

# Self-affirmation enhances processing of negative stimuli among threat-prone individuals

Anna J. Finley,<sup>1</sup> Adrienne L. Crowell,<sup>2</sup> and Brandon J. Schmeichel<sup>1</sup>

<sup>1</sup>Texas A&M University, College Station, TX, USA, and <sup>2</sup>Hendrix College, Conway, AR, USA

Correspondence should be addressed to Brandon Schmeichel, Department of Psychological and Brain Sciences, Texas A&M University, MS 4235, College Station, TX 77840, USA. E-mail: bjschmeichel@gmail.com

## Abstract

Self-affirmation reduces defensive responding to self-threats. The present study extended beyond self-threats to assess affirmation's influence on responses to negative emotional pictures as measured by the late positive potential (LPP), an event-related potential in the encephalogram that reflects motivational significance. Participants completed a trait measure of behavioral inhibition system (BIS) sensitivity. Then they affirmed (or did not affirm) a core personal value before viewing a series of emotionally evocative pictures. Affirming a core value increased LPP responses to negative emotional pictures among individuals higher in BIS. Self-affirmation thus appeared to alter the motivational significance of negative pictures among threat-prone individuals, consistent with a reduction in the defensive avoidance of aversive stimuli. These findings suggest that affirming values may influence responses associated with basic (non-self) motivational systems among individuals sensitive to threat.

**Key words:** self-affirmation; late positive potential; behavioral inhibition system

## Introduction

Simply reflecting upon an important personal value may affirm the integrity of the self-concept and reduce defensive motivations. Numerous experiments have found that thinking and writing about core values reduces defensiveness in response to self-threats (Steele, 1988). For example, individuals have been observed dismissing or ignoring threatening health information, but this tendency may be reduced by expressing a meaningful personal value or characteristic before encountering the threat (Harris and Napper, 2005). Affirming the self may also reduce defensive responses to aversive events that do not threaten the self-concept (i.e. startling noises and images of snakes and spiders; Crowell et al., 2015). This evidence suggests that self-affirmation not only bolsters the self-concept but also impacts more basic defensive motivational systems. The present study extended this line of investigation by testing the hypothesis that self-affirmation influences the processing of negative emotional images as measured by electroencephalographic (EEG) activity.

## Self-affirmation and psychological self-defense

One common means of affirming the self is to think and write about one's most cherished values or personal characteristics (see McQueen and Klein, 2006). Affirming core values changes how individuals react to an assortment of self-threats, including uncertainty about a personal dilemma (McGregor et al., 2001), mortality salience (Schmeichel and Martens, 2005), losing a sports competition (Sherman and Kim, 2005), unrealistically high performance expectations (Siegel et al., 2005), stereotype threat (Martens et al., 2006) and relational insecurity (Stinson et al., 2011). The results of these studies have consistently pointed to the conclusion that self-affirmation 'can sharply attenuate defensive biases, and encourage attitude and behavior change in potentially threatening or contentious domains' (Sherman and Cohen, 2002, p. 122).

Evidence that affirming the self increases the processing of self-threatening information is among the clearest evidence that affirmation reduces defensiveness. Individuals tend to be

Received: 20 November 2017; Revised: 17 May 2018; Accepted: 24 May 2018

© The Author(s) (2018). Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

disinclined to contemplate information that threatens the self-concept, but self-affirmed individuals appear to tolerate and even elaborate upon self-threatening information relative to non-affirmed individuals. Participants in one study, for example, encountered evidence that challenged their beliefs about capital punishment (Cohen et al., 2000). Self-affirmed participants were more persuaded by the counter-attitudinal information than participants who had not self-affirmed, suggesting that self-affirmation increases openness to belief-threatening information. In another study, participants were induced to fail at a cognitive task (Vohs et al., 2013). Those who had previously affirmed an important personal value reported feeling less intelligent and less efficacious after failure relative to non-affirmed participants, again suggesting that self-affirmation enables individuals to tolerate and process self-threatening information they may otherwise ignore or dismiss.

Most relevant for the current research is a study that tested the effects of self-affirmation on attention to self-threatening words in a sample of women who were moderately heavy alcohol consumers (Klein and Harris, 2009). In this study, participants viewed a health message linking alcohol consumption to breast cancer. Some participants affirmed a core value prior to reading the message and some did not. Then all participants completed a dot probe task assessing attention to self-threatening words embedded in the health message. Non-affirmed female drinkers showed an attentional bias away from threat words, suggesting avoidance or distancing from the self-threat. Affirmed participants, in contrast, showed a bias in attention toward threat words, consistent with orientation of attention toward self-threat under self-affirmation.

The current study assessed threat processing at the neural level. In testing the hypothesis that self-affirmation increases threat processing, we sought to extend prior research in three ways. First, rather than examining self-threats, the current study tested the effect of self-affirmation on the processing of negative emotional images that do not directly threaten self-concepts. Specifically, we examined neural responses to images of snakes, spiders, mutilated bodies, and other aversive stimuli commonly used to activate the defensive motivational system (Bradley et al., 2001). We reasoned that insofar as self-affirmation increases openness to threatening information, affirmation may increase processing of negative emotional images. Evidence to support this hypothesis would extend self-affirmation theory beyond threats to self-regard to include threatening events more generally. Second, we tested an unselected sample of both male and female participants, as opposed to the female drinkers tested in the study by Klein and Harris (2009), and thus sought to extend the generality of the predicted self-affirmation effects. Third, we tested the effects of affirmation on an electrocortical measure of threat processing. The current study thus moved beyond prior research to assess affirmation's effects on neural correlates of aversive picture processing.

