Self-affirmation alters the brain’s response to health messages and subsequent behavior change

Emily B. Falk, Matthew Brook O’Donnell, Christopher N. Cascio, Francis Tinney, Yoona Kang, Matthew D. Lieberman, Shelley E. Taylor, Lawrence An, Kenneth Resnicow, and Victor J. Strebcher

*Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA 19104; †Research Center for Group Dynamics, Institute for Social Research, University of Michigan, Ann Arbor, MI 48106; ‡Department of Psychology, University of California, Los Angeles, CA 90095; ††Center for Health Communications Research, University of Michigan, Ann Arbor, MI 48109; and ‡‡Department of Health Behavior and Health Education, University of Michigan School of Public Health, Ann Arbor, MI 48109

Contributed by Shelley E. Taylor, January 6, 2015 (sent for review December 8, 2014)

Health communications can be an effective way to increase positive health behaviors and decrease negative health behaviors; however, those at highest risk are often most defensive and least open to such messages. For example, increasing physical activity among sedentary individuals affects a wide range of important mental and physical health outcomes, but has proven a challenging task. Affirming core values (i.e., self-affirmation) before message exposure is a psychological technique that can increase the effectiveness of a wide range of interventions in health and other domains; however, the neural mechanisms of affirmation’s effects have not been studied. We used functional magnetic resonance imaging (fMRI) to examine neural processes associated with affirmation effects during exposure to potentially threatening health messages. We focused on an a priori defined region of interest (ROI) in ventromedial prefrontal cortex (VMPFC), a brain region selected for its association with self-related processing and positive valuation. Consistent with our hypotheses, those in the self-affirmation condition produced more activity in VMPFC during exposure to health messages and went on to increase their objectively measured activity levels more. These findings suggest that affirmation of core values may exert its effects by allowing at-risk individuals to see the self-relevance and value in otherwise-threatening messages.

**Significance**

Self-affirmation is a psychological technique that is effective in increasing receptivity to interventions across domains from promoting health behaviors in high-risk populations to improving academic performance in underrepresented groups. The neural mechanisms that lead to affirmation’s success, however, are not known. We show that neural responses associated with self-related processing and value in response to an otherwise-threatening health communication intervention can be changed using self-affirmation; furthermore, these neural responses predict objectively measured behavior change in the month following the intervention. These findings suggest that self-affirmation may exert its effects by allowing at-risk individuals to see the self-relevance and value in otherwise-threatening messages and provide a framework for studying neural effects of self-affirmation more broadly.

www.pnas.org/cgi/doi/10.1073/pnas.1500247112

PNAS Early Edition | 1 of 6
Potential Neural Mechanisms of Affirmation

The current investigation examines neural activity associated with the effects of self-affirmation on processing health risk messages related to sedentary behavior in sedentary adults. We focused on the brain’s ventromedial prefrontal cortex (VMPFC) during exposure to potentially threatening health messages emphasizing the need to be more active and less sedentary in a group of sedentary adults. VMPFC is the most common region implicated in self-related processing (51) and is also a key region, along with the ventral striatum, implicated in positive valuation of stimuli (52). In addition, VMPFC has been consistently associated with behavior change in response to health messages in prior work (44–46, 50). This prior research has suggested that the link between VMPFC activity during health message exposure and behavior change may stem from a recipient’s ability to process a health message as self-relevant or as having value to oneself. Thus, we hypothesized that if affirmation allows people to see otherwise-threatening information as more self-relevant and valuable, delivering self-affirmation before health messages should increase neural activity in VMPFC during message exposure. In addition, we hypothesized that internalizing the messages in this way would be associated with subsequent behavior change. More specifically, we hypothesized that increased neural responses within the VMPFC during message exposure would predict behavior change consistent with the messaging. In this case, we expected that increased activity in VMPFC in response to health messages should be associated with decreased sedentary behavior following the scan.

Results

Changes in Sedentary Behavior. We measured physical activity using wrist worn accelerometers (Methods and SI Methods). At baseline, participants were sedentary an average of 50.6% of their awake time (SD, 14.0%; range, 21–84%), which is close to the national average (53). On average, controlling for baseline sedentary behavior and demographics, participants showed significant declines in their sedentary behavior over time in the month following exposure to the health message intervention \( (γ_{time} = −0.001; t = −3.49; P = 0.0005) \).

