Self-Affirmation Activates the Ventral Striatum: A Possible Reward-Related Mechanism for Self-Affirmation

Janine M. Dutcher[1], J. David Creswell[2], Laura E. Pacilio[2], Peter R. Harris[3], William M. P. Klein[4], John M. Levine[4], Julienne E. Bower[1,5,6], Keely A. Muscatell[7], and Naomi I. Eisenberger[1]

[1]Department of Psychology, University of California, Los Angeles; [2]Department of Psychology, Carnegie Mellon University; [3]School of Psychology, University of Sussex; [4]Department of Psychology, University of Pittsburgh; [5]Department of Psychiatry and Biobehavioral Science, University of California, Los Angeles; [6]Semel Institute for Neuroscience and Human Behavior, Cousins Center for Psychoneuroimmunology, University of California, Los Angeles; and [7]Robert Wood Johnson Foundation Health and Society Scholars Program, University of California, San Francisco, and University of California, Berkeley

Abstract

Self-affirmation (reflecting on important personal values) has been shown to have a range of positive effects; however, the neural basis of self-affirmation is not known. Building on studies showing that thinking about self-preferences activates neural reward pathways, we hypothesized that self-affirmation would activate brain reward circuitry during functional MRI (fMRI) studies. In Study 1, with college students, making judgments about important personal values during self-affirmation activated neural reward regions (i.e., ventral striatum), whereas making preference judgments that were not self-relevant did not. Study 2 replicated these results in a community sample, again showing that self-affirmation activated the ventral striatum. These are among the first fMRI studies to identify neural processes during self-affirmation. The findings extend theory by showing that self-affirmation may be rewarding and may provide a first step toward identifying a neural mechanism by which self-affirmation may produce a wide range of beneficial effects.

Self-affirmation—the process of reflecting on important personal values or personal characteristics and strengths—has been shown to have a broad range of benefits in hundreds of studies (for reviews, see Cohen & Sherman, 2014; Sherman & Cohen, 2006). For example, self-affirmation has been shown to reduce defensiveness (Sherman, Nelson, & Steele, 2000) and stress (Creswell et al., 2005) and to improve academic outcomes (e.g., grade point average, problem-solving performance; Cohen, Garcia, Apfel, & Master, 2006; Creswell, Dutcher, Klein, Harris, & Levine, 2013). Self-affirmation has also been shown to have a range of positive effects on social and affective behavior, including improved self-control after rejection (Burson, Crocker, & Mischkowski, 2012), increased well-being (Nelson, Fuller, Choi, & Lyubomirsky, 2014), reduced rumination (Koole, Smeets, Van Knippenberg, & Dijksterhuis, 1999), and enhanced feelings of relational security (Stinson, Logel, Shepherd, & Zanna, 2011). However, we know little about the underlying neural mechanisms.

Recent research has focused on building mechanistic accounts of self-affirmation; studies have suggested that

Corresponding Author:
Janine M. Dutcher, Department of Psychology, University of California, Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563
E-mail: jdutcher@ucla.edu

self-affirmation triggers a psychological cascade of effects, such as increasing attention to threat, broadening perspective, increasing feelings of social connection, and enhancing coping resources (e.g., Correll, Spencer, & Zanna, 2004; Crocker, Niiya, & Mischkowski, 2008; Klein & Harris, 2009; Sherman & Hartson, 2011). In a recent review, Cohen and Sherman (2014) noted there could be multiple mechanisms for self-affirmation and using a variety of methods could illuminate these processes. However, most accounts have not identified the basic neural and cognitive processes leading to these psychological effects. In particular, no published research to date has considered which neural regions are involved during self-affirmation (although some work has examined the neural consequences of self-affirmation; Falk et al., 2015; Legault, Al-Khindi, & Inzlicht, 2012). Using neuroimaging as a tool to investigate the self-affirmation process could help advance self-affirmation theory, given that neuroimaging does not rely on self-report.

In two studies, we proposed and tested a novel self-affirmation neural-reward account. Specifically, we posited that engaging in self-affirmation activates not only self-related neural regions (medial prefrontal cortex, precuneus) but also neural reward pathways in the brain's mesolimbic dopamine system. Our proposed account was informed by studies showing that neural reward regions, such as the ventral striatum (VS) and the ventral tegmental area (VTA), are activated when people disclose self-traits or beliefs (Tamir & Mitchell, 2012). The VS is also more active when participants think about positive, compared with neutral, autobiographical memories (Speer, Bhanji, & Delgado, 2014). These studies provide evidence that thinking about positive aspects of the self may activate neural reward pathways.

This reward account may provide a mechanistic explanation for self-affirmation's effects on threat and stress responding, given that previous research has found that rewarding stimuli (e.g., sexual stimuli, sucrose) decrease physiological stress responses in humans (Creswell, Pacilio, Denson, & Satyshur, 2013) and rats (Ulrich-Lai et al., 2010). It is possible that, when participants perform a self-affirmation task, their reward-related neural activity increases, which diminishes their neural responses to threat, allowing them to be more resilient and open to self-related threats compared with nonaffirming participants.

We conducted two neuroimaging studies to explore neural activity specific to the process of self-affirmation; in particular, we examined whether self-affirmation led to increased activity in neural reward regions (VS and VTA) relative to nonaffirmation (in Study 1) or control (in Study 2).

In Study 1, college-age participants were randomly assigned to either a self-affirmation group, in which they made decisions about important personal values (experimental condition) and decisions about less important personal values (control condition), or a nonaffirmation group, in which they made decisions about attributes other people might prefer in toasters (experimental condition) and decisions about less important toaster attributes (control condition). Thus, this study used a 2 × 2 mixed design, with group (self-affirmation, nonaffirmation) manipulated between subjects and ranking of value or attribute (high, low) manipulated within subjects. We included a nonaffirmation group to evaluate whether neural reward activity was due specifically to self-affirmation or to the difference in value between the experimental conditions (important personal values or toaster attributes) and control conditions (less important personal values or toaster attributes). In Study 2, we tested whether these effects extended to a community sample of older adult women.