### The late positive potential

Scalp electrical potentials captured by EEG recordings reliably measure emotive processing. The scalp potential of interest in the current study was the late positive potential (LPP). The LPP is an event-related potential modulated by the motivational significance of an event. Numerous studies have found that the LPP is larger when viewing positive and negative emotional pictures as compared to neutral pictures (e.g. Cuthbert et al., 2000; Schupp et al., 2000; Hajcak and Olvet, 2008). The LPP is also

enhanced for personally relevant stimuli (e.g. Bayer et al., 2017). The LPP is thus thought to reflect the processing of motivationally significant events (Hajcak et al., 2010).

The current study represents the first attempt to test the hypothesis that self-affirmation influences the processing of negative emotional images at the neural level. We predicted that affirming the self would increase the LPP to negative images, suggesting increased motivational significance of aversive stimuli. Put differently, we expected affirmed participants to be more inclined to process and engage with the aversive images, rather than defend against, deny or otherwise downplay their motivational significance. As described below, we also anticipated that individual differences in the sensitivity of the defensive motivational system would influence responding to aversive images and thus may interact with self-affirmation to influence LPP magnitudes.

We found only one prior study testing the effects of self-affirmation on event-related potentials, and that study found evidence consistent with our hypothesis. More specifically, a study by Legault et al. (2012) assessed the effect of self-affirmation on the error-related negativity, an event-related potential that occurs within milliseconds of the commission of an error and is thought to reflect the activation of defensive motivations (e.g. Hajcak and Foti, 2008). Affirmed participants in that study exhibited larger error-related negativities, which is congruent with the idea that self-affirmation increases openness to aversive information.

### Individual differences in behavioral inhibition system sensitivity

The defensive motivational system is a neurobiological system that underpins responses to threats (e.g. predators, illness, injury), and can be contrasted with the appetitive motivational system, which is activated by sustenance and other rewards (e.g. water, nutrients, sexual partners; Lang and Davis, 2006; Lang and Bradley, 2013). An extensive body of research has revealed that threatening images elicit greater activation in the defensive motivational system than do neutral or positive images (e.g. Lang et al., 2008). The initial response to threatening images is to orient attention toward them, presumably to aid information intake and prepare the organism for an adaptive response to the threat. This orienting response (including the LPP) dissipates rather quickly for stimuli with low motivational relevance, whereas aversive or threatening stimuli that maintain motivational relevance elicit a more sustained response (e.g. Löw et al., 2008).

Some persons respond more intensely than do others to threatening images. In particular, individual differences in behavioral inhibition system (BIS) sensitivity moderate the intensity of emotional responding to threatening stimuli. According to Gray (1976, 1982), BIS is activated by aversive or threatening situations to prepare individuals for survival-related actions and shares conceptual overlap with the defensive motivational system. Simply put, persons higher in BIS are more sensitive to threats and show greater activation of the defensive motivational system in response to threatening stimuli (e.g. Carver and White, 1994; Leen-Feldner et al., 2004).

Most relevant for present purposes is evidence that persons higher in BIS sensitivity exhibit more extreme neural responses to threatening images. Specifically, negative emotional images elicit larger LPPs among individuals higher in BIS sensitivity (Balconi et al., 2012) and associated traits (e.g. trait anxiety; MacNamara and Hajcak, 2009). BIS and other anxiety-related

traits may also influence the time course of threat processing. Specifically, BIS may be associated with enhanced threat processing at relatively early stages of the LPP (i.e. prior to 1000 ms; Balconi et al., 2012) but reduced threat processing at later stages of the LPP (i.e. after 1000 ms). This pattern of change over time suggests initial engagement and attention to threatening stimuli among those higher in BIS, followed by subsequent avoidance of threatening information or defensive disengagement from the motivational significance of the threat (e.g. Mathews and MacLeod, 1994; Williams et al., 2004; Li et al., 2007).

Given the links between trait BIS and threat processing and self-affirmation and threat processing, we expected main effects of BIS and affirmation to be qualified by an interaction between the two variables in predicting the LPP to threatening images.<sup>1</sup> Specifically, we reasoned that self-affirmation would increase the motivational significance of threat-related (but not neutral) stimuli, and that this affirmation effect would be most pronounced among those prone to defend against threat (i.e. those higher in BIS). We considered two possible effects of self-affirmation on LPP magnitudes among high BIS individuals. The first possibility was that self-affirmation would magnify activity to threatening images during the early phase of the LPP among those higher in BIS. By this account, individuals higher in BIS show an enhanced LPP early in negative picture processing, and self-affirmation increases this enhanced LPP even more. The second possibility was that self-affirmation would prolong the LPP to negative images among those higher in BIS, suggesting more sustained processing to threatening images over time (i.e. during later phases of the LPP). Either form of enhanced LPP magnitudes would be consistent with the idea that self-affirmation enhances the motivational significance of threatening stimuli particularly for individuals who are most prone to defensive responding.

### The current study

Participants self-affirmed or not and then viewed a series of emotionally charged images while electrocortical activity was measured using EEG. Based on previous evidence we expected to find a positive relationship between trait BIS and the LPP to negative images (Balconi et al., 2012). Additionally, we hypothesized that self-affirmation would enhance the motivational relevance of negative images. This enhancement may manifest either as greater mean activity at LPP onset (during the time window 500–1000 ms after the picture appeared onscreen) or as a more sustained LPP over time (extending up to 3500 ms into picture viewing), particularly among more threat-prone persons (i.e. those higher in BIS). Participants also viewed positive images, but we did not make specific predictions regarding the effects of self-affirmation on LPP responses during positive images. The exploratory results pertaining to positive images are presented in the [Supplementary Material](#).

## Materials and methods

### Participants

One hundred and ten undergraduate students (54 women and 56 men; age  $M = 19.01$ ,  $s.d. = 0.96$ ) completed the experiment in exchange for credit toward a course requirement. Six additional participants completed the study but were excluded from

analyses for the following reasons: two had bad EEG recordings due to a malfunctioning grounding electrode, two had missing picture viewing data due to computer errors, one had missing questionnaire data due to computer errors and one completed the affirmation task in the incorrect order (i.e. after the picture viewing task) due to experimenter error.