Effects of Affirmation on Brain Activity and on Behavior Change. On average, participants who were affirmed showed greater activity during exposure to the health messages within our hypothesized VMPFC region of interest (ROI) compared with those who were unaffirmed, controlling for baseline sedentary behavior and demographics \( [B = 0.15, t_{(34)} = 2.10, P = 0.04; \text{Fig. } 1A] \). For whole-brain effects of self-affirmation during exposure to health messages, see Table 1. We next examined the effect of self-affirmation on changes in sedentary behavior over time, controlling for baseline levels of sedentary behavior and demographics. Those who were in the affirmation condition showed greater declines in sedentary behavior more over time following exposure to health messages (condition by time), compared with those in the control condition \( (γ_{time \times condition} = −0.002, t = −2.68, P = 0.008; \text{Fig. } 1B) \).

Neural Activity During Health Messages Predicts Changes in Sedentary Behavior Distinct from Self-Reports. We next examined whether neural activity in our hypothesized VMPFC ROI during message encoding predicted changes in sedentary behavior over time following the scanner intervention. Those who showed greater activity in VMPFC during health message exposure also showed greater declines in sedentary behavior after the scan \( (γ_{VMPFC \times time} = −0.006, t = −3.04, P = 0.002; \text{Fig. } 2B) \). Finally, we examined whether the variance explained by neural activity overlapped with that explained by participants’ self-reports following the intervention. We observed significant relationships between self-reported standards and behavior change and between self-reported attitudes and behavior change (SI Results). To determine whether neural activity during health messages captures different information that predicted by self-report measures, we next examined whether the relationships observed between affirmation condition and changes in sedentary behavior as well as neural activity in VMPFC during message exposure and behavior change held controlling for these measures (e.g., attitudes toward physical activity; self-standards as someone who can increase physical activity). All previously observed relationships between affirmation, neural activity, and behavior change remained significant when controlling for our attitude and self-standard measures. This suggests that the effects of affirmation and consequent neural activity in VMPFC during message exposure are explaining additional variance in the behavior change beyond those predicted by self-reports.

Discussion

Self-affirmation has been effective in augmenting interventions across a number of domains; however, the neural mechanisms supporting its effects have not been studied (29). We manipulated exposure to an fMRI-compatible affirmation intervention before exposure to health risk messages in a group of sedentary adults (54). Our results demonstrated that participants who reflected on their highest value during the self-affirmation exercise

Fig. 1. Effect of affirmation on neural activity in VMPFC and on behavior change in the month following the scan. Participants who were affirmed showed (A) greater activity in VMPFC during exposure to health messages and (B) greater declines in sedentary behavior in the month following the scan than participants who were unaffirmed, controlling for baseline sedentary behavior and demographics.
A community sample of sedentary adults (\(\bar{y} = 67; 41\) females; 0.42–0.44) was recruited for a study on health message exposure, which in turn was associated with declines in sedentary behavior. These data are consistent with a model of affirmation that emphasizes increased receptivity to otherwise-threatening health information as a function of successful affirmation (29, 55). Our data add clarity to the picture of one aspect of what it means to be successfully affirmed; the neural results obtained emphasize the idea that self-affirmation may allow increased processing of potentially threatening health information as more self-relevant and valuable to at-risk individuals. Neural data cannot speak to the specific types of self-worth upon which individuals might draw, nor whether such increases in self-related processing stem from drawing on one or multiple distinct sources of self-worth. They also do not address whether increases in self-processing stem from maintaining prior levels of self-worth in the face of new information that otherwise would pose a threat (55), or why thinking about important values should increase self-relevance and valuing of potentially threatening information (although others have offered evidence concerning these mechanisms; e.g., refs. 34 and 56–58). Further development of mappings between theory and evidence at different levels of analysis will allow more specific linkage of psychological and neural evidence.

