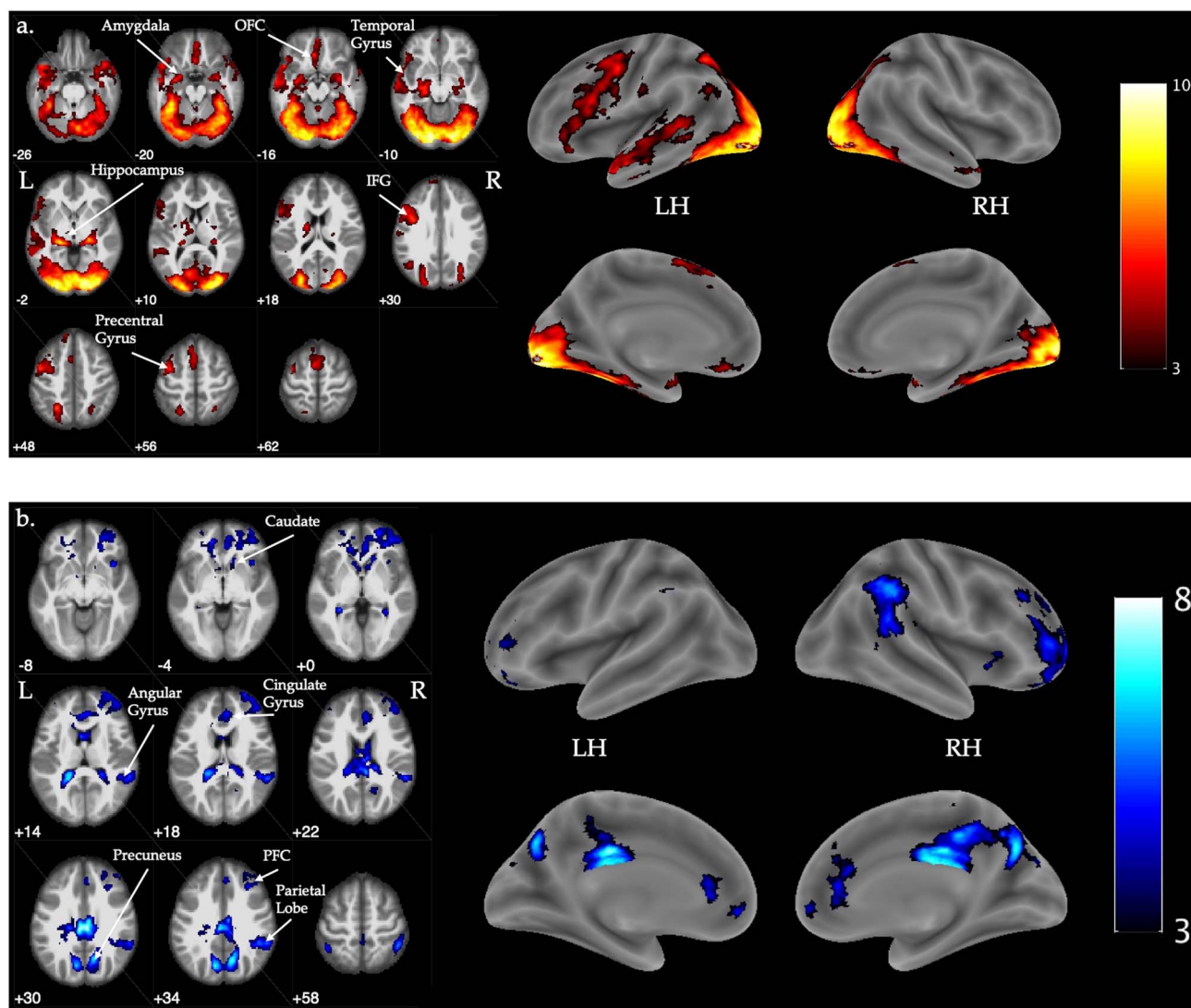


Fig. 4. Visualization of significant (a) increases and (b) decreases in activation for the contrast of tweets versus baseline. Flame1, $Z > 3.0$, FWE-corrected $p < 0.05$, outliers deweighted. Color bar indicates Z intensity values. Coordinates are z-plane in MNI space. LH=left hemisphere, RH=right hemisphere, OFC=orbitofrontal cortex, IFG=inferior frontal gyrus, PFC=prefrontal cortex. $N = 57$.

