OXFORD

# Exposure to negative stereotypes influences representations of monetary incentives in the nucleus accumbens

B. Locke Welborn,[1,2] Youngki Hong,[1] and Kyle G. Ratner[1,]

[1]Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106, USA, and [2]School of Psychology, University of East Anglia, Norwich, NR4 LTJ, UK

Correspondence should be addressed to B. Locke Welborn or Kyle. G Ratner. E-mail: b.welborn@uea.ac.uk or kyle.ratner@psych.ucsb.edu.

## Abstract

Contemporary society is saturated with negative representations of racial and ethnic minorities. Social science research finds that exposure to such negative stereotypes creates stress above and beyond pre-existing effects of income inequality and structural racism. Neuroscience studies in animals and humans show that life stress modulates brain responses to rewards. However, it is not known whether contending with negative representations of one's social group spills overs to influence reward processing. We used functional magnetic resonance imaging to examine the effects of stigmatizing negative stereotypes on neural responding to the anticipation and consumption of monetary gains and losses in a Mexican American sample. Machine learning analyses indicated that incentive-related patterns of brain activity within the nucleus accumbens differed between Mexican Americans subjected to negative stereotypes and those who were not. This effect occurred for anticipating both gains and losses. Our work suggests that rhetoric stigmatizing Latinos and other minorities could alter how members of such groups process incentives in their environment. These findings contribute to our understanding of the linkage between stigmatizing experiences and motivated behavior with implications for well-being and health.

Key words: stereotypes; stigma; incentive processing; reward; fMRI

## Introduction

Political and social divisions, amplified by Internet commentary and social media, have saturated popular culture with unflattering depictions of racial and ethnic minorities. Cable and network news have disproportionately portrayed African Americans as perpetrators of crimes and White Americans as law enforcement officers or victims (Dixon and Linz, 2000a, 2000b; Gilliam and Iyengar, 2000). Meanwhile, Latinos have been overrepresented in critical discussions about undocumented immigrants (Wilson *et al.*, 2003) and are almost twice as likely as Whites to appear in newspaper articles about social problems (Dixon and Williams, 2015). Moreover, these biased patterns of representation are not limited to traditional media (Dixon *et al.*, 2019). Latinos are often depicted as conforming to negative ethnic stereotypes on YouTube, and content with stereotypical depictions of Racial and ethnic groups receives more views and comments

(Guo and Harlow, 2014). On Twitter, Latinos are less visible than other groups, which likely hinders their ability to counteract falsehoods spread about them (Sui and Paul, 2017).

Such misrepresentations in media are rooted in historically ingrained stereotypes that devalue and derogate racial and ethnic minorities (Dixon, 2008). Understandably, these caricatures make many minority group members feel that they are viewed poorly by society at large (Steele, 1997; Crocker, 1999). As a result, encountering negative stereotypes can be a source of stigma-related stress for minority individuals (Major and O'Brien, 2005). Substantial work has examined detrimental effects of stigma-related stress on performance in a stigmatized domain (Steele, 1997; Krendl et al., 2008; Forbes et al., 2018). However, little research has examined lingering effects of stereotype exposure on unrelated neural and psychological processes. In the present research, we bring together methods and theorizing from social psychology and affective and clinical neuroscience to examine whether stigmatizing negative stereotypes in the media have 'spill-over' effects on neural systems related to incentive processing and motivated behavior.

### The deflating effect of negative stereotypes

Members of disadvantaged social groups not only contend with stressors that people in non-disadvantaged groups face (e.g. illness, the loss of a loved one, relationship conflicts), but they also confront stressors that are unique to their place in the social hierarchy (Major et al., 2013). Many of these unique stressors emerge from structural inequities in legal, healthcare and financial systems that bias the distribution of socioeconomic resources (Williams and Mohammed, 2009). However, social devaluation also creates less tangible psychological stressors, such as stigmatizing negative stereotypes, personal experiences with discrimination and expectations of bias, which can strain physiology and harm mental health (Mays et al., 2007; Ratner et al., 2013).

Stigmatization via media sources is particularly pernicious because it can spread widely and affect members of targeted groups who have not personally experienced discrimination (Davies et al., 2002; Saleem and Ramasubramanian, 2019). Many members of disadvantaged groups are able to maintain positive self-esteem and group image in the face of stigma by attributing away specific instances of discrimination (Crocker and Major, 1989). However, continually managing self-definition and avoiding psychological harm from stigma is cognitively and emotionally taxing (Schmader and Johns, 2003; Reynolds et al., 2010). The toll that these racial stressors take on African Americans and Latinos in the United States has been called 'racial battle fatigue' (Franklin et al., 2014) and compared to feelings of resignation and withdrawal that occur when people experience other persistent stressors that are beyond their direct control (Branscombe et al., 1999).

### Stress and disrupted neural processing of rewards

The 'weathering' effects of stress are well-recognized by researchers who study depression. Experiencing stressful life events is one of the most robust predictors of depression and is a risk factor for poor response to treatment and relapse (Tennant, 2002). Stress is particularly associated with anhedonic depression, which is characterized by a reduced sensitivity to reward (Pizzagalli, 2014). Considerable affective and clinical neuroscience research has examined the neural substrates of reward processing as well as the moderating effects of stress.

In humans, neuroscience research on reward processing has focused on the dopamine-rich nucleus accumbens (NAcc) in addition to the ventral medial prefrontal cortex (VMPFC) (Knutson et al., 2001; Dillon et al., 2008). Consistent with stress-related reward processing deficits in depression, anhedonic depression in adolescents is associated with reduced volume of the NAcc (Auerbach et al., 2017).

Critically, the NAcc and the VMPFC seem to have dissociable roles in reward processing. Berridge (2009) makes a distinction between anticipatory ('wanting') and consummatory ('liking') phases of reward processing. The anticipatory phase, which is thought to engage the NAcc, involves the incentive salience of a cue that signals the possibility of a reward. The consummatory phase, which engages both the NAcc and VMPFC, refers to the experience of receiving a reward. Surprisingly, only a few fMRI studies have directly examined how stress influences the anticipation and consumption of rewards. Dillon et al. (2009) reported that early life adversity was related to reduced basal ganglia activation during reward anticipation, consistent with stress dampening sensitivity to reward incentive cues. However, work with clinically depressed participants and also non-depressed individuals exposed to acute stress has reported blunted NAcc response during the consummatory phase (Pizzagalli et al., 2009; Kumar et al., 2014). Thus, it appears that stress spills over to influence the response of the NAcc to rewards, but whether this effect occurs at anticipation or consumption might depend on specific attributes of the stress experience.

## Current study

The present research examined whether stigma-related negative stereotypes act like other stressors to alter how the brain processes subsequent rewards. In part because of stigmatizing rhetoric toward Latinos in the United States, 63% of whom are Mexican American (Flores et al., 2017), we focused on stigma experienced by Mexican Americans. We specifically recruited college students in our research because higher education is a vehicle for social mobility, but Latino college students are often subjected to stigmatizing negative stereotypes that can undermine motivation and achievement (Smedley et al., 1993; Hwang and Goto, 2008; Huynh and Fuligni, 2010; Hout, 2012; Stephens et al., 2014).