### Materials and procedures

After providing informed consent, participants completed the BIS/BAS scales (Carver and White, 1994). In this sample, the average total score on the BIS subscale was  $M = 20.24$  ( $s.d. = 3.72$ ,  $\alpha = 0.77$ ).<sup>2</sup> After participants completed the questionnaire an experimenter attached sensors to participants' heads using 59 tin electrodes in a stretch-lycra electrode cap. Electrodes were also placed on participants' earlobes for offline re-referencing. EEG electrode impedances were kept below 5000 k $\Omega$ , and differences in impedance at homologous sites were kept below 1000 k $\Omega$ .

**Self-affirmation manipulation.** After cap placement and a 4-min period for recording electrical activity in the brain at rest, participants completed the self-affirmation manipulation. Participants spent five minutes thinking and writing about personal values. Participants in the 'no affirmation condition' ( $n = 52$ ) reviewed (but did not rank) a list of 12 values and personality characteristics and then wrote an essay on why artistic appreciation (i.e. one of the items on the list) may be important to other people.<sup>3</sup> Participants in the 'self-affirmation condition' ( $n = 58$ ) reviewed the same list of values, ranked them in order of personal importance and then wrote an essay explaining why their highest ranked value is important to them (a method borrowed from Cohen et al., 2000). Thus, one group of participants reflected upon a core personal value and the other group did not.

**Picture viewing paradigm.** Following the self-affirmation manipulation participants viewed a series of images on a computer screen. Participants were instructed to view pictures the entire time they appeared on screen. Trials began with a 3-s fixation cross, followed by a picture for 3.5–6 s and an intertrial interval (ITI) of 6–12 s. Picture stimuli consisted of 60 images from the International Affective Picture System (IAPS; Lang et al., 2008). The first four images were neutral practice trials and were not analyzed. Experimental trials included 18 positive, 19 neutral and 19 negative pictures from the IAPS.<sup>4</sup>

2 Participants also completed the BAS scale ( $M = 40.77$ ,  $s.d. = 3.97$ ,  $\alpha = 0.67$ ). Results pertaining to BASAS sensitivity are not relevant to the current investigation and are not reported here.

3 Prior research has found that participants rarely if ever select artistic appreciation as their top-ranked value (see Crowell et al., 2015). In the current sample, among participants who ranked the values, artistic appreciation had an average ranking of 9.85 out of 12 and a modal ranking of 12 out of 12.

4 Images were selected to be middling in valence and low in arousal (i.e. neutral), low in valence and high in arousal (i.e. negative) or high in valence and high in arousal (i.e. positive). Neutral pictures from IAPS (valence  $M = 5.08$ , arousal  $M = 3.03$ ): 2190, 2393, 2394, 2397, 2506, 2516, 2850, 5534, 7000, 7009, 7025, 7035, 7053, 7100, 7161, 7180, 7185 and 7236. Negative pictures from IAPS (valence  $M = 2.63$ , arousal  $M = 6.44$ ): 1052, 1205, 1270, 1300, 2811, 3000, 3022, 3071, 3130, 3150, 3250, 3400, 3550, 6230, 6550, 6560, 7380, 9300 and 9405. Positive pictures from IAPS (valence  $M = 6.90$ , arousal  $M = 5.96$ ): 4608, 4651, 4656, 4658, 4659, 4670, 4681, 4695, 5621, 7200, 7260, 7350, 7390, 7460, 7470, 8031, 8161, 8186 and 8260. Images were selected to ensure each valence category contained images of people.

1 More precisely, given that picture type was a factor in our analysis, we predicted an interaction among BIS, affirmation condition and image valence (negative vs neutral).

Following the image viewing task participants sat quietly for 4 min to permit another recording of resting brain activity. Participants then completed a modified flanker task adapted from Eriksen and Eriksen (1974). The flanker task measured individual differences in neural responses to errors; results associated with this task are not presented here.

**Psychophysiological recording and quantification.** EEG signals were amplified with Neuroscan SynAmps2 (El Paso, TX, USA), band-pass filtered (0.05–100 Hz), notch filtered (60 Hz) and digitized at 500 Hz. Eye movements were recorded from the supraorbit of left eye. Artifacts (e.g. horizontal eye movements and muscle movements) were first removed by hand. Then, a regression-based eye movement correction was applied to correct vertical eye movements and blinks (Semlitsch et al., 1986), after which the data were again visually inspected to ensure proper correction.

The stimulus-locked LPP was computed for the picture task. All data were epoched 100 ms prior to the stimulus to 3500 ms after the stimulus and were filtered with a lowpass of 16 Hz at 12 db. Waveforms were baseline corrected using pre-stimulus activity. Average voltages at midline centroparietal sites (i.e. Pz, CPz, Cz, FCz, Fz) for each participant were calculated for each stimulus type (i.e. positive, negative and neutral images), and the LPP was quantified by the area under the curve in six 500 ms time windows (i.e. 500–1000, 1000–1500, 1500–2000, 2000–2500, 2500–3000 and 3000–3500 ms after picture onset).