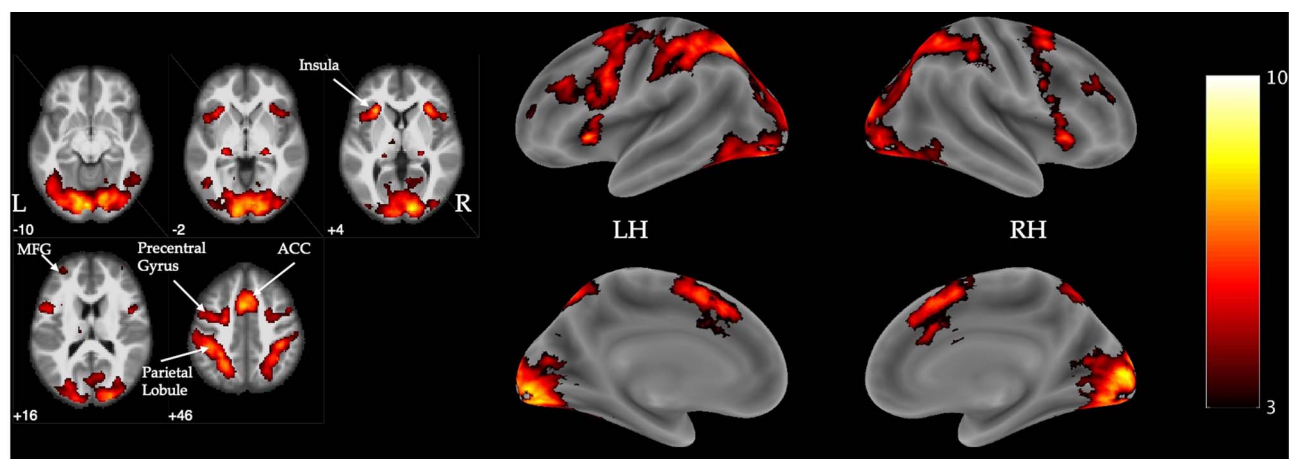


Fig. 5. Visualization of significant activation for the contrast of MR-post decisions versus baseline. Flame1, $Z > 3.0$, FWE-corrected $p < 0.05$, outliers deweighted. Color bar indicates Z intensity values. Coordinates are z-plane in MNI space. LH=left hemisphere, RH=right hemisphere, MFG=middle frontal gyrus, ACC=anterior cingulate cortex. $N = 57$.

Table 2. Significant clusters from group level analyses

Cluster Peak Region ¹	R/L	Peak MNI Coordinates			Max Z value	Voxels	p value
		X	Y	Z			
Tweets > baseline							
Occipital fusiform gyrus; Lingual gyrus; Occipital pole	R	18	-88	-6	10.40	37 636	<.001
Juxtapositional lobule cortex	L	-6	4	62	5.54	1270	<.001
Frontal medial cortex	L	-6	32	-16	6.01	534	<.001
Frontal pole	L	-12	48	48	4.19	326	<.001
Tweets < baseline							
Cingulate gyrus	R	4	-24	30	7.79	16 366	<.001
Superior parietal lobule	L	-44	-50	58	4.90	371	<.001
MR decisions > baseline							
Occipital pole; Intracalcarine cortex	R	14	-92	4	8.55	36 619	<.001
Thalamus	L	-24	-26	-2	5.82	247	<.001
Frontal pole	L	-30	54	16	4.15	153	.004
Thalamus	R	-26	-2	5.78	131	.009	
Tweet habituation > baseline							
Angular gyrus; Supramarginal gyrus	R	52	-48	42	6.48	12 540	<.001
Middle frontal gyrus	R	46	14	48	5.02	4019	<.001
Supramarginal gyrus; Angular gyrus	L	-44	-50	5.84	1074	<.001	
Insular cortex	R	32	16	0	4.34	467	<.001
Middle temporal gyrus	R	-36	-14	4.75	382	<.001	
Frontal pole	L	-34	56	12	4.52	256	<.001
Thalamus	R	26	-24	2	4.69	249	<.001
Frontal operculum cortex; Inferior frontal gyrus	L	-44	16	6	4.72	154	.010
Tweet habituation associated with affect change > baseline							
Frontal pole; Middle frontal gyrus	L	-42	42	16	4.37	1.046	

Note: N = 57. Z > 3.0, FWE-corrected p < 0.05, Flame1. R = Right hemisphere, L = Left hemisphere. ¹Regions based on the Harvard-Oxford Structural Atlas.