Study 1

Method

Participants. Forty university students (18 female; mean age = 24.13 years, *SD* = 5.72) completed study procedures. Data collection was stopped after each group reached a minimum of 20 participants, but data from 2 participants in the nonaffirmation group were not saved after the completion of the imaging session, so that group had 18 participants with usable data. All participants met eligibility criteria for functional MRI (fMRI) studies (i.e., they were right-handed, not claustrophobic, free of implanted metal, and not pregnant). Fifty-nine percent were White, 3% were Hispanic, 11% were Black, 11% were Asian American or Asian, and 16% were of "other" race. The Carnegie Mellon University institutional review board approved all study procedures.

Procedure. Before the scanning session, participants were randomly assigned to either the self-affirmation group (*n* = 20) or the nonaffirmation group (*n* = 18). We used a standard self-affirmation decision-making task (Steele, 1988; Steele & Liu, 1983); participants were given a series of paired personal-value statements and were asked to indicate their relative preference (adapted from Vernon & Allport, 1931; see Fig. 1). In the self-affirmation group, participants ranked five personal values (art, religion, science, social issues, politics) in order of importance. We used this list to create a scanner task that was specific to each participant's own personal values. In the nonaffirmation group, participants were asked to rank a list of five toaster attributes (e.g., slice capacity, color, size) in the order of the importance that they believed an average college student would rank them (for frequency tables
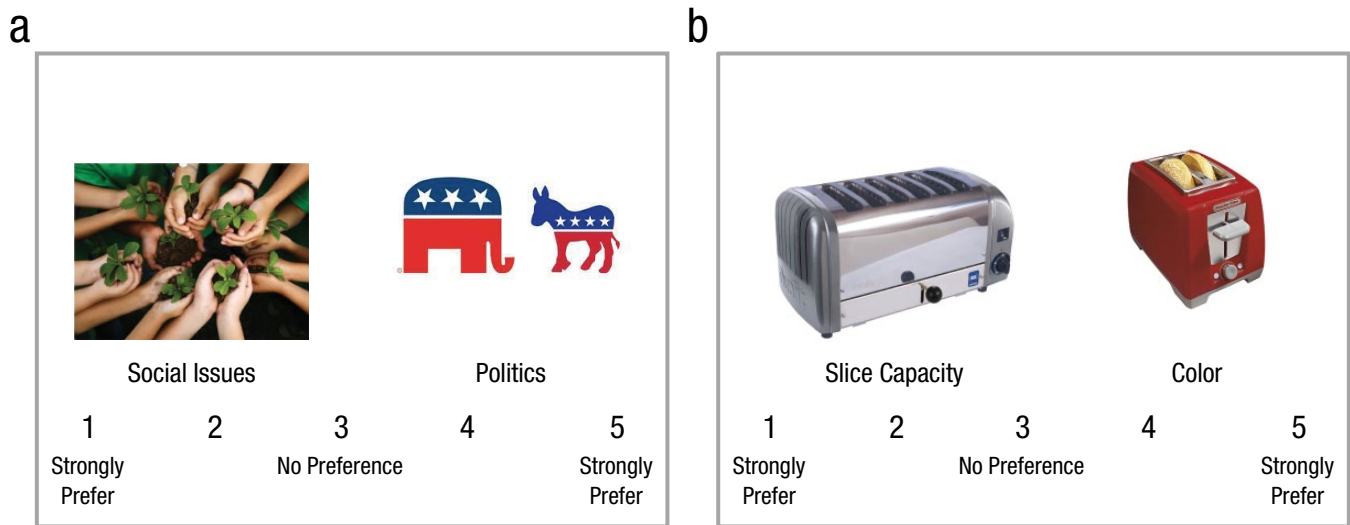
a

b

Social Issues   Politics

1	2	3	4	5
Strongly Prefer		No Preference		Strongly Prefer

Slice Capacity   Color

1	2	3	4	5
Strongly Prefer		No Preference		Strongly Prefer

Fig. 1. Examples of trials in the (a) self-affirmation and (b) nonaffirmation tasks in Study 1. In each trial, participants in the self-affirmation group (a) saw pictures associated with two personal values, with the values' labels below the pictures. Participants indicated which value they preferred, on a scale from 1 to 5 (1 = *strongly prefer* [the value on the left], 2 = *slightly prefer* [the value on the left], 3 = *no preference*, 4 = *slightly prefer* [the value on the right], 5 = *strongly prefer* [the value on the right]). In the experimental condition, one of the values was always the given participant's top-ranked value; in the yoked control condition, the values were always the given participant's two bottom-ranked values. Participants in the nonaffirmation group (b) saw pictures illustrating two toaster attributes, with the attributes' labels below the pictures. Participants indicated which attribute they thought the average college student would prefer, on the same scale from 1 to 5. In the experimental condition, one of the attributes was always the given participant's top-ranked attribute; in the yoked control condition, the attributes were always the given participant's two bottom-ranked attributes. The same pictures could represent either the experimental condition or the control condition, depending on a given participant's rankings.

reporting the personal values and toaster attributes selected by participants, see the Supplemental Material available online). We used this list to create a scanner task that was specific to each participant's beliefs about what other people prefer. Before the scanning session, participants were trained on the tasks assigned to them. In these training sessions, participants were familiarized with the pictures that represented each personal value or toaster attribute, so the pictures were not novel to participants at test.

Although many self-affirmation studies use a nonaffirmation control condition in which participants focus on a value of lower personal relevance and why it might be important to someone else (Sherman & Cohen, 2006), the design for this study was slightly different to accommodate standard imaging techniques. Specifically, fMRI contrasts require a within-subjects design, using comparisons between two conditions to isolate neural activity to the specific psychological differences between the conditions. Therefore, both groups in the current study included an experimental condition (self-affirmation group: high personal-value ratings; nonaffirmation group: high toaster attribute ratings) and a control condition (self-affirmation group: low personal-value ratings; nonaffirmation group: low toaster attribute ratings).

Consequently, our within-subjects self-affirmation conditions differed only in how the personal values displayed had previously been ranked by the participant, which controlled for any self-related processes. The within-subjects nonaffirmation conditions also differed only in how the toaster attributes displayed had previously been ranked by the participant. Including both within- and between-subjects levels of comparison offered a more specific test of whether self-affirmation relies on reward activity inherent in the self-affirmation process itself or just relies on self- or value-related reflection. This design also closely approximated the original paradigms used in behavioral self-affirmation studies, ensuring that we manipulated self-affirmation.