In our research, Mexican American participants were randomly assigned to view either stigmatizing or non-stigmatizing video clips in rapid succession. To maximize ecological validity, all the video clips were from actual news stories or documentary-style recordings that were available on the Internet. To measure effects of the video clips on neural processing of the different phases of reward processing, participants completed a monetary incentive delay (MID) task (Knutson et al., 2000) in an MRI scanner (Figure 1). The MID task has been used widely in affective and clinical neuroscience to assess how the brain anticipates and responds to (i.e. consumes) monetary gains and losses (e.g. Samanez-Larkin et al., 2008; Kumar et al., 2014; Balodis and Potenza, 2015; Knutson and Heinz, 2015).

Based upon contemporary work on the neural impact of stressful experiences on incentive processing, we offer several hypotheses regarding the effects of the present experimental manipulation. We predicted that participants in the stigma condition would show reduced subjective arousal, which would provide evidence for a 'weathering' effect of repeated exposure to negative stereotypes. We also expected stigmatized participants, relative to non-stigmatized participants, to demonstrate

## A. Ethnicity-related Social Stigma Manipulation:
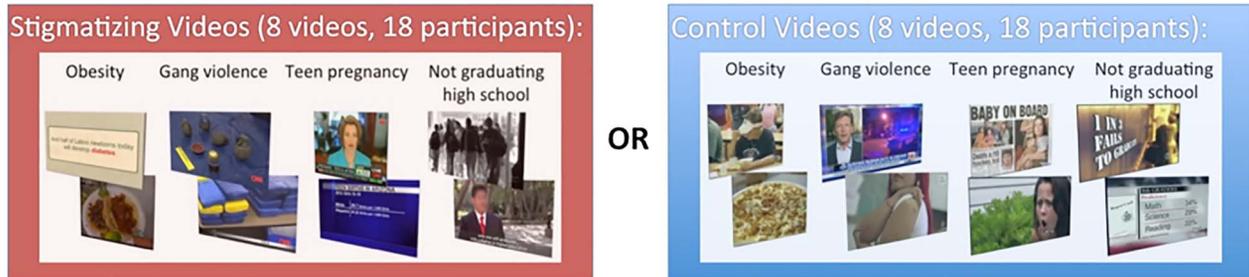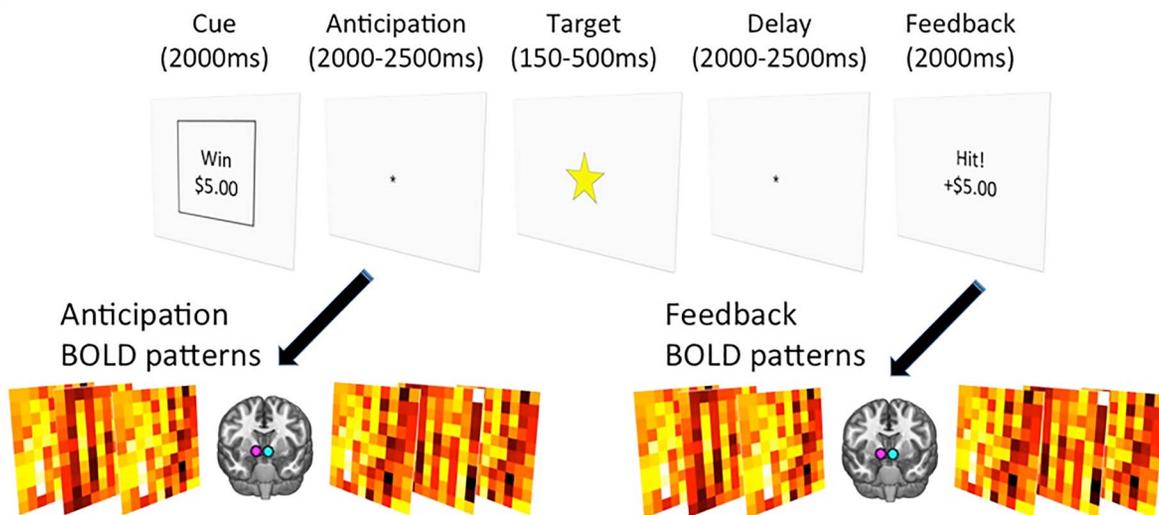


## B. Monetary Incentive Delay (MID) Task:



**Fig. 1**. Overview of experimental manipulation and monetary incentive delay (MID) task. (A) Screenshots from the video clips used as part of the experimental manipulation. Participants either saw the stigmatizing clips or the non-stigmatizing control clips. (B) Schematic of the MID Task. Analyses focused on BOLD patterns from the NAcc and VMPFC measured during the anticipation and feedback phases. The NAcc ROIs are shown.

lower overall NAcc activity when processing rewards. We further expected that patterns of activity within the NAcc would differ between the stigma and control conditions, which would suggest that representations of rewards can be influenced by recent exposure to negative stereotypes. We did not have *a priori* hypotheses about whether stigma-related NAcc effects would specifically influence the anticipation or consumption phases of reward processing because past research on stress and reward processing does not provide clear predictions in this regard. We also were not sure whether the stigma manipulation would increase self-reported feelings of negativity and subjugation because past research does not support conventional wisdom that stigma should undermine self-esteem and image about one's group (Crocker and Major, 1989).

## Methods

### Participants

Forty self-identified Mexican American participants were recruited from the psychology research subject pool at University of California, Santa Barbara (UCSB), and from other campus groups. All participants were right-handed and had normal or

corrected-to-normal vision. Exclusion criteria included self-reported psychiatric or neurological disorders, the use of psychoactive medications or drugs and metallic or electronic items embedded in the body that would render MR scanning unsafe. All participants were compensated $40 for their contribution to this research. Participants provided written informed consent approved by the UCSB Human Subjects Committee. Data from four participants were not included in analyses due to excessive motion (greater than 2 mm translation or 2 degrees rotation per run) or partial acquisition failure (final $N = 36$, 18 per experimental condition).

### Experimental manipulation

After completion of the informed consent process, participants were randomly assigned either to the stigma condition or to the control condition. This assignment determined the nature of the experimental manipulation to which participants were later subjected, but did not otherwise affect experimental procedures. The experimental manipulation in both conditions involved the presentation of eight brief (2–3 min) video clips focused on four topic areas: (i) childhood obesity, (ii) high school dropout rates, (iii) gang-related violence and (iv) teen pregnancy. These

topics were selected because they are issues of considerable social importance in contemporary American society and can be described either as predominately affecting Mexican Americans and other Latinos (consistent with common negative stereotypes) or as diffusely affecting American society as a whole. The video clips were introduced as stimulus materials unrelated to the MID task, which the experimenters wished to evaluate for use in future studies.

Video clips in the stigma condition focused on the prevalence of each social problem in the Latino community and the effects of these problems on Latinos specifically. Clips in this condition also emphasized negative social comparisons with other portions of American society, suggesting that Latinos were uniquely susceptible and disproportionately affected. Because these issues are associated with various stereotypes of Mexican Americans (Barreto *et al.,* 2012), presenting the clips in rapid succession was intended to create a stigma-related fatiguing sensation in our Mexican American participants. In contrast, video clips in the *control* condition focused on the prevalence of a given social problem and its deleterious effects within American society broadly construed, without focusing on any particular racial or ethnic community. Video clips in the stigma and control conditions were otherwise approximately matched for duration, content and tone. Importantly, while the content of the video clips was evocative in both conditions, in neither condition were the selected video clips polemical (i.e. video clips did not assign blame or strongly advocate specific policy solutions). Rather, the clips highlighted the pervasive nature of the social problems depicted and the adverse effects experienced by individuals (either within the Latino community in the United States or within American society in general). Prior to watching the video clips, and again before responding to questions on their personal experience with the depicted social problems, participants were reminded of their Mexican American (or American) identity.