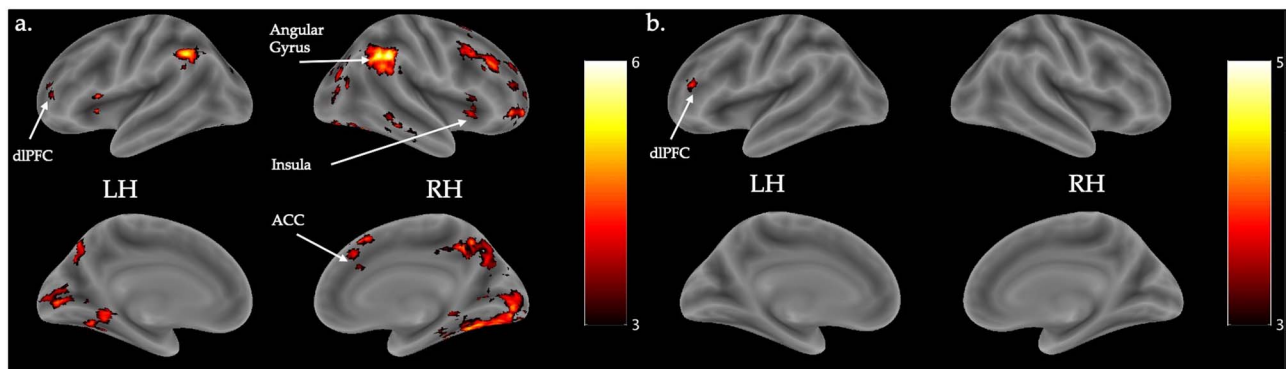


Fig. 6. (a) Visualization of activation showing linear decreases in activation across trials for the contrast of tweets versus baseline, irrespective of changes in negative affect. (b) Visualization of activation showing increased linear change across trials for the contrast of tweets versus baseline and correlated with more extreme changes in negative affect. Flame1, Z > 3.0, FWE-corrected p < 0.05, outliers deweighted. Color bar indicates Z intensity values. LH = left hemisphere, RH = right hemisphere. N = 57.

post-tweet, $\chi^2(1, N = 57) = 0.849$, $p = 0.357$. Participants did significantly differ by Condition in the extent to which they felt more tense completing MR trials, $\chi^2(1, N = 57) = 17.923$, $p < 0.001$, such that more participants in the Negative Condition reported increased tension after tweet presentation compared to those in the Neutral Condition.

Discussion

Results support the hypothesis that negative stimuli experienced through social media evoke negative affect, which relates to neural and executive functioning. Participants exposed to negative, discriminatory tweets reported worsening affect than those exposed to neutral tweets. Worsening affect was associated with greater habituation in the dIPFC, a neural region supporting higher-order cognition and affective control. Individual differences in dIPFC habituation were related to RT performance

on a spatial reasoning task such that those with greater habituation demonstrated no performance improvements over the course of the post-tweet task but those with less neural change showed linear decreases in RT as trials progressed. Results provide evidence of the detrimental effect of negative, emotional content via social media on cognitive processes.

Although distributed neural regions in the dorsal executive system exhibited linear decreases during tweet presentation, targeted habituation was observed in the dIPFC for those experiencing worsening affect as a result of the tweet presentation. Given relevance of the dIPFC for affect regulation, we interpret this activation change as a reallocation of resources resulting from emotional distraction. The dIPFC supports allocation of attentional resources and active maintenance of task-relevant information in working memory, thus, this hijacking of cognitive resources is relevant for a myriad of cognitive processes. Studies using non-invasive brain stimulation techniques draw direct

Table 3. Mixed effects regression model of neural change interacting with trial order to predict reaction time during mental rotation post-test trials, controlling for affect change and Condition

Fixed Effects	B	SE	β	p
Intercept	1.240	.218	5.678	<.001
Condition	.011	.226	.047	.963
Affect Change	-.031	.126	-.246	.807
Trial Order	-.005	.002	-2.306	.021
Neural Change	-.160	.085	-1.892	.062
Interaction	.007	.002	3.128	.002
Random Effects	Variance	SD		
Intercept	.223	.472		
Residual	.632	.795		

Note: N = 57, trials = 1710. Neural data was extracted from significant clusters using *fslmeants* and then z-scored to facilitate interpretation and improve scaling. Condition: 0 = neutral, 1 = negative. Affect change: higher scores = worsening affect post-tweet.

associations between dlPFC activation and RT during cognitively demanding tasks (Brunoni and Vanderhasselt, 2014), and habituation in the dlPFC in this study was associated with increased response latency post-tweet. Response latency may be adaptive as slowed responding provides additional time to engage cognitive systems necessary for accurate performance. No performance differences in accuracy by Condition were observed.