Imaging procedures for the self-affirmation and non-affirmation tasks. During each task, participants viewed instructions, pictures, and words via a high-resolution projector and were asked to make responses (when appropriate) using a five-button data glove.

The self-affirmation task used a block design. Each block included three decision-making trials, and each trial lasted for 8 s, for a total of 24 s per block. In the self-affirmation trials, participants were shown pictures associated with two personal values, one of which was always the top-ranked value; the values' labels appeared beneath the pictures (Fig. 1a). Participants were asked to think about the role of these two personal values in their lives and then indicate which of the two displayed values was more important to them. Participants responded using a 5-point scale (1 = *strongly prefer* [the value on the left], 2 = *slightly prefer* [the value on the left], 3 = *no*

preference, 4 = *slightly prefer* [the value on the right], 5 = *strongly prefer* [the value on the right]).

The control blocks had the same timing and instructions; however, participants were shown their fourth- and fifth-ranked personal values on each trial. Thus, the only difference between the experimental and control trials was whether subjects were able to affirm important values during the trials. Participants completed four experimental blocks in one run and four control blocks in another run. Run order was counterbalanced across participants. The two conditions were completed in separate runs to reduce carryover effects, given that self-affirmation has been shown to have sustained benefits over time (Cohen et al., 2006).

We wanted to have a comparison group to test whether the observed neural correlates of self-affirmation were due to the difference between making preference judgments regarding more important characteristics and making preference judgments regarding less important characteristics. To this end, we created a task that was similar to the self-affirmation task in design and demands but did not lead participants to self-affirm. As in the self-affirmation task, each nonaffirmation block included three 8-s decision-making trials, for a total of 24 s per block. In these trials, participants were shown pictures associated with toaster attributes with the attribute label beneath each picture (see Fig. 1b). Each experimental block contained three experimental trials. In the experimental trials, participants were asked to indicate which of the two displayed toaster attributes, one of which was always the top-ranked attribute, was more important to the average college student. Participants responded using a 5-point scale (1 = *strongly prefer* [the attribute on the left], 2 = *slightly prefer* [the attribute on the left], 3 = *no preference*, 4 = *slightly prefer* [the attribute on the right], 5 = *strongly prefer* [the attribute on the right]). Each control block included three trials. In the control trials, participants were given the same instructions, but were shown the fourth- and fifth-ranked attributes. Each of the trials and blocks was the same length as in the experimental condition. Participants completed four experimental blocks in one run and four control blocks in another run to match the self-affirmation group. Runs were counterbalanced.

Image acquisition. Data were acquired using a 3-T MRI scanner (Verio; Siemens Medical Solutions USA, Malvern, PA). Foam padding surrounded each participant's head to reduce head movement. For each participant, we first acquired a high-resolution structural magnetization-prepared rapid gradient-echo (MPRAGE) imaging volume—repetition time (TR) = 1,700 ms, echo time (TE) = 2.48, matrix size = 128 × 128, resolution = 1 × 1 × 1 mm, field of view (FOV) = 256 mm, 176 slices (1

mm thick), flip angle = 9°, and bandwidth = 170 Hz/pixel. For the tasks, two functional gradient-echo scans (3 min for self-affirmation/toasters and 3 min for control) were acquired—TR = 2,000 ms, TE = 29 ms, flip angle = 79°, matrix size = 64 × 64, resolution = 3 × 3 × 3 mm, FOV = 192 mm, 36 axial slices (3 mm thick), and bandwidth = 2232 Hz/pixel. These tasks were conducted in separate runs to give participants a break and to minimize carryover effects.

fMRI data analysis. Imaging data were analyzed using Statistical Parametric Mapping software (SPM8; Wellcome Department of Cognitive Neurology, Institute of Neurology, London, England). For preprocessing, we first manually reoriented the echoplanar images to align brains along a horizontal anterior commissure-posterior commissure, with an image origin at the anterior commissure. For functional images, the first run's first-image parameters were applied to each subsequent volume in the respective run to correct for head motion. Structural MPRAGE images were normalized to Montreal Neurological Institute (MNI) space using diffeomorphic anatomical registration through exponentiated lie (DARTEL) algorithms and were then smoothed using an 8-mm Gaussian kernel, full-width at half maximum. Before first-level analyses, images were visually inspected for accurate normalization. The 24 s of trials for each condition (self-affirmation experimental condition, self-affirmation control condition, nonaffirmation experimental condition, nonaffirmation control condition) were modeled as blocks. Rest periods, when participants viewed a fixation cross between blocks, comprised the implicit baseline.

We computed linear contrasts for the self-affirmation experimental condition compared with its control condition (high-rated personal values vs. low-rated personal values) for each participant. These individual contrast images were then used in group-level analyses. For the nonaffirmation group, we computed linear contrasts for the experimental condition compared with its control condition (high-rated toaster attributes vs. low-rated toaster attributes) for each participant. These individual contrast images were then used in group-level analyses. In addition, to determine whether there were differences in neural activation between the two groups, an independent (two-sample) *t* test was computed comparing the self-affirmation group (self-affirmation experimental minus yoked control contrast) to the nonaffirmation group (experimental minus yoked control contrast).

On the basis of a priori predictions that self-affirmation would activate reward-related regions, group-level results were examined using regions of interest (ROIs) of the left and right VS and the VTA. VS ROIs were structurally defined using the automated anatomical labeling atlas (Tzourio-Mazoyer et al., 2002); caudate nucleus and

putamen from the atlas were combined and constrained at x between 0 and −24, y between 4 and 18, and z between 0 and −12 for the left ROI and x between 0 and 24, y between 4 and 18, and z between 0 and −12 for the right ROI (based on ROIs from Inagaki & Eisenberger, 2012). Thus, we constrained the ROI to the ventral parts of the caudate nucleus and putamen to create this VS ROI. The VTA ROI was created in the MarsBar Toolbox for SPM (http://marsbar.sourceforge.net; Brett, Anton, Valabregue, & Poline, 2002) and centered at ($x = 2$, $y = −12$, $z = −8$) within a 3-mm radius on the basis of previous work investigating social reward (Aron et al., 2005). We examined activity within each of these ROIs for the self-affirmation and nonaffirmation conditions relative to their yoked control conditions. Parameter estimates representing the average activity across all voxels in the ROI were extracted and averaged. ROI analyses were run in MarsBar, which reports an fMRI standard statistical threshold of $p < .05$, one-tailed. Confidence intervals (CIs) for these tests were estimated using the bias-corrected and accelerated-percentile method (10,000 random samples with replacement), implemented using the BOOTCI function in MATLAB (The MathWorks, Natick, MA).