In addition to watching the video clips, participants were asked to respond to a series of questions using a scanner-compatible button box. After being reminded of their Mexican American (or American) identity, participants indicated, for each topic, whether they had been affected by the associated social problem (e.g. 'Did you suffer from childhood obesity?') because they were Mexican American (stigma condition) or American (control condition). For each item, participants were provided with two response options: 'Yes' and "No/I'd rather not say." Thus, participants were afforded the opportunity to unambiguously affirm that they had been affected by the focal problem, but not afforded the opportunity to unambiguously *deny* that they had been affected. For participants in the stigma condition, this aspect of the manipulation prevented participants from distancing themselves from the stereotype-consistent content in the videos. Participants were also asked to state their current valence ('How positive/negative do you feel?'), arousal ('How calm/excited do you feel?') and dominance ('How dominated/dominant do you feel?') using on-screen Self-Assessment Manikin scales (scales ranged from 1 to 9, with higher values indicating more positive valence, more excited arousal and perceptions of the self as more dominant than dominated). These questions were intended to indirectly examine participants' self-reported reactions to the experimental manipulation they experienced.

In both conditions, clips were viewed at two time points (T1, T2). T1 occurred prior to the first functional run (during the acquisition of structural images) and T2 occurred prior to run four (halfway through the MID task). At each time point, participants viewed four video clips from their assigned condition (one from each topic category), for a total of eight presentations (two clips for each topic). No video clip was presented more than once per participant. Subject to these constraints, video clips were presented in a randomized order. Participants did not know the content areas of the videos beforehand, nor did they know how many video clips would be presented. All video clip stimulus materials are available upon request.

## Monetary incentive delay (MID) task

While undergoing functional magnetic resonance imaging (fMRI), participants completed an adapted version of the monetary incentive delay task introduced by Knutson *et al.* (2000). In this task, participants have the opportunity to win money or avoid losing money by making a rapid button-press response to a briefly presented visual target. Across 6 separate functional runs, participants completed a total of 180 trials divided into 6 different conditions (30 each). On win trials, a successful response would result in a monetary gain—whereas on lose trials, a successful response would avoid a possible monetary loss. On both types of trials, a button-press response was required, and failure to respond resulted in a worse outcome (failure to gain money for win trials and loss of money for lose trials). In addition, the amount of monetary reward/loss at stake varied across trials in three levels of compensation: $0, $1 and $5. Crossing the two levels of win/lose with the three levels of compensation resulted in six total conditions (e.g. win $5, lose $1, etc.). Trials from each condition were equally represented across the six functional runs and were presented in a pseudorandomized order designed to maximize contrast efficiency. Analyses focus on the anticipation phase (in which participants prepare to make the required button-press response) and the feedback phase (i.e. consummatory phase in which participants learn whether they have gained or lost money). See the Supplementary Materials for details regarding trial structure.

## FMRI data acquisition

All imaging data were acquired using a 3.0-Tesla Siemens Prisma scanner at the UCSB Brain Imaging Center. Across six functional runs, approximately 3420 T2∗-weighted echo-planar images (EPI) were acquired (∼570 per run) during completion of the MID task described above, with interleaved acquisitions using a multiband acceleration factor of 8 (slice thickness = 2 mm, gap = 0 mm, 72 slices, TR = 720 ms, TE = 37 ms, flip angle = 52°, matrix = 104 × 104, field of view = 208 mm, phase encoding: anterior-to-posterior). An oblique slice angle was used in order to minimize signal dropout in ventral medial portions of the brain. A gradient echo field mapping image was acquired for each participant (slice thickness = 2 mm, gap = 0 mm, 72 slices, TR = 758 ms, TE = 4.92 ms, flip angle = 60°, matrix = 104 × 104, field of view = 208 mm), with the same slice angle as the EPI images. We also acquired a T1-weighted magnetically prepared rapid acquisition gradient echo anatomical image (slice thickness = 0.94 mm, 208 slices, TR = 2500 ms, TE = 2.22 ms, flip angle = 7°, field of view = 241 mm) during presentation of the first set of video clips and a T2-weighted anatomical image (slice thickness = 0.94 mm, 208 slices, TR = 3200 ms, TE = 566 ms, field of view = 241 mm) during presentation of the second set of video clips. Lastly, a diffusion spectrum image (slice thickness = 1.8 mm, 78 slices, TR = 4300 ms, TE = 100.20 ms, flip angle = 90°, field of view = 259 mm) was acquired, but is not analyzed herein.

## FMRI data preprocessing

Structural and functional data were processed using SPM12 (Wellcome Department of Cognitive Neurology, London, UK) (Penny *et al.,* 2007). Within each functional run, image volumes were realigned to correct for head motion (using a rigid body transformation with 6 degrees of freedom) and unwarped (based upon the anatomical gradient field mapping images acquired for each participant). Images were subsequently segmented by tissue type (based upon the T1 and T2 images acquired for each participant) and normalized into standard MNI stereotactic space (resampled at 2x2x2mm). Finally, functional images were smoothed with a 6 mm Gaussian kernel, FWHM.

## Region of interest (ROI) definition for univariate and multivariate analyses

Region of interest (ROI) analyses were conducted to directly assess the recruitment of specific reward-related brain regions during the MID task. *A priori* ROIs were employed in order to test for group differences in hemodynamic response following the stigma or control video manipulation in regions associated with reward anticipation and consumption in prior studies. In addition, pattern classification techniques were deployed using data from these ROIs in order to determine whether differences in multivariate patterns of hemodynamic response following the stigma manipulation would be sufficient for accurate classification of experimental condition. For these purposes, *a priori* ROIs were derived from automated meta-analysis through www.neurosynth.org (Yarkoni *et al.,* 2011), using association test masks associated with the term 'reward.' *A priori* ROIs were derived from association test peak voxels in the right (MNI 12, 10, −8) and left (MNI -12, 10, −8) nucleus accumbens (NAcc), as well as in the ventral medial prefrontal cortex (VMPFC, MNI 2, 58, −8). Eight millimeter radius spheres were defined around each of these peaks and resliced into the 2x2x2 voxel space. An additional functional ROI was defined in early visual cortex based upon whole-brain analysis, as described in the Supplementary Materials. As this early visual cortex ROI was not specified *a priori*, analysis of data from this region must be considered exploratory.

For univariate analyses, parameter estimates from the models described above were extracted from all ROIs using MarsBaR (Brett *et al.,* 2002) for statistical comparisons. Parameter estimates from each participant's first-level models were evaluated at the group-level using general mixed-effects models implemented using the lme4 package (Bates *et al.,* 2015) in the R statistical software (http://www.r-project.org/). The primary factors that were considered in univariate analyses were the compensation level (zero or non-zero), win/lose framing of incentives and the experimental condition (stigmatized or non-stigmatized). For trials during the feedback phase, outcome (success or failure) was also considered. Sequential mixed-effects models were employed to avoid over-fitting data with too many explanatory variables, given the limited sample size.