This work makes several advances to understanding how emotions can interfere with cognitive functioning. By utilizing actual tweets, this study increases ecological validity compared with prior studies probing emotion-cognition interactions. Assessing neural activation corresponding to self-reported affect change reduces experimenter imposed assumptions about arousal to identify neural systems taxed when negative emotions are elicited. This study did not involve overt activation of threat and thereby supports prior assertions that even subtle threats can induce performance interference (Stone and McWhinnie, 2008).

This study specifically focused on individuals for whom the negative, discriminatory tweet content was of personal relevance. Results demonstrate that widely read tweets may have deleterious effects on executive functioning in a large segment of the population: historically marginalized identity groups. Importantly, more negative preexisting feelings about President Trump were associated with increased negative affect post-tweet, suggesting preexisting conceptions about underlying subject matter circulated on social media can amplify affective responses to online content. Certain individuals may therefore be more vulnerable to the interfering effects of social media on executive functioning, a hypothesis worthy of continued investigation. Although news consumption was not associated with affective response to the negative tweets in this study, prior work links excessive social media use with impaired decision making (Meshi et al., 2019) and increased perceived stress (Turel et al., 2018). Unlike our measure of news consumption, these prior studies assess social media use in terms of addictive behavior and suggest there may be individual differences in the impact of social media content for individuals who demonstrate excessive social media usage.

Negative, discriminatory tweet stimuli in this study were emotionally arousing. Although tweets increased negative affect across multiple states of anger, depression, disgust, and fear, anger and disgust specifically contributed to RT effects. Emotional information, especially that which is perceived as a threat,

is prioritized and subject to processing regardless of the limited availability of attentional resources (LeDoux and Brown, 2017; Adolphs, 2018). Attentional capture is particularly problematic during emotional experiences of anger, which are known to impair decision making (Lerner and Tiedens, 2006). Marginalized groups may be more likely to be negatively impacted by emotionally arousing online content given disproportionate exposure to harassment and subsequent chilling effects on potentially positive social media engagement (see Brady and Crockett, 2019).

There are several mechanisms by which emotionally charged information can interfere with executive resources. First, threats elicit attempts to regulate negative emotions, taxing resources like the dlPFC through implicit and automatic emotion regulation (Braunstein et al., 2017). No explicit regulatory instructions were given in this study, implicating the dlPFC even when selection of regulatory techniques is implicit, voluntary, and heterogeneous. Second, emotional content is associated with increased attentional capture (Brady et al., 2020) and the dlPFC is involved in allocation of attentional resources (Sagliano et al., 2016). Attentional distraction by engaging with affectively arousing information online may be a mechanism through which executive functioning is impaired. Third, increased physiological stress as a result of negative emotional arousal can directly affect prefrontal processing through activation in the hypothalamic-pituitary-adrenal (HPA) axis (Arnsten, 2009). The prefrontal cortex in particular is sensitive to stress with evidence of impairments on spatial working memory function (Mika et al., 2012). Fourth, discriminatory content can make stereotypes about one's identity group salient consequently increasing threat and performance monitoring (Schmader et al., 2008). Although we did not explicitly activate stereotypes with respect to the MR task in this study, the negative tweets consisted of attacks against historically marginalized groups. However, the negative tweets included threats to multiple identity groups (gender, race, immigration status) meaning that not all tweets would activate personal stereotypes. Future work would benefit from disaggregating negative affect generally and that generated by stimuli of personal relevance. These results demonstrate that emotional stimuli can capture and divert attentional resources even in the absence of experimentally-manipulated salience. Lastly, work on emotional contagion via social media suggests individuals are more likely to adopt positive or negative emotions if that emotion valence is over-expressed in their social network (Kramer et al., 2014). It is possible that compounding effects exist for a subset of individuals already vulnerable to negative emotion exposure. Future research should explore these potential mechanisms to determine the extent to which they make distinct and interrelated contributions to neural functioning in the context of social media.