Because this was the first study to explore neural activity during self-affirmation, we also conducted exploratory whole-brain analyses to provide a complete picture of the neural regions involved in this process. Thus, we performed whole-brain analyses contrasting activity in the self-affirmation experimental trials (relative to activity in the yoked control trials) with activity in the nonaffirmation experimental trials (relative to activity in the yoked control trials). We used an exploratory threshold of $p < .005$ and 20 voxels (Lieberman & Cunningham, 2009). We then explored the post hoc simple effects. To do so, we created ROIs based on the clusters of activity in the whole-brain analysis and extracted and averaged parameter estimates for the contrast between experimental and yoked control trials in the self-affirmation group and for the contrast between experimental and yoked control trials in the nonaffirmation group. Analyses were run in MarsBar, which reports an fMRI standard statistical threshold of $p < .05$, one-tailed.

Results

ROI analyses. To examine differences in neural activity between the self-affirmation and nonaffirmation groups, we investigated whether activity in the self-affirmation experimental condition (measured as activity relative to its yoked control condition) was greater than activity in the nonaffirmation experimental condition (measured as activity relative to its yoked control condition). We then further examined these effects by examining neural activity in each group separately.

Results showed greater activity in the left VS in the self-affirmation group compared with the nonaffirmation group, $t(36) = 2.04$, $p = .025$ (Fig. 2). Consistent with our hypotheses, within the self-affirmation group, there was greater activity in the left VS during the experimental trials than in the control trials, $t(19) = 2.12$, $p = .024$, mean parameter estimate = 0.538, 95% CI = [0.022, 0.987], whereas within the nonaffirmation group, there was no difference in left VS activity during the experimental and control trials, $t(17) = −0.53$, $p > .250$, mean parameter estimate = −0.099, 95% CI = [−0.424, 0.280].

Results showed greater activity in the right VS in the self-affirmation group than in the nonaffirmation group, but the difference was not statistically significant, $t(36) = 1.02$, $p = .157$ (Fig. 2). As we found for the left VS, this effect was driven by the self-affirmation group; there was greater activity, albeit not significantly greater, in the right VS during the experimental trials than during the control trials, $t(19) = 1.06$, $p = .152$, mean parameter estimate = 0.306, 95% CI = [−0.257, 0.837], but there was no difference between the experimental and control trials in the nonaffirmation group, $t(17) = −0.32$, $p > .250$, mean parameter estimate = −0.072, 95% CI = [−0.483, 0.350].

Finally, there were no differences in VTA activity between the self-affirmation and nonaffirmation groups, $t(36) = 0.13$, $p > .250$ (Fig. 2). Specifically, in the self-affirmation group, there was marginally greater (albeit nonsignificant) VTA activity in the experimental trials than in the control trials, $t(19) = 1.34$, $p = .098$, mean parameter estimate = 0.731, 95% CI = [−0.054, 2.119], and the same pattern was seen in the nonaffirmation group, $t(17) = 1.22$, $p = .119$, mean parameter estimate = 0.632, 95% CI = [−.203, 1.806].

Whole-brain analyses. Like the ROI analyses, whole-brain analyses revealed significantly greater activation in the VS during the self-affirmation experimental trials (relative to activation in the yoked control) than in the nonaffirmation experimental trials (relative to activation in the yoked control; Table 1). There was also significant activation in clusters in the medial prefrontal cortex and precuneus/posterior cingulate cortex, regions previously shown to play a role in self-processing (Heatherton et al., 2006; for a full list of activations, see Table 1). To further investigate these findings, we used post hoc tests of simple effects to explore which trials drove the effect. As expected for the VS, these tests revealed that there was greater activity in the VS during self-affirmation experimental trials than during the yoked control trials, $t(19) = 2.26$, $p = .018$, but there was no difference in activity between the nonaffirmation experimental and yoked control trials, $t(17) = −2.90$, $p > .250$. The other clusters showed the same pattern: There was greater activity during the self-affirmation experimental trials than during
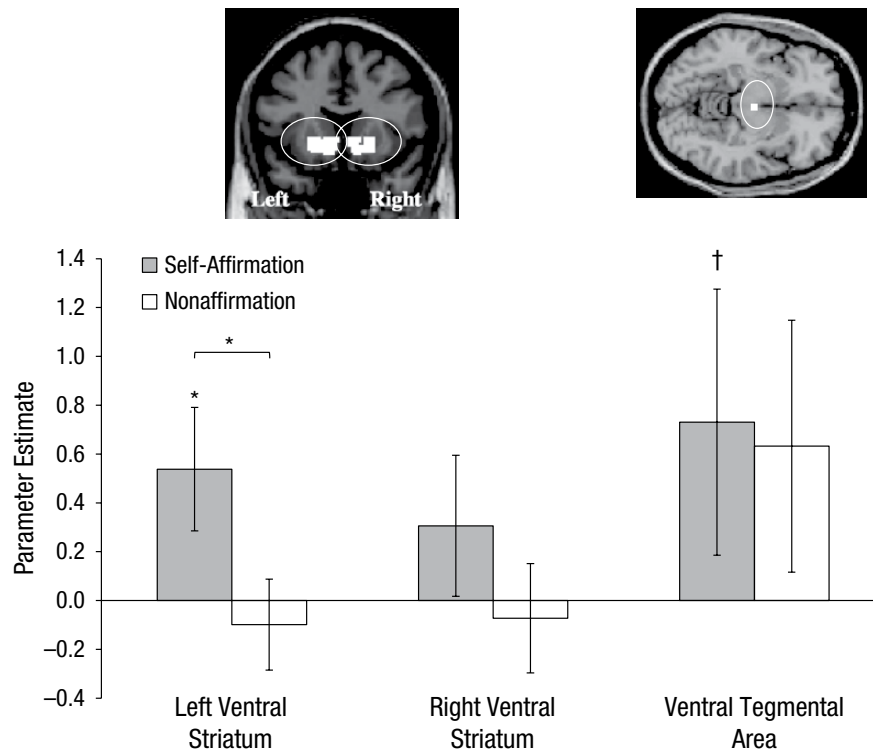
Fig. 2. Results for Study 1: parameter estimates for activity in the left ventral striatum, the right ventral striatum, and the ventral tegmental area during experimental trials are graphed separately for each condition. Activation was measured relative to the corresponding control trials. Error bars represent ±1 *SE*. Symbols indicate parameter estimates that are significantly or marginally significantly different from 0 as well as a significant difference between conditions (†*p* < .10, **p* < .05). The coronal scan (top left) shows the location of the regions of interest in the left ventral striatum and the right ventral striatum, and the transverse scan (top right) shows the location of the region of interest in the ventral tegmental area.

the yoked control trials (*p*s < .024; medial prefrontal cortex: *p* = .143), but there was no difference in activity between the nonaffirmation experimental and yoked control trials (*p*s > .559).