## Univariate analytic approach

General linear models (GLMs) were defined for each participant, in which trials were modeled with separate functions corresponding to (i) the anticipation phase (corresponding to the fixation period prior to target presentation), (ii) the target phase (corresponding to the temporal window in which the target 'star' appeared and successful button presses occurred) and (iii) the feedback phase (corresponding to the final portion of each trial, in which success or failure feedback was presented to participants). The anticipation and target phases were modeled as variable-duration epochs (as their precise timing differed from trial to trial as well as across participants), while the feedback phase was modeled as a fixed duration epoch lasting 2 s. Each task phase was convolved with the canonical (double gamma) hemodynamic response function.

All models controlled for six estimated motion parameters (three translations and three rotations), as well as linear trends and differences between runs. The time series was high-pass filtered using a cutoff period of 128 s, and serial autocorrelations were modeled as an autoregressive AR(1) process. Individual-level statistics were aggregated for group-level comparisons and evaluated with a mixed-effects model. For whole-brain analyses, correction for multiple comparisons was implemented based upon Gaussian random field theory, in order to yield a cluster family-wise error (FWE) of $P < 0.05$ based upon an initial (voxel-wise) cluster-formation threshold of $P < 0.001$. See Supplementary Materials for further details.

## Multivariate analytic approach

In order to conduct multivariate pattern analysis of brain imaging data from the task, a series of beta images were first derived from univariate general linear models (Rissman *et al.,* 2004; see Supplementary Materials). Multivariate classification analyses were conducted using scikit-learn version 0.20.2 (Pedregosa *et al.,* 2011) pipelines employing three steps: (i) missing voxel imputation, (ii) standard scaling and (iii) classification using support vector machines (SVM) with a radial basis function (RBF) kernel. Importantly, each of the above steps was performed within cross-validation folds for each analysis so that there would be no data leakage from the training dataset to the testing dataset. Missing voxel imputation replaced missing values with the mean value from each voxel. Standard scaling involved mean-centering the data from each voxel and dividing by the voxel's standard deviation. Finally, SVM classification assigned a class label to each case (i.e. trial) within the dataset, using a radial basis function with default cost and gamma hyperparameters ($C = 1$ and gamma = $1/n$ features).

Cross-validation accuracy was assessed using a 'leave-one-pair-out' strategy. First, participants were randomly assigned to pairs for each classification, with one member of each pair in the stigma condition and one member in the control condition. As each of the 36 participants completed 180 trials (over the course of all functional runs), this procedure yields 18 cross-validation folds each of which includes a testing dataset with 360 ($180 \times 2$) cases equally divided between stigma and control participants. Importantly, to successfully classify cases (trials) in the testing dataset on each fold, the classification algorithm must generalize relationships between features (voxels) and stigma/control category labels, learned from the training set, to a test set consisting of data from participants that are entirely novel (for that cross-validation fold). A leave-one-subject-out cross-validation strategy could have been used but would have led to a maximally unbalanced distribution of class labels in the test dataset, as all trials for a single subject share a common stigma/control class. With the leave-one-pair-out strategy, class labels are balanced for both training and testing, and crucially, no participant is included in both training and testing datasets for any given fold. Separate randomized pairings were performed for each classification analysis, in order to minimize the possibility that specific pairings could help or hinder classification.

Besides distinguishing between stigma and control trials, other classification analyses were independently performed to distinguish between levels of compensation, whether trials afforded possibility of gain or of loss and (for feedback only) whether participants succeeded or failed. These classifications also employed the leave-one-pair-out cross-validation strategy for consistency, although in such cases this strategy is not necessary to ensure class balance. For multi-class classifications (i.e. with more than two categories), a one-vs-rest approach was used for the decision function.

Permutation testing was employed to determine whether cross-validation classifier accuracy differed significantly from chance. For each permutation, dataset class labels were randomly permuted within cross-validation pairs, and the classification algorithm was repeated. The $P$-value was then estimated from the proportion of permutations for which classifier accuracy exceeded the accuracy with non-permuted (true) labels (Ojala and Garriga, 2010).

## Results

### Subjective responses to the manipulation

The effects of the stigma manipulation were assessed at two time points: (T1) after the presentation of the first set of four video clips, prior to the start of the MID task, and (T2) after the presentation of the second set of four video clips, halfway through completion of the MID task (i.e. after three of six functional runs). On-screen queries assessed subjective experience related to arousal, valence and dominance (see *Methods* above). Valence and dominance did not show any differences across conditions at either time point ($Ps = 0.324$). However, at T1, participants in the stigma condition showed significantly reduced arousal relative to non-stigmatized controls ($t = -2.668$, $P = 0.0122$). Moreover, for participants in the stigma condition, arousal at T1 was significantly below the midpoint of the scale ($M = 3.72$ vs midpoint scale value 5; $t = -4.10$, $P = 0.0007$). For non-stigmatized controls at T1, arousal was slightly (non-significantly) above the midpoint ($M = 5.22$ versus 5; $t = 0.474$, $P = 0.641$). At T2, arousal for stigmatized participants was also below the scale midpoint, but this effect was only marginally statistically significant ($M = 4.11$; $t = -2.10$, $P = 0.0513$). For non-stigmatized controls, activation was slightly (non-significantly) below the midpoint at T2 ($M = 4.44$, $t = -1.317$, $P = 0.205$). Mean responses for the measures are plotted with standard errors of the mean in Figure 2A.

Results of self-reported affective responses suggest that stigmatizing media messages had an effect on the arousal of participants. Specifically, participants subjected to stigmatizing materials seem to have begun the MID task with lower arousal than participants who viewed non-stigmatizing materials, consistent with a fatiguing effect of stigma. Although at T2 there were not statistically significant arousal differences between the stigma and control participants, this effect seems to be driven by a reduction of arousal for the control participants at T2. It is possible that the controls reported less arousal at T2 compared to T1 simply because they had been in the scanner for a prolonged period of time. For the stigmatized participants, the T1 and T2 levels were not significantly different from each other, and both means were below the midpoint of the scale. It is unclear why there were no effects of the stigma condition on valence and dominance, but this result is consistent with work showing that people who experience stigma do not report lower self-esteem or regard for one's social group (Crocker and Major, 1989).

## Performance during the monetary incentive delay task

On average, participants performed slightly below the target two-third success rate, with an overall mean success of 62.1% across all runs of the MID task and all conditions, 95% CI [61.3%, 63.1%] (see Figure 2B). Overall performance did not depend upon the condition of the participants, with stigmatized participants' mean success rate at 62.7% and non-stigmatized participants' mean success rate at 61.7% ($t = -1.12$, $P = 0.27$). Aggregating across participants, a mixed-effects model with win/lose and compensation as fixed effects (and with participant as a random effect) found no significant effects (win/lose: $F = 0.942$, $P = 0.333$, compensation: $F = 0.492$, $P = 0.612$). The MID task script was coded so as to update (increment/decrement) the target response window based upon participant performance independently for each combination of win/lose and compensation levels, so no significant differences in performance based upon these factors indicate that the task operated as designed.