## Results

The results are organized as follows. First, we examined LPP activity at each time window collapsing across affirmation condition and electrode site to verify the expected effect of picture type. Next, we assessed the effect of BIS on the LPP to negative images. Last, we used a multilevel model to test the effects of affirmation condition and BIS on LPPs to neutral vs negative pictures. Additional analyses examining the relationship between BIS and LPP during positive images and a multilevel model including all picture types is reported in the [Supplementary Material](#).<sup>5</sup>

### LPP as a function of picture type and time window

To ensure we replicated the standard emotion-modulated LPP response, we ran a 3 (Picture Type)  $\times$  6 (Time Window) repeated measures analysis of variance (ANOVA) collapsed across affirmation condition and electrode site. We observed main effects of picture type,  $F(2, 1070) = 17.16$ ,  $P < 0.001$ , and time,  $F(5, 1070) = 51.63$ ,  $P < 0.001$ , which were qualified by the Picture Type  $\times$  Time Window interaction,  $F(10, 1070) = 14.39$ ,  $P < 0.001$ . Planned comparisons found that positive and negative pictures elicited larger LPPs than neutral pictures at all time windows except for last two time windows from 2500 to 3000 ms,  $F(2, 214) = 2.49$ ,  $P = 0.086$ , and from 3000 to 3500 ms,  $F(2, 214) = 2.30$ ,  $P = 0.103$ , respectively (for all other time windows,  $ps < 0.015$ ). LPPs to negative pictures were larger than LPPs to positive pictures during the first time window from 500 to 1000 ms,  $t(107) = 5.382$ ,  $P = 0.001$ , and the last time window from 3000 to 3500 ms,  $t(107) = 2.64$ ,  $P = 0.009$ , respectively (for all other time

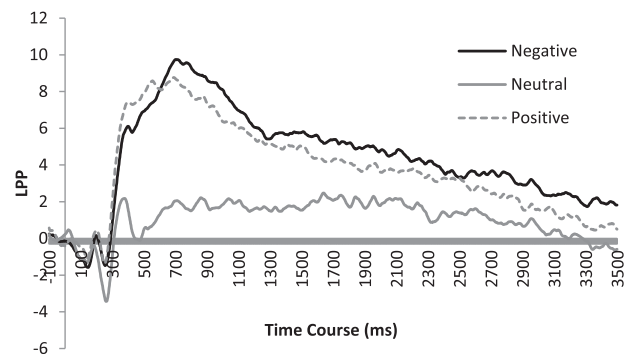


Fig. 1. LPP waveform by picture type, collapsing across affirmation condition, BIS, and electrode site.

windows,  $ps > 0.140$ ) (see Figure 1). This finding of larger LPPs to emotional vs neutral images and only occasional differences between LPPs to negative and positive images replicates prior research (e.g. Hajcak et al., 2007; Foti and Hajcak, 2008).

### LPP to negative images as a function of BIS at the early time window (500–1000 ms)

Next, we assessed the relationship between BIS and early-stage LPPs to negative images. We focused on the LPP at electrode site Pz during the earliest time window because this was when and where the LPP was maximal. We regressed mean-centered BIS scores onto LPPs to negative pictures, controlling for LPPs to neutral pictures and observed a significant effect of BIS,  $B = 0.634$ ,  $t(104) = 2.20$ ,  $P = 0.038$ , such that higher BIS scores predicted larger LPPs to negative pictures. Hence, consistent with past research (Balconi et al., 2012), persons higher (vs low) in BIS had larger initial LPP amplitudes to threatening images, consistent with the idea that high BIS individuals are more sensitive to threatening information.

### Multilevel modeling of LPP magnitudes as a function of BIS, affirmation condition, time window, and electrode site

We examined LPP amplitudes to negative vs neutral images in a multilevel model as a function of time window, time window-squared, electrode site, affirmation condition and BIS. More precisely, we constructed a three-level multilevel model with LPP magnitudes (i.e. area under the curve) nested within time (level 1), nested within image valence and electrode site (cross-classified level 2), nested within participant (BIS and affirmation condition; level 3). We included random intercepts to account for the dependence of LPP responses within participant, electrode site and valence. We also included random slopes for time and time squared (to allow for non-linear changes) within valence (level 2) and participant (level 3). Due to model convergence issues, random slopes for valence within participant (level 3) were not included in the model. We included cross-level interactions between electrode site and time because LPP magnitudes vary by electrode site and time since picture onset (e.g. Hajcak et al., 2007; Foti and Hajcak, 2008). We also included cross-level interactions among time, image valence (negative vs neutral), affirmation condition, and BIS. A diagonal covariance matrix and Satterthwaite estimation of degrees of freedom were used to compute the model in SPSS. For additional information, including the model equations, SPSS syntax and full model results, please see the [Supplementary Material](#).

5 Briefly, BIS did not relate to LPP magnitudes to positive images ( $P = 0.259$ ), and Multilevel model S1 found no effect of affirmation condition on LPPs to positive images ( $ps > 0.259$ ). See the [Supplementary Material](#) for additional information.



Overall, the three-level model explained a moderate portion of the variance in LPP responses at the level of individual time windows,  $\text{pseudo } R^2_{\text{Level 1}} = 0.116$ , picture type,  $\text{pseudo } R^2_{\text{Level 2 valence}} = 0.097$ , electrode site,  $\text{pseudo } R^2_{\text{Level 2 Site}} = 0.078$ , and individual participant,  $\text{pseudo } R^2_{\text{Level 3}} = 0.061$  (Cohen, 1992; Edwards et al., 2008).

The key hypothesis was tested in the four-way interaction among BIS, affirmation condition, image valence (negative vs neutral) and time-squared. This interaction was statistically significant,  $F(1, 39.24) = 14.58$ ,  $P < 0.001$ ,  $\text{semi-partial } R^2 = 0.271$ . See Figure 2. To break down the four-way interaction and to test our specific predictions regarding the effects of self-affirmation among threat-prone (high BIS) persons, we probed the simple effects of affirmation condition on LPPs to negative vs neutral images using predicted values at high (+1 s.d.) and low (−1 s.d.) levels of BIS at each time window (see Aiken and West, 1991). This is the multilevel model equivalent of conducting a simple slopes test in linear regression. As shown in Table 1, self-affirmation did not affect LPPs at any time window for participants low in BIS. Affirmation affected LPP magnitudes only for participants who scored high in BIS. Among high BIS individuals, the effect of affirmation was evident on LPPs at all time windows from 1000 to 3000 ms after image onset,  $ps < 0.04$ . As shown in Figure 2, affirmed individuals high in BIS showed a sustained LPP to negative vs neutral images compared to non-affirmed high BIS individuals.