More immediately, however, the neural results do link pathways of successful affirmation with other successful methods for health intervention. For example, these results also speak to potential mechanisms explaining prior findings that demonstrate relationships between neural activity in VMPFC and health behavior change (e.g., refs. 42, 45, 46, and 59). The current study provides the first experimental evidence (to our knowledge) that changing activity within VMPFC alters subsequent trajectories of health behavior change. In past work, neural activity in VMPFC during exposure to messages designed to increase sunscreen use (45) and decrease smoking (46, 59) has predicted message-consistent behavior change up to 1 mo after message exposure, but these results have been correlational. The experimental manipulation of neural activity in VMPFC via a self-affirmation paradigm adds substantially to prior findings. It is possible that, in past work, those individuals whose responses to health messages were spontaneously more like those of our affirmed participants (i.e., experiencing self-relevant value in health messages) would have the most success in changing their behaviors. Furthermore, neural responses within VMPFC have forecasted population-level public health campaign success (44). It is possible that, despite not containing explicit affirmations, messages that elicit less defensiveness and more self-related processing across participants are also those that go on to have success at larger scales.

Thus, the current study adds additional confidence concerning psychological processes that may underpin previously observed effects of self-affirmation and also suggests an experimental method for altering activity within VMPFC during other types of interventions that may be of use to investigators in other domains. The use of a scanner-compatible affirmation paradigm allowed us to uncover mechanisms of affirmation on subsequent receptivity to health messaging measured objectively via accelerometers. The study’s longitudinal design also allows examination of affirmation effects that are initially triggered during the scanned intervention and reinforced over time via short message service (SMS) text messages. As with any neuroimaging study, the psychological functions of the neural activity observed should be interpreted with caution pertaining to reverse inference as the VMPFC serves many psychological functions (60). Our strong a priori hypothesis, and convergence with prior research on behavior change, however, suggests that the explanations offered are parsimonious.

### Conclusion

In sum, self-affirmation is a method for up-regulating activity within the VMPFC during exposure to health messages. The present findings support a model in which affirmation allows people to see otherwise-threatening information as more self-relevant and valuable. Affirmed (compared with unaffirmed) participants showed greater activity within VMPFC during exposure to targeted health messages and the degree of this activity predicted the trajectory of objectively measured sedentary behavior in the subsequent month. These findings begin to build a picture of the neural mechanisms of affirmation and suggest promise in creation of interventions that prime participants to view information in ways that they can internalize.

### Methods

**Participants.** A community sample of sedentary adults (\(n = 67; 41\) females; mean age, 33.42 yr; SD, 13.04; 44 white, 12 black, 3 Asian, 1 Hispanic, 7 other) was recruited for a study on “daily activities” to avoid biasing recruitment in favor of people who would want to sign up for a physical activity study.

### Table 1. Whole-brain results comparing neural activity during health message exposure for affirmed > unaffirmed participants, \(P < 0.005, k = 20\)

<table>
<thead>
<tr>
<th>Region</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>k</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ventral striatum</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>33</td>
<td>3.41</td>
</tr>
<tr>
<td>Posterior cingulate</td>
<td>−9</td>
<td>−36</td>
<td>13</td>
<td>78</td>
<td>3.41</td>
</tr>
<tr>
<td>Precuneus</td>
<td>1</td>
<td>−70</td>
<td>37</td>
<td>31</td>
<td>3.49</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>15</td>
<td>36</td>
<td>55</td>
<td>26</td>
<td>3.22</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>−60</td>
<td>5</td>
<td>−8</td>
<td>33</td>
<td>3.64</td>
</tr>
</tbody>
</table>

No cerebral activity was greater for unaffirmed > affirmed participants.

Fig. 2. (A) VMPFC ROI. (B) Participants who showed higher levels of VMPFC activity during exposure to health messages subsequently decreased their sedentary behavior more in the month following the scan, controlling for baseline sedentary behavior and demographics.
and might be less defensive (Table 2). To be included at the baseline screening, participants had to report engaging in less than 195 min per week of walking, moderate, and vigorous physical activity (using short-form International Physical Activity Questionnaire (IPAQ) criteria; mean reported minutes of activity at intake, 123.53; SD, 49.52). Participants also met standard fMRI safety criteria (no metal in body, not claustrophobic, not pregnant) and were right handed. Participants with histories of major health problems or mental illness were excluded. On average, participants were overweight [mean body mass index (BMI), 27.99; SD, 6.84; range, 18.2–54.86]. Due to attrition, the final sample consisted of 67 participants at T1, 61 at T2, and 60 participants at T3. In addition, we lost data from an additional 15 subjects due to excessive movement (n = 1), reporting technical difficulties in scanning (n = 1), equipment failure (n = 11), or damage (n = 1), resulting in a final sample of 45 participants with both neuroimaging and accelerometer data. Years of education was not reported for three participants and age was not reported for one participant. These participants were thus excluded from models collating these variables; results remain substantively unchanged, however, with or without these participants. This research was approved by the institutional review board at the University of Michigan.