On average, participants responded during the target phase with a reaction time (RT) of 302 ms across all runs of the MID task and all conditions, 95% CI [0.285, 0.319] (see Figure 2C). Overall RT did not depend upon the condition of the participants, with stigmatized participants' mean RT at 298 ms and non-stigmatized participants' mean RT at 307 ms ($t = 0.510$, $P = 0.614$). A mixed-effects model with win/lose and compensation as fixed effects (and with participant as a random effect) was employed to assess the impact of these factors on reaction time. Win/lose did not have a significant influence on RT ($F = 1.583$, $P = 0.210$). However, the level of compensation had a strong effect on RT ($F = 28.629$, $P < 0.001$), with greater compensation associated with faster response. Models including interaction terms between win/lose and compensation, as well as between both factors and stigma condition, did not find any significant interactions (all $P > 0.05$). Overall, the correlation between mean RT and mean success rate was not significant ($r = -0.188$, $P = 0.278$). Together, the RT data suggest that participants in both conditions approached the task similarly.

## Evidence for the canonical NAcc/VMPFC responses to the MID task

Before examining the effects of our manipulation, we began by conducting a traditional univariate analysis in our NAcc and VMPFC ROIs. To our knowledge, this is the first time the MID task has been used with an exclusively Mexican American sample, so it is important to verify that NAcc and VMPFC show MID task responses expected from the literature in this subpopulation. Mixed-effects models (with participants as random effects) were employed to evaluate the effects of factors of interest on the BOLD signal in these regions during the anticipation and feedback phases. All runs were analyzed in general linear models together. All of these univariate results are reported in Table 1.

In general, the univariate analyses confirmed that NAcc and the VMPFC were sensitive to the magnitude of the incentives and feedback, results that are consistent with the MID literature and validated our use of this task (see Figure 3 and Supplementary Figures S3 and S4 for ROI effects as well as Supplementary Figures S1 and S2 and Supplementary Tables S1 and S2 for whole-brain effects). These results demonstrate that the canonical effects in the literature do generalize to a Mexican American sample.

Moreover, consistent with a blunting effect of negative stereotypes on incentive processing, the mean response in the left NAcc ROI was lower for the stigma vs control condition,
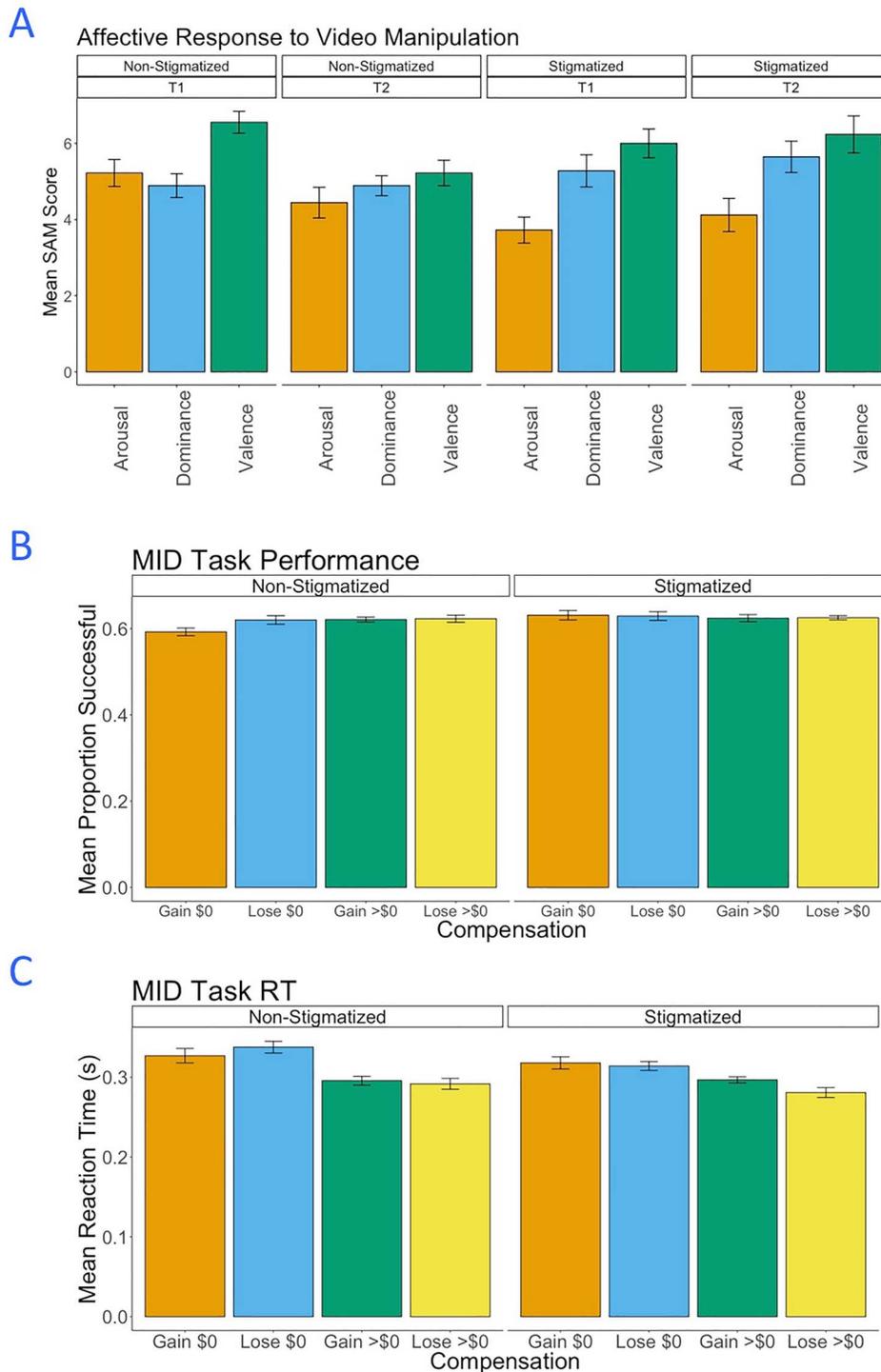
**Fig. 2**. (A) Mean Self-Assessment Manikin (SAM) responses for stigmatized and non-stigmatized participants at T1 and T2 (error bars indicate standard error of the mean). Decreased activation is evident for stigmatized participants but only at T1. (B) Performance on the MID task for stigmatized and non-stigmatized participants. Because reaction windows for successful response were adjusted adaptively during performance, success should approach two-third for all participants and trial types. (C) Reaction times for the MID task for both participant groups. Reaction times are faster with greater compensation at stake, but do not differ between groups.

although this effect was only marginally significant. Interestingly, this effect was moderated by magnitude of compensation but not win or loss frame and only during the anticipation phase. There were no effects of condition on the VMPFC response at either phase. Although the univariate analyses suggest that stigma may have influenced the NAcc response during reward anticipation, we are cautious about interpreting the condition effects because relevant *P*-values were marginally significant and only one exceeded the 0.05 significance level.