On the other hand, whole brain analyses revealed no significant clusters of activity during the reverse contrast, which tested for greater activity in the experimental condition (relative to the yoked control condition) in the non-affirmation group compared with the self-affirmation group. In addition, within the self-affirmation group, there were no clusters with significantly greater activity in the control condition than in the experimental condition.

Table 1. Results From Study 1: Brain Regions More Active During Self-Affirmation Trials Than During Nonaffirmation Trials

Anatomical region	Brodmann's area	Hemisphere	MNI coordinates of peak voxel			*t*(36)	Number of voxels (*k*)
			x	*y*	*z*		
Ventral striatum	—	Left	−3	15	−6	3.48	29
Medial prefrontal cortex	10		0	54	−6	4.15	330
Occipital cortex	18	Right	6	−78	−3	4.02	321
Subgenual anterior cingulate cortex	24	Left	−18	24	−6	3.40	22
Thalamus	—	Right	3	−9	3	4.04	46
Middle temporal cortex	39	Left	−36	−63	12	3.45	36
Rostrolateral prefrontal cortex	9	Left	−24	30	36	4.41	40

Note: Activations were measured relative to activation in the yoked control trials. The table reports significant activations (*p* < .005) of clusters with a minimum size of 20 voxels. The *t* tests were conducted at peak coordinates. MNI = Montreal Neurological Institute.

Interim Summary

Results indicated that the self-affirmation group showed greater VS activity than did the nonaffirmation group. This effect was driven by increased VS activity specific to the self-affirmation condition. Whole-brain results revealed that neural regions implicated in self-processing (precuneus/posterior cingulate cortex) and reward processing (VS) were more active in the self-affirmation group than in the nonaffirmation group. Results from the VTA were inconclusive. Most self-affirmation paradigms do not include pictures, so we conducted a second study without pictures to rule out the possibility that the VS activity was due to viewing pictures. Moreover, to test the generalizability of our findings, we conducted this study with a community sample.

Study 2

Method

Participants. Twenty-one female participants (mean age = 55.5 years) completed the study procedures. All participants were deemed eligible for fMRI (i.e., right-handed, not claustrophobic, free of implanted metal, and not pregnant). Participants were recruited from a larger study investigating the neurobiological pathways linking psychological stress and inflammation in breast cancer survivors and healthy control participants; thus, 6 of the participants were breast-cancer survivors. Because there were no differences in neural activity between the healthy participants and breast cancer survivors in any of the analyses for any of the ROIs (*p*s > .20), we collapsed the data across participants for all analyses reported here. We aimed to collect data from a minimum of 20 participants, and data collection stopped at the targeted enrollment for the larger study. Data from 1 participant were excluded because she did not follow task instructions. Seventy-six percent of participants were White, 14% were of "other" race, 5% were Hispanic, and 5% were Asian American. The University of California, Los Angeles, institutional review board approved all procedures.

Procedure

Imaging procedures for the self-affirmation task. Before the scanning session, participants were emailed a survey in which they were asked to rank a list of 11 personal values (e.g., art, religion, friends and family) in order of importance (for a frequency table reporting the personal values selected by participants, see the Supplemental Materials). From this, we were able to create tasks that were specific to each participant's most important personal value for the scanning session. During the scan, participants viewed instructions and words through scanner-compatible goggles and were asked to make responses (when appropriate) using a four-button button box.

The self-affirmation task was similar to that used in Study 1; however, for this study, all participants completed the self-affirmation and control trials only (i.e., there was no nonaffirmation group). For this experiment, participants selected which of the two personal values shown on screen was most important to them on each trial. However, in this study, participants were shown only the personal-value label, without a picture; we hoped to ensure that the results observed in Study 1 were not being driven by participants' seeing pictures of important values. Each self-affirmation block included three trials, lasting 7 s each, separated by a 1-s fixation cross, for a total of 23 s per block. During the self-affirmation trials, participants were shown their top-ranked personal value and another highly ranked value. They were asked to indicate which of the two personal values displayed was more important using a 4-point scale (1 = *strongly prefer* [the value on the left], 2 = *slightly prefer* [the value on the left], 3 = *slightly prefer* [the value on the right], 4 = *strongly prefer* [the value on the right]). Each control block included three trials. During the control trials, participants were given the same instructions, but were shown two personal values that they had rated as being less important to them, with the same timing as the self-affirmation trials and blocks. Participants completed three self-affirmation blocks and three control blocks. Blocks were randomized once with the constraint that there were never three consecutive blocks of the same condition for any task.[1] The blocks could be ordered in seven different ways, but each participant saw only one order.

Image acquisition. Data were acquired on a Siemens Trio 3-T MRI scanner. Foam padding surrounded each participant's head to reduce head movement. For each participant, we acquired a high-resolution structural matched-bandwidth scan—TR = 5,000 ms, TE = 34, matrix size = 128 × 128, resolution = 1.6 × 1.6 × 3 mm, FOV = 200 mm, 36 slices (3 mm thick), flip angle = 90°, and bandwidth = 1302 Hz/pixel. The self-affirmation task was completed in one functional scan lasting 436 s (about 7 min, 16 s)—TR = 2,000 ms, TE = 25 ms, matrix size = 64 × 64, resolution = 3.1 × 3.1 × 4.0 mm, FOV = 200 mm, 33 axial slices (3 mm thick with 1-mm gap), flip angle = 90°, and bandwidth = 2604 Hz/pixel.

fMRI data analysis. Imaging data were analyzed using SPM8. For preprocessing, functional and anatomical images were realigned, coregistered to the structural scan, and normalized using the DARTEL procedure in SPM8. For each participant, the 23 s of self-affirmation decision-making trials were modeled as the self-affirmation blocks, and the 23 s of control trials were modeled as the

control blocks. Implicit baseline consisted of the rest periods (viewing a fixation cross).