Results should be interpreted within constraints of the study design. Although this paradigm involved arousal of negative affect, we make no claims about the exclusivity of the effects to the regulation of negative affect. Regulation of positive affect also relies on cognitive resources (Seo et al., 2014) and poor control of positive affect can lead to problematic behaviors like impulsivity. Cognitive effects of positive affect induction is an interesting question for future research. Despite that these results should generalize to other tasks tapping executive functioning, the present study did not probe multiple executive domains. A spatial reasoning task was selected given spatial reasoning skills have been linked to success in Science, technology, engineering, and mathematics domains (STEM; Wai et al., 2009). This association is particularly relevant for disadvantaged individuals, like those included in this study,

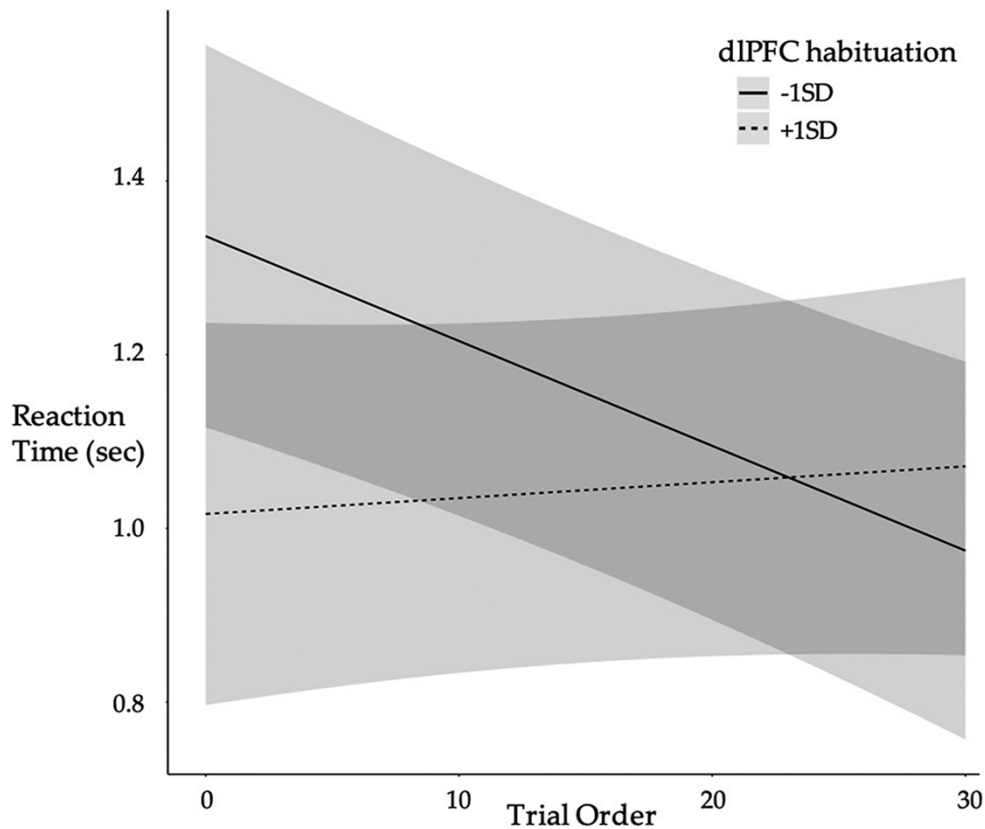


Fig. 7. Individuals with greater (+1SD) dorsolateral prefrontal cortex (dlPFC) habituation demonstrated no RT improvement over the course of the MR post-test whereas those with less (-1SD) habituation showed RT improvements over the course of the post-test, controlling for Condition and affect changes. Lines depict predicted values (marginal effects) for the regression model with 95% confidence interval bands shown in gray. $N = 57$.

who are likely to be most affected by negative emotional distraction from discriminatory content on social media and who are also underrepresented in STEM. Two primary alternative explanations may be responsible for the observed effects: task frustration and mental fatigue. However, examination of pre-tweet performance and post-task reported motivation lends support for our interpretation of emotional interference on cognitive performance. Future studies may probe interaction effects of emotional interference and motivation by providing incentives for performance. We told participants they would be asked about the tweet stimuli in order to ensure the stimuli were attended to, which may have resulted in active maintenance. Prior work implicates active maintenance of unpleasant stimuli in reduction of dlPFC activation (Perlstein et al., 2002). A young adult sample was chosen because of political identity development during this time (Schwartz et al., 2013). Cognitive interference is likely age-variant and should be explored in a larger sample spanning different developmental phases. There were some qualitative differences in negative and neutral tweets including length and personal relevance. Positive affect reports suggest the neutral tweets were not less engaging despite their shorter length and generic subject matter, but that they existed on a different point on the valence continuum.