**Table 1.** Univariate models of activity during the MID task

| Region | Factor(s) | Phase | F | $\chi^2$ improvement | P |
|---|---|---|---|---|---|
| Left NAcc | Compensation | Anticipation | 5.310 | | 0.0220 |
| | Compensation + win/lose | | | 1.418 | 0.492 |
| | Compensation + time | | | 4.919 | 0.0854 |
| | Stigma condition | Anticipation | 3.983 | | 0.054 |
| | Stigma condition + compensation | | | 5.275 | 0.0216 |
| | Compensation | Feedback | 18.321 | | <0.001 |
| | Compensation + win/lose | | | 0.517 | 0.472 |
| | Compensation + stigma condition | | | 1.798 | 0.180 |
| | Outcome | Feedback | 24.423 | | <0.001 |
| | Outcome + win/lose | | | 5.695 | 0.0580 |
| | Outcome + stigma condition | | | 1.836 | 0.399 |
| Right NAcc | Compensation | Anticipation | 85.793 | | <0.001 |
| | Compensation + win/lose | | | 1.203 | 0.548 |
| | Stigma condition | Anticipation | 1.737 | | 0.197 |
| | Stigma condition + compensation | | | 2.857 | 0.240 |
| | Compensation | Feedback | 32.578 | | <0.001 |
| | Compensation + win/lose | | | 3.257 | 0.0711 |
| | Compensation + stigma condition | | | 5.959 | 0.0508 |
| | Outcome | Feedback | 76.849 | | <0.001 |
| | Outcome + win/lose | | | 6.656 | 0.0359 |
| | Outcome + stigma condition | | | 1.512 | 0.470 |
| VMPFC | Compensation | Anticipation | 4.297 | | 0.0406 |
| | Compensation + win/lose | | | 7.693 | 0.02114 |
| | Compensation | Feedback | 8.816 | | <0.001 |
| | Compensation + win/lose | | | 3.858 | 0.145 |
| | Compensation + stigma condition | | | 4.458 | 0.108 |
| | Outcome | Feedback | 10.550 | | 0.002 |
| | Outcome + win/lose | | | 17.715 | <0.001 |
| | Outcome + stigma condition | | | 3.504 | 0.173 |

Univariate responses of the right and left NAcc and the VMPFC during the MID task were analyzed using mixed-effects models with participant as a random factor. The impact of various factors, alone or in combination, was assessed for each region during the anticipation phase and the feedback phase. For base models (with only one factor), F-values and P-values are reported. For other models, improvements to model fit are reported with $\chi^2$ and P-values. See Results for details, as well as Figure 3 and Supplementary Figures S1–S4. NAcc, nucleus accumbens; VMPFC, ventromedial prefrontal cortex.

## Multivariate patterns differentiate stigmatized and non-stigmatized participants

Contemporary research has increasingly employed multivariate methods in the decoding and analysis of incentive-related brain processes (Kahnt, 2018). These methods are sensitive to differences in hemodynamic response patterns that vary within and between regions on a variety of different spatial scales. Importantly, research on reward processing using MVPA decoding suggests that multivariate patterns reflect more than the mere value of a stimulus: they can also encode information about the identity of the expected outcome, the actions necessary to achieve that outcome and contextual factors related to the appraisal of incentives. For this reason, we conducted a multivariate pattern classification analysis of incentive processing during anticipation and feedback, with the goal of determining whether information in patterns of activity in NAcc would differ based upon stereotype exposure. According to this analysis, patterns of activity within the NAcc significantly differentiated stigmatized and control participants during the critical anticipation period. This pattern classification result replicated in both the right and left NAcc, although accuracy was greater in the right NAcc (left NAcc accuracy = 55.5%, P < 0.0099, right NAcc accuracy = 59.8%, P < 0.0099, Figure 3). Whether the incentives were framed as gains or losses did not moderate these effects. To examine whether the pattern classification results at anticipation could be attributed specifically to incentive processing, we examined a region in auditory cortex as a control, given that the task was visual and not auditory in nature. The auditory cortex could not classify stigmatized from non-stigmatized participants during anticipation (accuracy = 50.2%, P < 0.406).

The results also appear to be temporally specific, emerging during anticipation of incentives but not during subsequent feedback. Pattern classification did not distinguish participants by condition during feedback (left NAcc accuracy = 48.6%, P < 0.956, right NAcc accuracy = 50.5%, P < 0.317). The pattern classification analysis of VMPFC voxels showed no significant effect of condition, either at anticipation or at feedback.

Although we focused our analyses on the NAcc and VMPFC because of empirical work connecting these regions to incentive processing, an exploratory univariate whole-brain analysis indicated that the early visual cortex also differed between the stigma and control conditions (Supplementary Figure S5). During the feedback phase, signal in the early visual cortex was higher in control versus stigmatized participants (right lingual gyrus, peak t = 5.088, cluster FWE P < 0.05, k=260; see Supplementary Table S3 and Supplementary Figure S5). Pattern classification analyses also indicated that this region of the early visual cortex could classify stigmatized from non-stigmatized participants at feedback (accuracy = 53.2%, P < 0.0099). This early visual cortex effect was not expected, so we are hesitant to interpret this result. It is possible, however, that stigmatized participants were not encoding feedback to the same extent as non-stigmatized
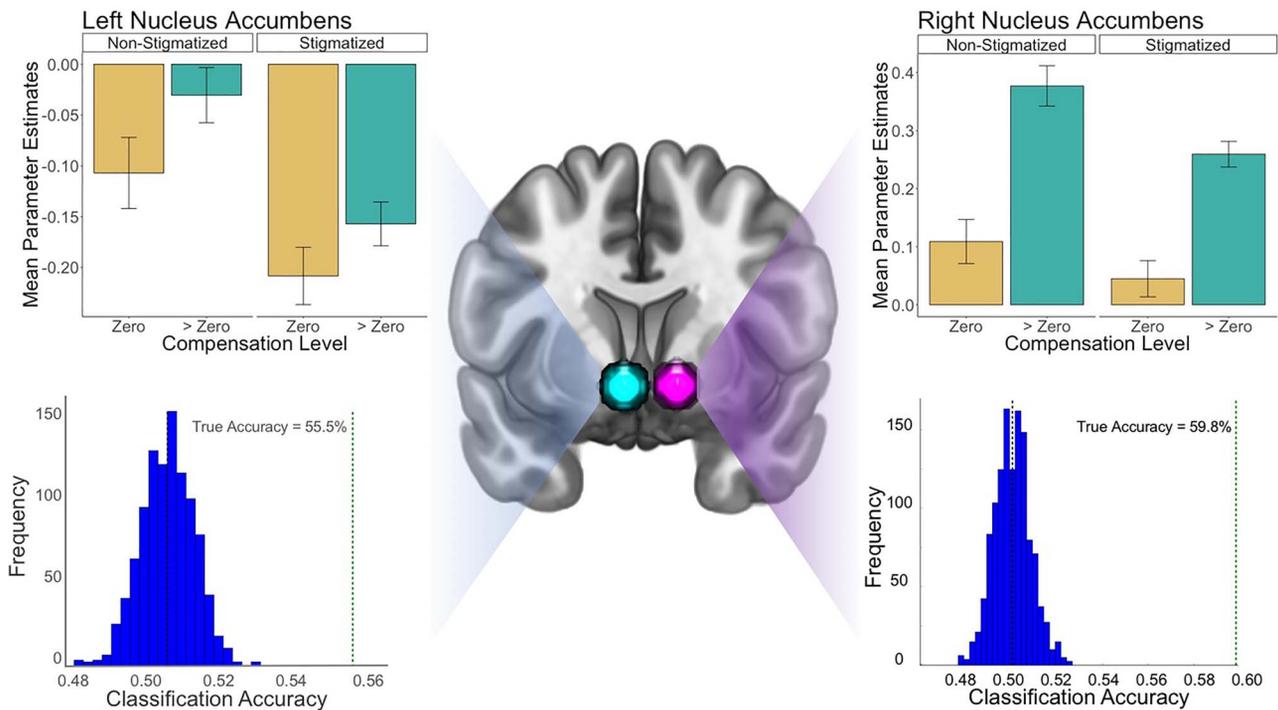
**Fig. 3**. Results from analyses in the left and right NAcc during anticipation. For both the left and right ROIs, the top graphs show the results from univariate analyses contrasting the stigmatizing and non-stigmatizing conditions. The bottom graphs show results from permutation tests of the multivariate pattern classification analysis, which indicate that stigmatization did influence representation of incentives in both NAcc ROIs. None of the permutations (histogram of accuracy scores with permuted condition labels in blue) achieved accuracy greater than or equal to the classification accuracy with the true labels (indicated by the green dotted line).