We computed linear contrasts comparing the self-affirmation trials with the control trials for each participant. These individual contrast images were then used in group-level analyses. We examined activity within each of the ROIs used in Study 1 for the self-affirmation trials compared with the control trials. Parameter estimates representing the average activity across all voxels in the ROI were extracted and averaged. Analyses were run in Mars-Bar. A standard statistical threshold of $p < .05$ was used for these ROI analyses. CIs for these tests were estimated using the bias-corrected and accelerated-percentile method (10,000 random samples with replacement; implemented using the BOOTCI function in MATLAB).

To supplement the ROI analyses, we performed whole brain analyses on the self-affirmation compared with control contrast. Whole-brain analyses used an exploratory threshold ($p < .005$, $k = 20$; Lieberman & Cunningham, 2009). All coordinates are reported in MNI space.

Results

ROI analyses. As in Study 1, compared with the control trials, the self-affirmation trials produced significantly more left VS activity, $t(19) = 1.84$, $p = .041$, mean parameter estimate = 0.261, 95% CI = [0.001, 0.548], and marginally more right VS activity, $t(19) = 1.62$, $p = .061$, mean parameter estimate = 0.263, 95% CI = [−0.046, 0.560] (Fig. 3). As in Study 1, there was no difference in VTA activity between the self-affirmation and control trials, $t(19) = 0.72$, $p = .240$, mean parameter estimate = 0.190, 95% CI = [−0.111, 0.597] (Fig. 3).

Whole-brain analyses. As in Study 1, results from the whole-brain analyses revealed a significant cluster in the VS (putamen) as well as in the medial prefrontal cortex (for a full list of activations, see Table 2). Results from the posterior cingulate cortex at this threshold did not reach statistical significance. The control condition > self-affirmation condition contrast revealed no significant clusters of activation.

Discussion

Reflecting on important personal values during self-affirmation activated neural reward pathways (VS) across two studies with different age groups, using either personal-value pictures with text or text only. VS activation was not due simply to making judgments about preferences or personal values; rather, it was specific to thinking about one's most important personal value. This is the first fMRI study to identify neural systems engaged during self-affirmation, which extends self-affirmation theory by suggesting that self-affirmation is rewarding. This is a first step toward identifying the neural mechanisms by which self-affirmation reduces threat and stress responding, improves performance, reduces defensiveness, and alters social and health behaviors.

The VS is a key region in the mesolimbic dopamine reward pathway, which suggests that affirming important personal values is rewarding and may lead to a cascade of effects associated with reward processing. Indeed, when we investigated the term "reward" using Neurosynth (http://neurosynth.org/), a large-scale database of neuroimaging studies that provides meta-analytic reverse-inference analyses, the z score identified in the VS by
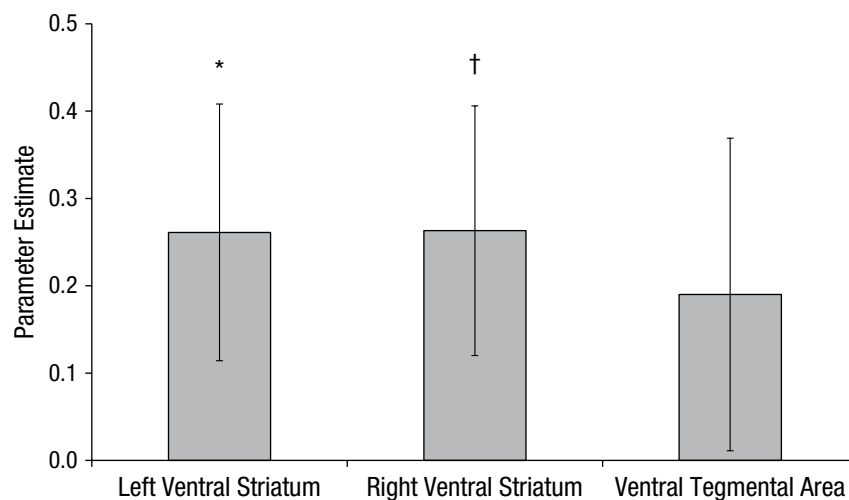

Fig. 3. Results for Study 2: parameter estimates for activity in the left ventral striatum, the right ventral striatum, and the ventral tegmental area during experimental trials. Error bars represent ±1 *SE*. Symbols indicate the results of tests comparing the experimental trials with the control trials (†$p < .10$, *$p < .05$).

Table 2. Results From Study 2: Brain Regions More Active During Self-Affirmation Than During Control Trials

Anatomical region	Brodmann's area	Hemisphere	MNI coordinates			$t(20)$	Number of voxels (k)
			x	y	z		
Ventrolateral prefrontal cortex	47	Left	−51	27	−3	4.14	135
Ventral striatum or putamen[a]		Left	−21	21	−6	3.63	
Ventrolateral prefrontal cortex	47	Right	33	27	−9	3.65	24
Medial prefrontal cortex	10	Right	9	60	18	3.74	27
Rostrolateral prefrontal cortex	10	Right	21	63	9	4.18	30
Dorsolateral prefrontal cortex	46	Right	45	30	21	3.82	62
Dorsomedial prefrontal cortex	9	Left	−12	51	21	3.36	24
Supplementary motor area	6	—	−6	24	54	4.13	276
Angular gyrus	40	Left	−54	−66	39	3.74	24

Note: Activations were measured relative to activation in the yoked control trials. The table reports significant activations ($p < .005$) of clusters with a minimum size of 20 voxels. The t tests were conducted at peak coordinates. MNI = Montreal Neurological Institute.
[a]This activation extended from the larger cluster listed in the previous row.

Neurosynth suggested that neural activity reliably indicated reward processing. In fact, the z score for inferring that activity in the VS is indicative of reward ($z = 26.11$) was higher than the z score for inferring that activity in the occipital cortex was indicative of vision ($z = 13.36$) or that activity in the amygdala was indicative of affect ($z = 6.41$), fear ($z = 13.10$), or emotion ($z = 18.01$). These findings from Neurosynth provide further support for our evidence that self-affirmation elicits reward-related processing.