These results contribute to existing knowledge by providing new evidence for the effect of social media content on neural and cognitive functioning in young adults. Behavioral data in tandem with fMRI indicate negative, emotionally-charged information temporarily impairs executive function. Because of the frequency with which younger individuals engage with

social media, there is significant potential for daily exposure to content that can limit cognitive and neural resources available for executive control, particularly in those undergoing the final stretch of brain development through the mid-20s. The detrimental effect of this cognitive interference is exacerbated by the fact that emotional online content, particularly that which evokes negative affect, is prone to wider sharing thereby compounding attentional capture (Brady et al., 2020). These findings have important practical implications for arbitrating discriminatory online content that may have deleterious effects on cognition for historically marginalized identity groups.

Funding

This work was supported by a grant from the Russell Sage Foundation (G-6711) and the Jeffrey/Wenzel Chair in Behavioral Neuroscience to A.G. Preparation of this manuscript was supported in part by the National Science Foundation Graduate Fellowship to S.M.T.

Conflict of interest statement: The authors declare no competing financial interests.

Author contributions

All authors contributed to developing study concept and study design. Data collection, analyses, and interpretation were performed by S.M.T. under the supervision of A.G. S.M.T. drafted

with manuscript and A.G. provided critical revisions. All authors approved the final version of the manuscript for submission.