participants because of the potential threat-confirming nature of adverse feedback.

## Discussion

Our results provide evidence that exposure to negative stereotypes about one's group can spill over and influence neural processing of incentives and suggest that such effects may be both regionally and temporally specific. We found that exposure to negative stereotypes bilaterally influenced patterns of responding within the NAcc during the anticipation of incentives. Moreover, this finding was not moderated by gain vs loss frames. There was no evidence that the VMPFC response at anticipation or feedback was influenced by stereotype exposure. Given that clinical disorders, such as major depressive disorder, are characterized by dysfunctional incentive processing, our work suggests that stereotype exposure effects on the NAcc when anticipating incentives could be important for understanding mental health disparities experienced by negatively stereotyped groups.

It is particularly notable that the NAcc effects of stigma during the anticipation phase pertained to monetary incentives generally and were not specific to gains. Our research was motivated by existing work showing that stress blunts reward processing, so we did not predict loss trials to show similar effects as gain trials. In fact, loss trials were included in our design because they are a standard part of the MID task that we used. However, in hindsight, our results are consistent with a recent meta-analysis that found that the NAcc is responsive to both positively and negatively valenced incentives (Oldham *et al.,* 2018). Moreover, the finding that negative stereotype exposure influences NAcc activity to the anticipation of both gains and

losses is in line with theorizing that depression is characterized by a general disengagement from the environment, resulting in blunted sensitivity to both positive and negative stimuli (Rottenberg *et al.,* 2005). Contrary to our hypothesis, exposure to negative stereotypes did not robustly blunt the overall NAcc reactivity to incentivizing cues, although the magnitude of the response in this region was marginally lower in the stigma condition. While we hesitate to interpret non-significant results, it is possible that stronger evidence for the predicted affective blunting might be obtained using a considerably larger sample size and/or a more potent exposure to aversive stereotypes.

Conclusions regarding the generality of our effects should be tempered both with respect to the populations considered and the nature of stigmatizing stimuli. Our Mexican American student participants represent only one group of Latino demographic, albeit a demographic that is critical to closing socioeconomic gaps in America. Additionally, the stereotype exposure in this research represents only one type of ethnicity-related stressor that disadvantaged minorities experience. It is possible that stressors not investigated in this research, such as hate speech, elicit different affective responses from those reported here, and thus, the downstream effects on incentive processing might not be the same. It is also unknown how our participants' past experience with stigma and discrimination influenced their response to the stigma manipulation. Future large-scale studies powered for investigating individual differences would provide insight into how effects of negative stereotype exposure on neural processing of incentives are modulated by a variety of life experiences and coping strategies. Additional research should also explore whether non-Latino individuals demonstrate a neural response similar to that of

our Mexican American participants when exposed to negative stereotypes about Latinos. It is possible that non-Latinos with high egalitarian values would demonstrate similar responses as our participants; however, it is also possible that the responses would be different because of ethnicity-related differences in lived experience. Given these limitations, our work should be viewed simply as a proof of principle that negative stereotype exposure can influence representations of rewards and costs in the NAcc.

Although it is well established that stress influences incentive processing, the ways that marginalized groups cope with stigmatizing experiences and the impact of these experiences on incentive processing remain understudied. Incentive processing is crucial to many life outcomes, including financial decision-making, academic achievement, job performance, depressive disorders, drug addiction and health behaviors. Thus, delineating the association between stigmatization and incentive processing may help advance understanding of psychological mechanisms that perpetuate societal disadvantage.

## Supplementary data

Supplementary data are available at *SCAN* online.

## Acknowledgments

## Funding

## Conflict of interest

The authors declare no conflict of interest.

## References

Auerbach, R.P., Pisoni, A., Bondy, E., *et al.* (2017). Neuroanatomical prediction of anhedonia in adolescents. *Neuropsychopharmacology*, **42**(10), 2087–95.

Balodis, I.M., Potenza, M.N. (2015). Anticipatory reward processing in addicted populations: a focus on the monetary incentive delay task. *Biological Psychiatry*, **77**, 434–44.

Barreto, M.A., Manzano, S., Segura, G. (2012). *The impact of media stereotypes on opinions and attitudes towards Latinos*. Available: http://archive.nhmc.org/wp-content/uploads/2014/01/LD_NHMC_Poll_Results_Sept.2012.pdf

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Berridge, K.C. (2009). Wanting and liking: observations from the neuroscience and psychology laboratory. *Inquiry (Oslo, Norway)*, **52**(4), 378.

Branscombe, N.R., Schmitt, M.T., Harvey, R.D. (1999). Perceiving pervasive discrimination among African Americans: implications for group identification and well-being. *Journal of Personality and Social Psychology*, **77**(1), 135–49.

Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox [abstract] Presented at the 8th International Conference on Functional Mapping of the Human Brain. Available on CD-ROM in NeuroImage, **16**(2), 497.

Crocker, J. (1999). Social stigma and self-esteem: situational construction of self-worth. *Journal of Experimental Social Psychology*, **35**, 89–107.

Crocker, J., Major, B. (1989). Social stigma and self-esteem: the self-protective properties of stigma. *Psychological Review*, **96**(4), 608–30.

Davies, P.G., Spencer, S.J., Quinn, D.M., Gerhardstein, R. (2002). Consuming images: how television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, **28**(12), 1615–28.

Dillon, D.G., Holmes, A.J., Jahn, A.L., Bogdan, R., Wald, L.L., Pizzagalli, D.A. (2008). Dissociation of neural regions associated with anticipatory versus consummatory phases of incentive processing. *Psychophysiology*, **45**(1), 36–49.

Dillon, D.G., Holmes, A.J., Birk, J.L., Brooks, N., Lyons-Ruth, K., Pizzagalli, D.A. (2009). Childhood adversity is associated with left basal ganglia dysfunction during reward anticipation in adulthood. *Biological Psychiatry*, **66**, 206–13.

Dixon, T.L. (2008). Network news and racial beliefs: exploring the connection between national television news exposure and stereotypical perceptions of African Americans. *Journal of Communication*, **58**, 321–37.