These two studies used different populations, which suggests that the neural correlates of self-affirmation extend from undergraduates to community-dwelling adults. This is consistent with the wealth of previous research on self-affirmation, which has found benefits for a range of ages and ethnicities (Cohen & Sherman, 2014). Study 2's self-affirmation task differed from the one used in Study 1 in that it used a different list of personal values and no pictures, but it yielded similar effects. This provides converging evidence that increased VS activity during self-affirmation is not specific to viewing personal-value pictures. Rather, the VS is sensitive to thinking about one's most important value. Although the difference in VS activity between the self-affirmation and yoked control conditions was statistically significant for the left VS but not quite significant for the right VS, the pattern was in the same direction for both regions.

Whole-brain results showed that, in addition to activating the VS, self-affirmation led to greater activity in regions typically associated with thinking about the self (the precuneus in Study 1; the medial prefrontal cortex in Studies 1 and 2). Indeed, self-affirmation requires participants to reflect on their preferences, which makes these findings unsurprising. However, self-affirmation also led

to greater activity in self-processing regions compared with its yoked control, which still required participants to think about their own preferences. The difference was that the self-affirmation experimental condition had participants think about their top-ranked personal value. Self-affirmation may lead to a deeper processing of self-related information, which may also be a key ingredient in self-affirmation's effects on behavior.

The present studies employed stringent control conditions. For the self-affirmation task, participants received the same instructions for both conditions; the only difference was the participant's prior ratings of the personal values shown. This allowed us to conclude that it was not preference judgments about important values per se that activated reward circuitry; rather, it was about focusing on one's most important personal values. Whereas many standard self-affirmation manipulations use a control condition in which participants think about why a less important value might be important to someone else, our design provided a more specific test of the neural activity involved in the self-affirmation condition, providing insight into a possible mechanism for self-affirmation.

One important question for future studies is whether this neural-reward account of self-affirmation can explain the subsequent cascade of neural and psychological effects observed in previous studies (Creswell et al., 2005; Falk et al., 2015; Legault et al., 2012; Sherman et al., 2000). Falk et al. (2015) showed that activity in the ventromedial prefrontal cortex in response to health messages was greater in participants who performed a self-affirmation task than in participants who performed a control task, and their findings may be consistent with our findings. Specifically, the ventral striatum shows functional connectivity with the ventromedial prefrontal

cortex (Di Martino et al., 2008). It is possible, then, that VS activity during self-affirmation activates a cascade of neural effects, including increased ventromedial prefrontal cortex activity, that in turn lead to psychological and behavioral changes. Our findings could also be consistent with the finding that self-affirmation leads to greater neurophysiological error-related negativity during subsequent tasks (Legault et al., 2012). Specifically, there is evidence that reward and positive affect increase correlates of error-related negativity event-related potentials, which might relate to improved conflict adaptation (Larson, Perlstein, Stigge-Kaufman, Kelly, & Dotson, 2006; Stürmer, Nigbur, Schacht, & Sommer, 2011). Therefore, it is plausible that self-affirmation activates reward processing, which increases error-related negativity and causes a shift in strategy that leads to improved performance (Legault et al., 2012). Further research should investigate the subsequent neural and behavioral processes that follow self-affirmation's reward activity.

There are a few limitations to these studies. In Study 1, we chose to have participants think about the toaster attributes most important to an average college student to ensure this task was not inadvertently affirming. Future work should have participants make decisions that are important to them, but not in the values domain. In Study 1, the experimental and control conditions were in separate runs, which is nonstandard for fMRI studies. Because self-affirmation findings sometimes show lasting effects (Cohen et al., 2006), this design was selected to provide participants with a break to prevent carryover. Study 2 employed a more typical design with blocks of affirmation and control randomly ordered in one run. However, the results were the same for both studies. In these studies, self-affirmation was manipulated via values affirmation. Although this is the most common self-affirmation operationalization used, it is not the only one. Our results may be specific to this values-affirmation procedure, and future studies could determine whether all self-affirmation operationalizations rely on similar neural substrates.

These findings provide insight into the neural mechanism by which self-affirmation reduces threat responding. Recent work has found that rewarding stimuli (e.g., sweet foods, sweet drinks, or sexual stimuli) lead to reductions in stress responding (Creswell, Pacilio, et al., 2013; Ulrich-Lai et al., 2010). This effect extends to social rewards also, given that social support activates reward regions, such as the VS (Inagaki & Eisenberger, 2013; Strathearn, Fonagy, Amico, & Montague, 2009), and has been shown to reduce threat-related neural activity (Eisenberger et al., 2011; Younger, Aron, Parke, Chatterjee, & Mackey, 2010). It is possible that self-affirmation relies on similar neural mechanisms to reduce threat responding. In the present article, self-affirmation (vs. control) led to greater VS activity, which could correspond with activation decreases in neural threat regions during subsequent tasks. In turn, this could contribute to the array of threat reduction benefits that self-affirmation has been shown to foster. Future studies should assess whether this proposed reward-system mechanism underlies the stress-buffering effects shown in previous self-affirmation studies. Although we suggest a reward-related mechanism, it is possible that the reward associated with self-affirmation may be distinct from nonself-related reward (e.g., food, winning money). Future studies could investigate the possible distinct and overlapping neural reward circuitry underlying different reward processes.

Author Contributions

All authors contributed to the concept and to the design of the studies. J. M. Dutcher, L. E. Pacilio, and K. A. Muscatell contributed to testing and data collection. J. M. Dutcher performed the data analysis. J. M. Dutcher, J. D. Creswell, and N. I. Eisenberger interpreted the data. J. M. Dutcher, J. D. Creswell, P. R. Harris, W. M. P. Klein, J. M. Levine, and N. I. Eisenberger drafted the manuscript. All authors approved the final version of the manuscript for submission.

Supplemental Material

Additional supporting information can be found at http://pss .sagepub.com/content/by/supplemental-data

Note

1. Participants also completed another self-affirmation task in which they thought about important values for 23 s (this task is similar to self-affirmation manipulations involving writing; Cohen, Aronson, & Steele, 2000). However, because the neuroimaging results suggested this manipulation was not successful, we do not focus on this task here. The blocks in this task

were modeled separately from the blocks in the task of interest. Details of this task procedure and results are included in the Supplemental Material.