## Discussion

The current study tested the hypothesis that affirming core values influences responding to negative emotional stimuli. Specifically, we considered two possible moderating effects of self-affirmation (vs no affirmation) on the LPP to negative pictures particularly among threat-prone individuals. One possibility was that affirmation would amplify early LPP responses. We found enhanced early LPP magnitudes to negative images among high BIS individuals, consistent with the idea that high BIS individuals are more attuned to threats (Balconi et al., 2012), but self-affirmation did not influence these early responses. Instead, consistent with the second possibility, self-affirmation influenced later LPP responses among high BIS individuals.

Hence, threat-prone individuals strongly registered the motivational significance of negative emotional images, exhibiting a magnified orienting response as revealed by the LPP (see Bradley, 2009). This exaggerated early LPP response quickly dissipated among non-affirmed individuals, which suggests disengagement or defensive distancing from aversive content later on in picture processing. Following self-affirmation, however, high BIS individuals showed a sustained LPP to negative images. This pattern suggests that high BIS individuals assigned greater, more sustained motivational significance to threatening images after self-affirmation.

The current findings are consistent with evidence that self-affirmation increases processing of threats to self-regard (e.g. Correll et al., 2004; Klein and Harris, 2009; Vohs et al., 2013). Like previous research, the current findings suggest that self-affirmation reduces defensive denial or disengagement and thereby increases threat processing. Unlike previous research, however, participants in the current study viewed negative pictures that posed little or no threat to their self-concepts, and self-affirmation enhanced neural responses to those pictures particularly among those prone to threat (i.e. high BIS individuals). The effects of self-affirmation thus appear to extend beyond responses to self-threats.

## Implications for self-affirmation theory

The current findings advance research and theory on self-affirmation in at least two ways. First, the current findings recommend expanding the scope of self-affirmation theory to include a broader array of threats than threats to the self-concept alone. Negative pictures do not pose a threat to self-regard, but they do elicit a defensive response, as evidenced by the increased LPP response to negative images among high BIS individuals. But this enhanced early LPP response quickly dissipates. This pattern of early engagement with and subsequent disengagement from negative images is consistent with attentional biases associated with anxiety-proneness and avoidance (e.g. Bar-Haim et al., 2007).

We found that self-affirmation alters LPP magnitudes in a manner that suggests sustained engagement with the aversive images instead of avoidance or disengagement in those participants who are most likely to disengage (i.e. high BIS individuals). The LPP results thus suggest that self-affirmation enables processing of threatening information even when the information does not directly impugn the self-concept. We propose that self-affirmation may act not only through self-esteem or identity-related mechanisms but also through basic systems underlying avoidance motivation. More precisely, we suggest that self-affirmation reduces defensive responding among those who respond most strongly to threats and thus increases the processing of threatening information by suppressing the defensive motivation system (see also Crowell et al., 2015).

This palliative or soothing effect would be consistent with prior research on self-affirmation's effects in the context of self-threats. Presumably, the threats to self-regard at issue in prior self-affirmation research (e.g. health threats; Klein and Harris, 2009) also trigger the defensive motivational system—the system we tried to activate by showing participants negative emotional images. Insofar as high BIS individuals have a more reactive defensive motivational system, it is plausible that high BIS individuals also show enhanced responses to self-threats. Future research may examine whether BIS is relevant for threats to self-regard and further explore the idea that self-affirmation suppresses defensive responding especially among those with more reactive defensive motivational systems.

Second, we found that self-affirmation influences the electroencephalogram. The current study is not the first to assess self-affirmation's effects on the brain, but it joins with only a handful of other studies to reveal neural aftereffects of self-affirmation. Previous studies using functional magnetic resonance imaging have found brain changes during self-affirmation tasks. For example, Cascio et al. (2016) found that activity in brain areas associated with positive valuation increases during self-affirmation. Another study assessing neural activity during self-affirmation found that making judgments about important values activates brain areas associated with reward processing (Dutcher et al., 2016). These studies suggest that affirming core values is rewarding. Perhaps the rewarding aspects of affirming the self helps to explain why threat-prone individuals showed a less engaged defensive motivational system during negative picture viewing in the current study (at least later in negative picture processing), insofar as increased activation of reward circuits during self-affirmation helps to suppress defensiveness. However, this idea is speculative, and we did not assess brain activity during the self-affirmation task.

Rather, we assessed brain activity after the self-affirmation manipulation, and, therefore, the current findings connect more clearly to prior studies of brain activity after self-affirmation.