Procedure (Fig. 3). During screening, participants answered self-report measures of their exercise behavior during the week prior to IPAQ to identify sedentary adults most in need of intervention (and most likely to be defensive in response to risk messages). They also reported their weight and height from which BMI scores were derived. Eligible participants were recruited to complete a baseline appointment (T1), an fMRI appointment (T2) approximately 1 wk later [mean (M), 9.35 d; SD, 6.16], and an endpoint appointment (T3) approximately 1 mo after T2 (M, 35.92 d; SD, 7.19).

At T1, participants completed an initial values ranking that was used in the affirmation intervention, a range of individual difference measures (SI Methods), and baseline measures to calibrate their later activity. More specifically, during the T1 appointment, they were fitted with a wrist-worn accelerometer device used during the duration of the study to monitor participants as they completed a range of activities including walking at their usual pace along a hallway, climbing stairs, and sitting for at least 30 min (calibration). Participants continued to wear the wrist-worn accelerometers for the week between T1 and T2, which served as their baseline (preintervention) activity period.

During the T2 fMRI session, participants completed a series of tasks (described below) including a values-based self-affirmation or control task, and a health message intervention. All tasks were presented on a scanner-compatible screen at 800 x 600-pixel resolution using Presentation (NeuroBehavioral Systems), and responses were collected using a five-button response device attached to the participant’s right wrist.

After the T2 intervention, participants continued to wear their accelerometers for an additional month. In addition, participants received one value affirmation and one health message per day, drawn from the same value and health messages shown during the primary health message intervention, via their mobile phones for a month leading up to their third visit. At the final T3 endpoint visit, participants returned their accelerometers, completed a final set of surveys, and were debriefed, paid, and thanked for their participation.

Measures. Physical activity behavior. Our primary outcome of interest was changes in objectively measured sedentary behavior using accelerometers; given our focus on sedentary adults, and the fact that exchanging sedentary for even light activity is known to have physiological and psychological benefits (17–20), we focused on the proportion of each day that participants were sedentary. More specifically, we collected accelerometer data during a baseline measurement week and for 1 mo following the intervention (Fig. 3) using a triaxial GENEA accelerometer (61) worn on the left wrist (all participants are right-handed; see SI Methods for details). Subjects were encouraged to maintain ≥247 wear of the water-proof accelerometers for the baseline week before the fMRI appointment and during the month following (62–65).

We defined sedentary behavior according to measurements taken during the T1 laboratory calibration in which participants performed a number of activities including at least 30 min of sedentary activities such as completing surveys while seated at a computer terminal; the peak acceleration during this 30-min period was used to determine appropriate cut points for each participant such that activity below that threshold was tagged as “sedentary.” Using the sedentary cut points defined during the T1 laboratory session, we computed the proportion of each day that participants were sedentary. Due to attrition, the final sample consisted of 67 participants at T1, 61 at T2, and 60 participants at T3. In addition, we lost data from an additional 15 subjects due to excessive movement (n = 1), reporting technical difficulties in scanning (n = 1), equipment failure (n = 11), or damage (n = 1), resulting in a final sample of 45 participants with both neuroimaging and accelerometer data. Years of education was not reported for three participants and age was not reported for one participant. These participants were thus excluded from models collating these variables; results remain substantively unchanged, however, with or without these participants. This research was approved by the institutional review board at the University of Michigan.

Fig. 3. Overall study design.
people who sit less are at lower risk for certain disease
people who sit less are less likely to have diabetes, and heart disease than people who are more active
supporting claim (7 sec)
initial suggestion (5 sec)
press when applied (6 sec)
customization (6 sec)
Fig. 4. fMRI task design.

Analysis.

fMRI data acquisition and analysis. The imaging data were acquired on a 3-T GE Signa MRI scanner. During the acquisition of the high-resolution structural images, participants were asked to complete a mental self-affirmation task following prompts relating to their highest (experimental group) or lowest (control group) value. This was immediately followed by one functional run of the affirmation task (323 volumes total) and two runs of the message task (308 volumes each; 616 volumes total). (Note: For the first six participants, a slightly longer version of the task was used, in which the affirmation task was split into two runs of 209 volumes each and the physical activity task was split into three runs of 257 volumes each.) The data were acquired and preprocessed using a standard processing stream (see SI Methods for details).