References

- Adolphs, R. (2018). *The Neuroscience of Emotion: A New Synthesis*, Princeton: Princeton University Press.
- Arnsten, A.F.T. (2009). Stress signaling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, **10**, 410–22.
- Bartońg, K. (2019). *MuMIn: Multi-model inference*. R package version 1.43.6. <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.
- Beckmann, C.F., Jenkinson, M., Smith, S.M. (2003). General multilevel linear modeling for group analysis in fMRI. *NeuroImage*, **20**, 1052–63.
- Brady, W.J., Crockett, M.J. (2019). How effective is online outrage? *Trends in Cognitive Sciences*, **23**, 79–80.
- Brady, W.J., Gantman, A.P., Van Bavel, J.J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, **149**, 746–56.
- Braunstein, L.M., Gross, J.J., Ochsner, K.N. (2017). Explicit and implicit emotional regulation: a multi-level framework. *Social Cognitive and Affective Neuroscience*, **12**, 1545–57.
- Brunoni, A.R., Vanderhasselt, M.A. (2014). Working memory improvement with non-invasive brain stimulation of the dorsolateral prefrontal cortex: a systematic review and meta-analysis. *Brain & Cognition*, **86**, 1–9.
- Cromheeke, S., Mueller, S.C. (2014). Probing emotional influences on cognitive control: an ALE meta-analysis of cognition emotion interactions. *Brain Structure and Function*, **219**, 995–1008.
- Dolcos, F., McCarthy, G. (2006). Brain systems mediating cognitive interference by emotional distraction. *Journal of Neuroscience*, **26**, 2072–9.
- Dolcos, F., Iordan, A.D., Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: a review of evidence from brain imaging investigations. *Journal of Cognitive Psychology*, **23**, 669–94.
- Earle, F., Hockey, B., Earle, K., Clough, P. (2015). Separating the effects of task load and task motivation on the effort-fatigue relationship. *Motivation and Emotion*, **39**, 467–76.
- Ekman, P. (1999). Basic emotions. In: Dalgleish, T., Power, M., editors. *Handbook of Cognition and Emotion*, Sussex, UK: John Wiley & Sons, Ltd.
- Ganis, G., Kievit, R.A. (2015). A new set of three-dimensional shapes for investigating mental rotation processes: validation data and stimulus set. *Journal of Open Psychology Data*, **3**, e3.
- Greve, D.N., Fischl, B. (2009). Accurate and robust brain image alignment using boundary based registration. *NeuroImage*, **48**, 63–72.
- Iordan, A.D., Dolcos, F. (2017). Brain activity and network interactions linked to valence-related differences in the impact of emotion distract. *Cerebral Cortex*, **27**, 731–49.
- Kohn, N., Eickhoff, S.B., Scheller, M., Laird, A.R., Fox, P.T., Habel, U. (2014). Neural network of cognitive emotion regulation – an ALE meta-analysis and MACM analysis. *NeuroImage*, **87**, 345–55.
- Kramer, A.D., Guillory, J.E., Hancock, J.T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, **111**, 8788–90.
- Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, **82**, 1–26.
- LeDoux, J.E., Brown, R.A. (2017). Higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences*, **114**, E2016–25.
- Lerner, J.S., Tiedens, L.Z. (2006). Portrait of the angry decision maker: how appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making*, **19**, 115–37.
- McNair, D.N., Lorr, M., Droppleman, L.F. (1971). *Manual for the Profile of Mood States*, San Diego, CA: Educational and Industrial Testing Services.
- Meshi, D., Elizarova, A., Bender, A., Verdejo-Garcia, A. (2019). Excessive social media users demonstrate impaired decision making in the Iowa gambling task. *Journal of Behavioral Addictions*, **8**, 169–73.
- Mika, A., Mazur, G.J., Hoffman, A.N., Talboom, J.S., Bimonte-Nelson, H.A., Sanabria, F., Conrad, C.D. (2012). Chronic stress impairs prefrontal cortex-dependent response inhibition and spatial working memory. *Behavioral Neuroscience*, **126**, 605–19.
- Mumford, J.A., Poline, J.B., Poldrack, R.A. (2015). Orthogonalization of regressors in fMRI models. *PLoS One*, **10**, e0126255.
- Ochsner, K.N., Silvers, J.A., Buhle, J.T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, **1251**, E1–24.
- Perlstein, W.M., Elbert, T., Stenger, V.A. (2002). Dissociation in human prefrontal cortex of affective influences on working memory-related activity. *Proceedings of the National Academy of Sciences*, **99**, 1736–41.
- Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences*, **13**, 160–6.
- Sagliano, L., D'Olimpio, F., Panico, F., Gagliardi, S., Trojano, L. (2016). The role of the dorsolateral prefrontal cortex in early threat processing: a TMS study. *Social Cognitive and Affective Neuroscience*, **11**, 1992–8.
- Schmader, T., Johns, M., Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, **115**, 336–56.
- Schwartz, S.J., Zamboanga, B.L., Luyckx, K., Meca, A., Ritchie, R. (2013). Identity in emerging adulthood: reviewing the field and looking forward. *Emerging Adulthood*, **1**, 96–113.
- Schweizer, S., Satpute, A.B., Atzil, S., et al. (2019). The impact of affective information on working memory: a pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, **145**, 566–609.
- Seo, D., Olman, C.A., Haut, K.M., Sinha, R., MacDonald, A.W., Patrick, C.J. (2014). Neural correlates of preparatory and regulatory control over positive and negative emotion. *Social Cognitive and Affective Neuroscience*, **9**, 494–504.
- Shear, M.D., Haberman, M., Confessore, N., Yourish, K., Buchanan, L., Collins, K. (2019). How trump reshaped the presidency in over 11,000 tweets. *The New York Times*. Retrieved from <https://www.nytimes.com/interactive/2019/11/02/us/politics/trump-twitter-presidency.html>.
- Stone, J., McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology*, **44**, 445–52
- Tashjian, S.M., Galván, A. (2018). The role of mesolimbic circuitry in buffering election-related distress. *Journal of Neuroscience*, **38**, 2887–98.

- Turel, O., Cavagnaro, D.R., Meshi, D. (2018). Short abstinence from online social network sites reduces perceived stress, especially in excessive users. *Psychiatry Research*, *270*, 947–53.
- Twitter, Inc. Q4 and Fiscal Year 2018 Letter to Shareholders. (2019). Retrieved from: https://s22.q4cdn.com/826641620/files/doc_financials/2018/q4/Q4-2018-Shareholder-Letter.pdf
- Vohs, K.D., Baumeister, R.F., Schmeichel, B.J. (2012). Motivation, personal beliefs, and limited resources all contribute to self-control. *Journal of Experimental Social Psychology*, *48*, 943–7.
- Wai, J., Lubinski, D., Benbow, C.P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817–35.
- Woolrich, M. (2008). Robust group analysis using outlier inference. *NeuroImage*, *41*, 286–301.