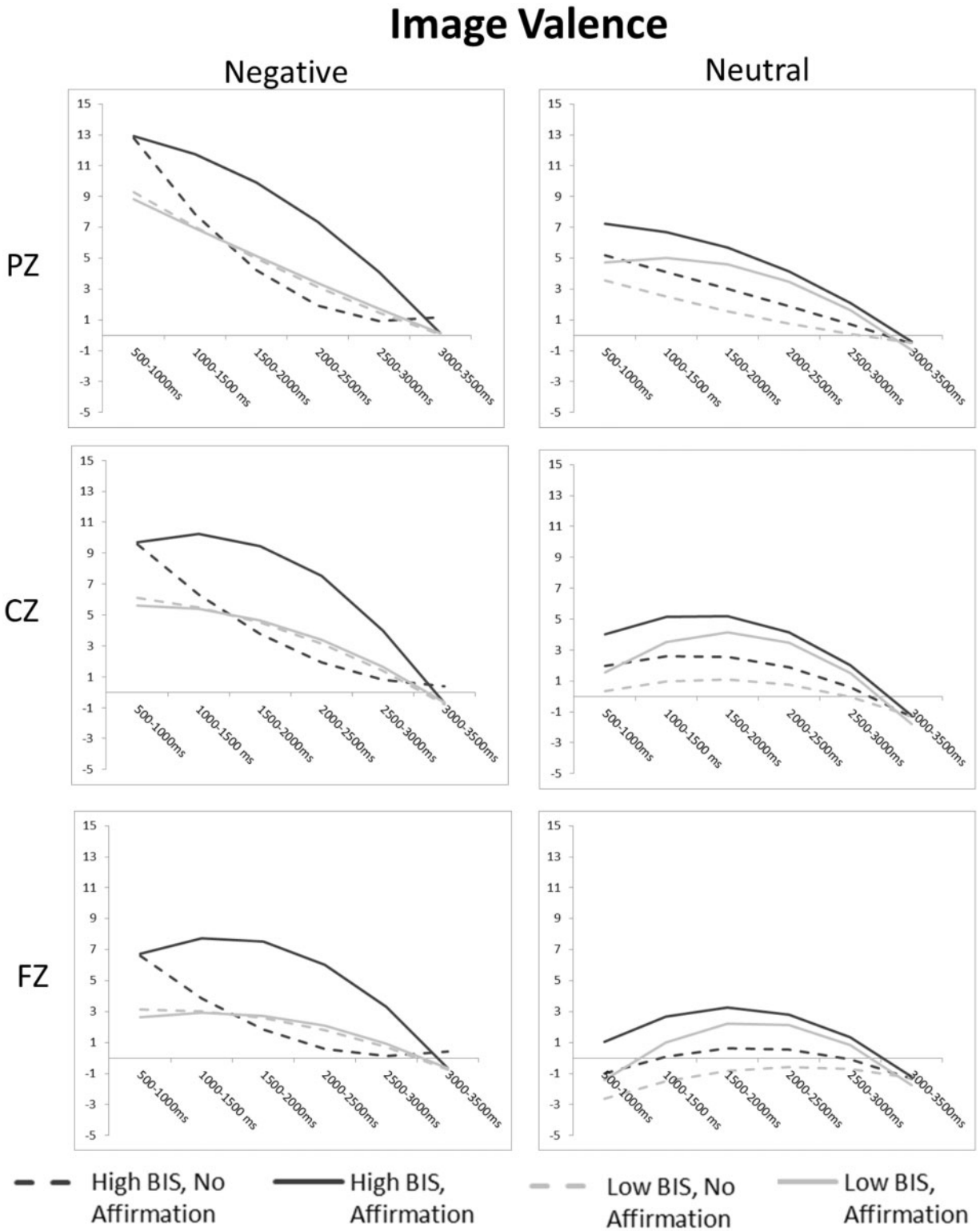


Fig. 2. Predicted values of LPP waveforms from the multilevel model at Pz, Cz, and Fz electrode sites by affirmation condition and high and low BIS for negative and neutral images. CPz and FCz are not included in the figure but show patterns similar the ones depicted here.

For example, Falk et al. (2015) found that ventromedial prefrontal cortex, a brain area related to self-processing, increases among affirmed individuals while viewing threatening health information. More relevant to the current research is the study by Legault et al. (2012) that found increased error processing (i.e. larger ERNs)

after self-affirmation. The current findings, along with those prior studies, indicate that affirming the self enhances attention to negative events as measured at the level of the brain. Future research should address the extent to which these neural changes can account for the attitudinal and behavioral effects

**Table 1.** Simple effects of self-affirmation condition at each time window for negative and neutral images at low and high BIS

	Simple Effect <i>b</i>	<i>df</i> <sub>denom</sub>	<i>F</i>	<i>P</i>
Negative images, low BIS (−1 s.d.)				
500–1000 ms	−0.24	181.05	0.07	.789
1000–1500 ms	−0.05	169.65	0.003	.960
1500–2000 ms	0.07	168.75	0.006	.940
2000–2500 ms	0.13	169.02	0.02	.884
2500–3000 ms	0.12	182.09	0.03	.856
3000–3500 ms	0.05	289.84	0.007	.932
Negative images, high BIS (+1 s.d.)				
500–1000 ms	0.06	181.37	0.004	.949
1000–1500 ms	1.95	169.84	4.45	.036
1500–2000 ms	2.84	168.93	8.49	.004
2000–2500 ms	2.72	169.25	9.66	.002
2500–3000 ms	1.59	182.49	5.94	.016
3000–3500 ms	−0.55	291.08	1.09	.297
Neutral images, low BIS (−1 s.d.)				
500–1000 ms	0.60	181.05	0.46	.499
1000–1500 ms	1.27	169.65	1.89	.171
1500–2000 ms	1.53	168.75	2.46	.118
2000–2500 ms	1.36	169.02	2.42	.121
2500–3000 ms	0.77	182.09	1.39	.239
3000–3500 ms	−0.23	289.84	0.20	.658
Neutral images, high BIS (+1 s.d.)				
500–1000 ms	1.03	181.37	1.36	.245
1000–1500 ms	1.29	169.65	1.57	.165
1500–2000 ms	1.33	168.93	1.85	.175
2000–2500 ms	1.13	169.25	1.68	.197
2500–3000 ms	0.71	182.49	1.17	.280
3000–3500 ms	0.04	291.08	0.007	.933

Note: All simple effects  $f_{\text{numerator}}=1$ . Because of Satterthwaite estimation,  $df_{\text{denom}}$  may not be whole numbers. Simple effects calculated based on full sample,  $N = 110$ .

associated with self-affirmation. One possibility is that the neural patterns observed here (i.e. increased LPP to aversive stimuli) and in prior research (e.g. increased ERNs) mediate the effects of self-affirmation on attitudinal and behavioral responses to threat.

### Limitations and future directions

We have assumed that the LPP indexes motivational significance, which led us to conclude that self-affirmation boosts the motivational significance of threatening images among high BIS individuals. But we did not include an independent, non-physiological measure of motivational significance nor did we include a self-report measure of perceptions of threat, so we cannot say with certainty that the LPP effects reflect changes in motivational significance as opposed to some other psychological process. A future study that includes a measure of motivational significance independent of the LPP and ratings of perceptions of threat could help to solidify the conclusion that self-affirmation encourages high BIS individuals to assign greater motivational significance to threatening stimuli.