Fixed-effects models of health message exposure were constructed in SPM8 for each participant that included regressors for each message type (risk messages, how to be active, how to be less sedentary, why to be active, why to be less sedentary, how to perform other daily activities, why to perform other daily activities) and the corresponding response periods. Movement parameters (a total of six rigid-body parameters, three for translation and three for rotation) derived from spatial realignment were also included as nuisance regressors in all first-level models. Data were high-pass filtered with a cutoff of 128 s. Contrasts were computed, averaging across the 50 health messages focusing on being more active and less sedentary, and comparing activity during those messages to rest. Second-level random-effects models were constructed that averaged across participants and were subjected to further ROI and between-groups analysis (described below).

ROI analysis and prediction of health behavior change. Our primary hypothesis was focused on a subregion of VMPFC that has been associated with health behavior change in a number of prior investigations (45, 46, 59). This VMPFC ROI encompasses 1,232.00 mm² at the border of Brodmann areas 10 and 11 (Fig. 2A). Parameter estimates of activity during the 50 health messages compared with rest were extracted using MarsBar (67), an ROI toolbox for SPM. We then computed a series of models in R (68) that examined our hypothesis regarding relationships between self-affirmation, neural activity in VMPFC, and changes in sedentary behavior in the month following the scan. All models controlled for centered baseline levels of sedentary behavior and demographic variables (centered age, sex, centered years of education, ethnicity). All time series mixed-effects models account for nonindependence of data within participants using lmer (69), and allow random-effect variable intercepts for participant and day postscan. More specifically, we hypothesized that those in the affirmation condition compared with the control condition would show greater behavior change following exposure to the messages. To test this hypothesis, we examined affirmation condition (affirmed vs. unaffirmed) as a predictor of the trajectory of sedentary behavior in the month following the scan, controlling for baseline levels of sedentary behavior and demographic covariates using linear mixed models [lme4 and ImeRtest packages in R (69, 70)]. Next, we hypothesized that one mechanism leading to behavior change is that affirmation allows people to see otherwise-threatening information as more self-relevant and valuable; as such, delivering self-affirmation before health messages should increase neural activity in VMPFC during message exposure. To test this hypothesis, we compared neural activity in our VMPFC ROI during exposure to the messages (relative to rest) between participants in the affirmation and control conditions, adjusting for demographics and baseline sedentary behavior, using linear regression in R. Finally, we hypothesized that this same neural activity in VMPFC during message exposure would predict changes in sedentary behavior following the intervention. To test this hypothesis, we examined neural activity in our VMPFC ROI during message exposure as a predictor of the trajectory of sedentary behavior in the month following the scan, controlling for baseline levels of sedentary behavior and demographic covariates using linear mixed models [lme4 and ImeRtest packages in R (69, 70)]. Finally, to examine whether neural activity overlaps with what is explained by self-report measures common in major theories of health behavior change, we examined each of our collected self-report variables as a potential predictor of sedentary behavior change in models alone, and with VMPFC.

Acknowledgments.

We thank Holly Derry, Ian Moore, and Michele Demers for assistance in developing intervention materials, and the staff of the University of Michigan fMRI Center for support and assistance with fMRI data acquisition. We thank Angela Fagerlin, Thad Polk, and the Michigan Center of Excellence in Cancer Communications Research for support in realizing this project, and Sonya Dal Cin and Sara Konrath for helpful discussions. We thank Kristin Shumaker, Nicolette Gregor, and Alison Sagon for assistance with data collection, Jonathan Mitchell for advice regarding the trajectory of sedentary behavior in the month following the scan, controlling for baseline levels of sedentary behavior and demographic covariates using linear mixed models [lme4 and ImeRtest packages in R (69, 70)]. Further, to examine whether neural activity overlaps with what is explained by self-report measures common in major theories of health behavior change, we examined each of our collected self-report variables as a potential predictor of sedentary behavior change in models alone, and with VMPFC.
70. Kuznetsova A, Brockhoff PB, Christensen RHB (2014) lmerTest: Tests for random effects for linear mixed effect models (lmer objects of lme4 package). Available at CRAN.R-project.org/package=lmerTest.