References

Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., & Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology*, *94*, 327–337.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *NeuroImage*, *16*, S497.

Burson, A., Crocker, J., & Mischkowski, D. (2012). Two types of value-affirmation: Implications for self-control following social exclusion. *Social Psychological & Personality Science*, *3*, 510–516.

Cohen, G. L., Aronson, J., & Steele, C. M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self. *Personality and Social Psychology Bulletin*, *26*, 1151–1164.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310.

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*, 333–371.

Correll, J., Spencer, S. J., & Zanna, M. P. (2004). An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength. *Journal of Experimental Social Psychology*, *40*, 350–356.

Creswell, J. D., Dutcher, J. M., Klein, W. M. P., Harris, P. R., & Levine, J. L. (2013). Self-affirmation improves problem-solving under stress. *PLoS ONE*, *8*(5), Article e62593. doi:10.1371/journal.pone.0062593

Creswell, J. D., Pacilio, L. E., Denson, T. F., & Satyshur, M. (2013). The effect of primary sexual reward manipulation on cortisol responses to psychosocial stress in men. *Psychosomatic Medicine*, *75*, 397–403.

Creswell, J. D., Welch, W., Taylor, S. E., Sherman, D. K., Gruenewald, T., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Psychological Science*, *16*, 846–851.

Crocker, J., Niiya, Y., & Mischkowski, D. (2008). Why does writing about important values reduce defensiveness? Self-affirmation and the role of positive other-directed feelings. *Psychological Science*, *19*, 740–747.

Di Martino, A., Scheres, A., Margulies, D. S., Kelly, A. M. C., Uddin, L. Q., Shehzad, Z., . . . Milham, M. P. (2008). Functional connectivity of human striatum: A resting state fMRI study. *Cerebral Cortex*, *18*, 2735–2747.

Eisenberger, N. I., Master, S. L., Inagaki, T. K., Taylor, S. E., Shirinyan, D., Lieberman, M. D., & Naliboff, B. D. (2011). Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proceedings of the National Academy of Sciences, USA*, *108*, 11721–11726.

Falk, E. B., O'Donnell, M. B., Cascio, C. N., Tinney, F., Kang, Y., Lieberman, M. D., . . . Strecher, V. J. (2015). Self-affirmation alters the brain's response to health messages and subsequent behavior change. *Proceedings of the National Academy of Sciences, USA*, *112*, 1977–1982. doi:10.1073/pnas.1500247112

Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., & Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience*, *1*, 18–25.

Inagaki, T. K., & Eisenberger, N. I. (2012). Neural correlates of giving support to a loved one. *Psychosomatic Medicine*, *74*, 3–7.

Klein, W. M. P., & Harris, P. R. (2009). Self-affirmation enhances attentional bias toward threatening components of a persuasive message. *Psychological Science*, *20*, 1463–1467.

Koole, S. L., Smeets, K., Van Knippenberg, A., & Dijksterhuis, A. (1999). The cessation of rumination through self-affirmation. *Journal of Personality and Social Psychology*, *77*, 111–125.

Larson, M. J., Perlstein, W. M., Stigge-Kaufman, D., Kelly, K. G., & Dotson, V. M. (2006). Affective context-induced modulation of the error-related negativity. *NeuroReport*, *17*, 329–333.

Legault, L., Al-Khindi, T., & Inzlicht, M. (2012). Preserving integrity in the face of performance threat: Self-affirmation enhances neurophysiological responsiveness to errors. *Psychological Science*, *23*, 1353–1360.

Lieberman, M., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*, *4*, 423–428.

Nelson, S. K., Fuller, J. A., Choi, I., & Lyubomirsky, S. (2014). Beyond self-protection: Self-affirmation benefits hedonic and eudaimonic well-being. *Personality and Social Psychology Bulletin*, *40*, 998–1011.

Sherman, D. A. K., Nelson, L. D., & Steele, C. M. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin*, *26*, 1046–1058.

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 38, pp. 183–242). San Diego, CA: Academic Press.

Sherman, D. K., & Hartson, K. A. (2011). Reconciling self-protection with self-improvement: Self-affirmation theory. In M. Alicke & C. Sedikides (Eds.), *The handbook of self-enhancement and self-protection* (pp. 128–151). New York, NY: Guilford Press.

Speer, M. E., Bhanji, J. P., & Delgado, M. R. (2014). Savoring the past: Positive memories evoke value representations in the striatum. *Neuron*, *84*, 847–856.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). San Diego, CA: Academic Press.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*, 5–19.

Stinson, D. A., Logel, C., Shepherd, S., & Zanna, M. P. (2011). Rewriting the self-fulfilling prophecy of social rejection: Self-affirmation improves relational security and social behavior up to 2 months later. *Psychological Science*, *22*, 1145–1149.

Strathearn, L., Fonagy, P., Amico, J., & Montague, P. R. (2009). Adult attachment predicts maternal brain and oxytocin response to infant cues. *Neuropsychopharmacology, 34,* 2655–2666.

Stürmer, B., Nigbur, R., Schacht, A., & Sommer, W. (2011). Reward and punishment effects on error processing and conflict control. *Frontiers in Psychology, 2,* Article 335. doi:10.3389/fpsyg.2011.00335

Tamir, D. I., & Mitchell, J. P. (2012). Disclosing information about the self is intrinsically rewarding. *Proceedings of the National Academy of Sciences, USA, 109,* 8038–8043.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage, 15,* 273–289.

Ulrich-Lai, Y. M., Christiansen, A. M., Ostrander, M. M., Jones, A. A., Jones, K. R., Choi, D. C., . . . Herman, J. P. (2010). Pleasurable behaviors reduce stress via brain reward pathways. *Proceedings of the National Academy of Sciences, USA, 107,* 20529–20534.

Vernon, P. E., & Allport, G. W. (1931). A test for personal values. *The Journal of Abnormal and Social Psychology, 26,* 231–248.

Younger, J., Aron, A., Parke, S., Chatterjee, N., & Mackey, S. (2010). Viewing pictures of a romantic partner reduces experimental pain: Involvement of neural reward systems. *PLoS ONE, 5*(10), e13309. doi:10.1371/journal.pone.0013309