Similarly, we have assumed that larger early LPP responses to negative pictures and smaller later LPP responses among non-affirmed high BIS individuals reflect early orientation of attention toward and subsequent disengagement from negative pictures. We have conceptualized disengagement of attention later in negative picture processing as a defensive response akin to denying health risks (Harris and Napper, 2005) or

downplaying the implications of negative feedback (e.g. Vohs et al., 2013). Those assumptions are consistent with Koole and Rothermund's (2011) perspective on emotion regulation, which claims that 'if the intensity (or direction) of an emotion changes over time, although the emotion eliciting stimulus remains the same, and no instruction regarding emotion regulation is given, then it is likely that spontaneous or implicit emotion regulation processes are responsible for the change' (p. 393). We suggest that individuals high in BIS engaged in spontaneous emotion regulation (e.g. attentional disengagement) to down-regulate their responses to the negative images, unless they had previously self-affirmed, in which case they continued to engage with the negative images. However, a study specifically designed to test this hypothesis is needed to support the idea that the smaller later LPP response to negative images among non-affirmed high BIS individuals is a regulatory maneuver akin to denying health risks or other threats to self-regard. Future studies linking LPP responses to behavioral responses to threats would be especially valuable.

On a related note, we did not track where participants looked when the pictures appeared onscreen. We therefore cannot be certain that looking time or gaze fixation points did not differ between conditions. However, we rejected artifacts in the EEG signal consistent with large horizontal ocular movements (suggesting averted gaze), and the LPP amplitudes we observed are consistent with past literature (i.e. enhanced during emotional images, shifting forward on the scalp over time, with a positive association between BIS and LPP amplitudes to negative images). We thus do not believe our participants looked at the images differently from participants in other similar studies, but a future study using eye-tracking technology could test these claims and provide additional insight into attentional engagement and disengagement dynamics under self-affirmation.

The sustained motivational significance (i.e. enhanced LPP) we observed in response to negative pictures among affirmed high BIS individuals may represent a maladaptive or counter-productive response, insofar as it interferes with adaptive defensive responding or increases negative emotional experience. Indeed, prior research has found that increased openness to threatening information among self-affirmed individuals can increase the negativity of self-evaluations (e.g. Vohs et al., 2013). But an increased openness or reduced defensiveness to threatening information may also help individuals to cope with threats directly rather than denying or distorting them (cf. problem-focused vs emotion-focused coping; Lazarus and Folkman, 1984). Additional research may shed light on the extent to which high BIS individuals benefit or suffer from assigning increased motivational significance to threatening stimuli.

### Conclusions

Self-affirmation appeared to sustain the motivational significance of threat-related stimuli among threat-prone individuals. This finding represents novel support for the idea, based on self-affirmation theory, that affirming core personal values reduces defensive responding, which would entail moving away from or disengaging from the threat. The results suggest that self-affirmation's effects extend beyond ego defenses to influence basic mechanisms of motivation and emotion. More research is needed into the consequences of increasing engagement with threatening stimuli under self-affirmation. For example, would self-affirmation be useful for phobic patients, who studiously avoid the triggers of their phobia? Acceptance and commitment therapy, a form of cognitive behavior therapy,

involves having clients discuss their core personal values in a manner not unlike the self-affirmation manipulations used in the current experiment (for an introduction, see Twohig, 2012). More research on the motivational and emotional consequences of self-affirmation may prove beneficial for threat-prone individuals.

## Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

## References

- Aiken, L.S., West, S.G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage.
- Balconi, M., Falbo, L., Conte, V.A. (2012). BIS and BAS correlates with psychophysiological and cortical response systems during aversive and appetitive emotional stimuli processing. *Motivation and Emotion*, *36*(2), 218–31.
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M.J., Van Ijzendoorn, M.H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, *133*(1), 1–24.
- Bayer, M., Ruthmann, K., Schacht, A. (2017). The impact of personal relevance on emotion processing: evidence from event-related potentials and pupillary responses. *Social Cognitive and Affective Neuroscience*, *12*(9), 1470–9.
- Bradley, M.M. (2009). Natural selective attention: orienting and emotion. *Psychophysiology*, *46*(1), 1–11.
- Bradley, M.M., Codispoti, M., Cuthbert, B.N., Lang, P.J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276–98.
- Carver, C.S., White, T.L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*(2), 319–33.
- Cascio, C.N., O'Donnell, M.B., Tinney, F.J., et al. (2016). Self-affirmation activates brain systems associated with self-related processing and reward and is reinforced by future orientation. *Social Cognitive and Affective Neuroscience*, *11*(4), 621–9.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–9.
- 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*(9), 1151–64.
- 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*(3), 350–6.
- Crowell, A., Page-Gould, E., Schmeichel, B.J. (2015). Self-affirmation breaks the link between the behavioral inhibition system and the threat-potentiated startle response. *Emotion*, *15*(2), 146–50.
- Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J. (2000). Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111.
- Dutcher, J.M., Creswell, J.D., Pacilio, L.E., et al. (2016). Self-affirmation activates the ventral striatum: a possible reward-related mechanism for self-affirmation. *Psychological Science*, *27*(4), 455–66.
- Edwards, L.J., Muller, K.E., Wolfinger, R.D., Qaqish, B.F., Schabenberger, O. (2008). An  $R^2$  statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, *27*(29), 6137–57.
- Eriksen, B.A., Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–9.
- Falk, E.B., O'Donnell, M.B., Cascio, C.N., et al. (2015). Self-affirmation alters the brain's response to health messages and subsequent behavior change. *Proceedings of the National Academy of Sciences*, *112*(7), 1977–82.
- Foti, D., Hajcak, G. (2008). Deconstructing reappraisal: descriptions preceding arousing pictures modulate the subsequent neural response. *Journal of Cognitive Neuroscience*, *20*(6), 977–88.
- Gray, J.A. (1976). The behavioural inhibition system: a possible substrate for anxiety. In: Feldman, M.P., Broadhurst, A.M., editors. *The Theoretical and Experimental Bases of Behaviour Modification*, pp. 3–41. London: Wiley.
- Gray, J.A. (1982). Precise of the neuropsychology of anxiety: an enquiry into the function of the septo-hippocampal system. *Behavioral and Brain Sciences*, *5*(3), 469–84.
- Hajcak, G., Dunning, J.P., Foti, D. (2007). Neural response to emotional pictures is unaffected by concurrent task difficulty: an event-related potential study. *Behavioral Neuroscience*, *121*(6), 1156–62.
- Hajcak, G., Foti, D. (2008). Errors are aversive defensive motivation and the error-related negativity. *Psychological Science*, *19*(2), 103–8.
- Hajcak, G., MacNamara, A., Olvet, D.M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology*, *35*(2), 129–55.
- Hajcak, G., Olvet, D.M. (2008). The persistence of attention to emotion: brain potentials during and after picture presentation. *Emotion*, *8*(2), 250–5.
- Harris, P.R., Napper, L. (2005). Self-affirmation and the biased processing of threatening health-risk information. *Personality and Social Psychology Bulletin*, *31*(9), 1250–63.
- Klein, W.M.P., Harris, P.R. (2009). Self-affirmation enhances attentional bias toward threatening components of a persuasive message. *Psychological Science*, *20*(12), 1463–7.
- Koole, S.L., Rothermund, K. (2011). “I feel better but I don't know why”: the psychology of implicit emotion regulation. *Cognition and Emotion*, *25*(3), 389–99.
- Lang, P.J., Bradley, M.M. (2013). Appetitive and defensive motivation: goal-directed or goal-determined? *Emotion Review*, *5*(3), 230–4.
- Lang, P.J., Bradley, M.M., Cuthbert, B.N. (2008). *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*. Technical Report a-8. University of Florida, Gainesville, FL.
- Lang, P.J., Davis, M. (2006). Emotion, motivation, and the brain: reflex foundations in animal and human research. *Progress in Brain Research*, *156*, 3–29.
- Lazarus, R.S., Folkman, S. (1984). *Stress, Appraisal, and Coping*. New York: Springer.
- Leen-Feldner, E.W., Zvolensky, M.J., Feldner, M.T. (2004). Behavioral inhibition sensitivity and emotional response suppression: a laboratory test among adolescents in a fear-relevant paradigm. *Journal of Clinical Child and Adolescent Psychology*, *33*(4), 783–91.
- Löw, A., Lang, P.J., Smith, J.C., Bradley, M.M. (2008). Both predator and prey: emotional arousal in threat and reward. *Psychological Science*, *19*(9), 865–73.
- 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**(12), 1455–60.
- Li, W., Zinbarg, R.E., Paller, K.A. (2007). Trait anxiety modulates supraliminal and subliminal threat: brain potential evidence for early and late processing influences. *Cognitive, Affective, & Behavioral Neuroscience*, **7**(1), 25–36.
- MacNamara, A., Hajcak, G. (2009). Anxiety and spatial attention moderate the electrocortical response to aversive pictures. *Neuropsychologia*, **47**(13), 2975–80.
- Martens, A., Johns, M., Greenberg, J., Schimel, J. (2006). Combating stereotype threat: the effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, **42**(2), 236–43.
- Mathews, A., MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology*, **45**(1), 25–50.
- McGregor, I., Zanna, M.P., Holmes, J.G., Spencer, S.J. (2001). Compensatory conviction in the face of personal uncertainty: going to extremes and being oneself. *Journal of Personality and Social Psychology*, **80**(3), 472–88.
- McQueen, A., Klein, W.M. (2006). Experimental manipulations of self-affirmation: a systematic review. *Self and Identity*, **5**(4), 289–354.
- Schmeichel, B.J., Martens, A. (2005). Self-affirmation and mortality salience: affirming values reduces worldview defense and death-thought accessibility. *Personality and Social Psychology Bulletin*, **31**(5), 658–67.
- Schupp, H.T., Cuthbert, B.N., Bradley, M.M., Cacioppo, J.T., Ito, T., Lang, P.J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology*, **37**(2), 257–61.
- Semlitsch, H.V., Anderer, P., Schuster, P., Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, **23**(6), 695–703.
- Sherman, D.K., Cohen, G.L. (2002). Accepting threatening information: self-affirmation and the reduction of defensive biases. *Current Directions in Psychological Science*, **11**(4), 119–23.
- Sherman, D.K., Kim, H.S. (2005). Is there an “I” in “team”? The role of the self in group serving judgments. *Journal of Personality and Social Psychology*, **88**(1), 108–20.
- Siegel, P.A., Scillitoe, J., Parks-Yancy, R. (2005). Reducing the tendency to self-handicap: the effect of self-affirmation. *Journal of Experimental Social Psychology*, **41**(6), 589–97.
- Steele, C.M. (1988). The psychology of self-affirmation: sustaining the integrity of the self. In: Berkowitz, L., editor. *Advances in Experimental Social Psychology*, Vol. 21, pp. 261–302. New York: Academic Press.
- 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**(9), 1145–9.
- Twohig, M.P. (2012). Acceptance and commitment therapy: introduction. *Cognitive and Behavioral Practice*, **19**(4), 499–507.
- Vohs, K.D., Park, J.K., Schmeichel, B.J. (2013). Self-affirmation can enable goal disengagement. *Journal of Personality and Social Psychology*, **104**(1), 14–27.
- Williams, L.M., Liddell, B.J., Rathjen, J., et al. (2004). Mapping the time course of nonconscious and conscious perception of fear: an integration of central and peripheral measures. *Human Brain Mapping*, **21**(2), 64–74.