Dixon, T.L., Linz, D. (2000a). Overrepresentation and underrepresentation of African Americans and Latinos as lawbreakers on television news. *Journal of Communication*, **50**, 131–54.

Dixon, T.L., Linz, D. (2000b). Race and the misrepresentation of victimization on local television news. *Communication Research*, **27**, 547–73.

Dixon, T.L., Williams, C.L. (2015). The changing misrepresentation of race and crime on network and cable news. *Journal of Communication*, **65**, 24–39.

Dixon, T.L., Weeks, K.R., Smith, M.A. (2019). Media constructions of culture, race, and ethnicity. *Oxford Research Encyclopedia of Communication*.

Flores, A., Lopez, G., Radford, J. (2017). *Facts on Latinos in America: current data*. Available: http://www.pewhispanic.org/2017/09/18/facts-on-u-s-latinos-current-data/

Forbes, C.E., Amey, R., Magerman, A.B., Duran, K., Liu, M. (2018). Stereotype-based stressors facilitate emotional memory neural network connectivity and encoding of negative information to degrade math self-perceptions among women. *Social Cognitive and Affective Neuroscience*, **13**, 719–40.

Franklin, J.D., Smith, W.A., Hung, M. (2014). Racial battle fatigue for Latina/o students: a quantitative perspective. *Journal of Hispanic Higher Education*, **13**(4), 303–22.

Gilliam, F.D., Iyengar, S. (2000). Prime suspects: the influence of local television news on the viewing public. *American Journal of Political Science*, **44**, 560–73.

Guo, L., Harlow, S. (2014). User-generated racism: an analysis of stereotypes of African Americans, Latinos, and Asians in YouTube videos. *Howard Journal of Communications*, **25**(3), 281–302.

Hout, M. (2012). Social and economic returns to college education in the United States. *Annual Review of Sociology*, **38**, 379–400.

Huynh, V.W., Fuligni, A.J. (2010). Discrimination hurts: the academic, psychological, and physical well-being of adolescents. *Journal of Research on Adolescence*, **20**, 916–41.

Hwang, W.C., Goto, S. (2008). The impact of perceived racial discrimination on the mental health of Asian American and Latino college students. *Cultural Diversity and Ethnic Minority Psychology*, **14**, 326–35.

Kahnt, T. (2018). A decade of decoding reward-related fMRI signals and where we go from here. *NeuroImage*, **180**, 324–33.

Knutson, B., Heinz, A. (2015). Probing psychiatric symptoms with the monetary incentive delay task. *Biological Psychiatry*, **77**, 418–20.

Knutson, B., Westdorp, A., Kaiser, E., Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage*, **12**, 20–7.

Knutson, B., Adams, C.M., Fong, G.W., Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience*, **21**(16), RC159.

Krendl, A.C., Richeson, J.A., Kelley, W.M., Heatherton, T.F. (2008). The negative consequences of threat: a functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science*, **19**, 168–75.

Kumar, P., Berghorst, L.H., Nickerson, L.D., *et al.* (2014). Differential effects of acute stress on anticipatory and consummatory phases of reward processing. *Neuroscience*, **266**, 1–12.

Major, B., O'Brien, L.T. (2005). The social psychology of stigma. *Annual Review of Psychology*, **56**, 393–421.

Major, B., Mendes, W.B., Dovidio, J.F. (2013). Intergroup relations and health disparities: a social psychological perspective. *Health Psychology*, **32**(5), 514–24.

Mays, V.M., Cochran, S.D., Barnes, N.W. (2007). Race, race-based discrimination, and health outcomes among African Americans. *Annual Review of Psychology*, **58**, 201–25.

Ojala, M., Garriga, G. (2010). Permutation tests for studying classifier performance. *Journal of Machine Learning Research*, **11**, 1833–63.

Oldham, S., Murawski, C., Fornito, A., Youssef, G., Yücel, M., Lorenzetti, V. (2018). The anticipation and outcome phases of reward and loss processing: a neuroimaging meta-analysis of the monetary incentive delay task. *Human Brain Mapping*, **39**, 3398–418.

Pedregosa, F., Varoquaux, G., Gramfort, A., *et al.* (2011). Scikit-learn: machine learning in python. *Journal of Machine Learning Research*, **12**, 2825–30.

Penny, W., Friston, K., Ashburner, J., Kiebel, S., Nichols, T. (2007). *Statistical parametric mapping: The analysis of functional brain images*, London: Academic Press.

Pizzagalli, D.A. (2014). Depression, stress, and Anhedonia: toward a synthesis and integrated model. *Annual Review of Clinical Psychology*, **10**, 393–423.

Pizzagalli, D.A., Holmes, A.J., Dillon, D.G., *et al.* (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated subjects with major depressive disorder. *The American Journal of Psychiatry*, **166**(6), 702–10.

Ratner, K.G., Halim, M.L., Amodio, D.M. (2013). Perceived stigmatization, ingroup pride, and immune and endocrine activity: evidence from a community sample of black and Latina women. *Social Psychological and Personality Science*, **4**(1), 82–91.

Reynolds, A.L., Sneva, J.N., Beehler, G.P. (2010). The influence of racism-related stress on the academic motivation of black and Latino/a students. *Journal of College Student Development*, **51**(2), 135–49.

Rissman, J., Gazzaley, A., D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, **23**, 752–63.

Rottenberg, J., Gross, J.J., Gotlib, I.H. (2005). Emotion context insensitivity in major depressive disorder. *Journal of Abnormal Psychology*, **114**(4), 627–39.

Saleem, M., Ramasubramanian, S. (2019). Muslim Americans' responses to social identity threats: effects of media representations and experiences of discrimination. *Media Psychology*, **22**(3), 373–93.

Samanez-Larkin, G.R., Hollon, N.G., Carstensen, L.L., Knutson, B. (2008). Individual differences in insular sensitivity during loss anticipation predict avoidance learning. *Psychological Science*, **19**, 320–3.

Schmader, T., Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, **85**, 440–52.

Smedley, B.D., Myers, H.F., Harrell, S.P. (1993). Minority-status stresses and the college adjustment of ethnic minority freshmen. *The Journal of Higher Education*, **64**, 434–52.

Steele, C.M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist*, **52**, 613–29.

Stephens, N.M., Hamedani, M.G., Destin, M. (2014). Closing the social-class achievement gap: a difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological Science*, **25**, 943–53.

Sui, M., Paul, N. (2017). Latino portrayals in local news media: underrepresentation, negative stereotypes, and institutional predictors of coverage. *Journal of Intercultural Communication Research*, **46**(3), 273–94.

Tennant, C. (2002). Life events, stress and depression: a review of recent findings. *The Australian and New Zealand Journal of Psychiatry*, **36**(2), 173–82.

Williams, D.R., Mohammed, S.A. (2009). Discrimination and racial disparities in health: evidence and needed research. *Journal of Behavioral Medicine*, **32**(1), 20.

Wilson, C.C., II, Gutierrez, F., Chao, L. (2003). *Racism, Sexism, and the Media: The Rise of Class Communication in Multicultural America*, Thousand Oaks, CA: SAGE Publications.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, **8**, 665–70.