How Self-Affirmation Reduces Defensive Processing of Threatening Health Information: Evidence at the Implicit Level

Guido M. van Koningsbruggen
Utrecht University

Enny Das
VU University Amsterdam

David R. Roskos-Ewoldsen
The Ohio State University

Objective: Self-affirmation reduces defensive responses to threatening health information, but little is known about the cognitive processes instigated by self-affirmation. This study tested whether self-affirmation increases responsiveness to threatening health information at the implicit level. *Design:* In an experimental study ($N = 84$), the authors presented high- (coffee drinkers) and low-relevance (noncoffee drinkers) participants with threatening health information linking caffeine consumption to health problems. Prior to reading this information, the authors manipulated self-affirmation. *Main Outcome Measures:* Participants completed an unobtrusive lexical decision task to measure the accessibility of threat-related cognitions and reported their perceptions of message quality and intentions to take precautions. *Results:* Among high-relevance participants, self-affirmation increased the accessibility of threat-related cognitions, increased perceptions of message quality, and promoted adaptive behavioral intentions. *Conclusion:* The findings suggest that self-affirmation can increase implicit responsiveness to threatening health information among a target audience, that is, people for whom the health information is highly relevant.

Keywords: self-affirmation, defensive processing, health risk information, accessibility, persuasion

People are often confronted with threatening information telling them they live an unhealthy lifestyle that poses a serious risk to their health. Although people should respond by adopting healthy behaviors or stop risky behaviors, people often persevere with their unhealthy behaviors (Sherman & Cohen, 2006). For instance, the more personally significant a health message is, the more people are likely to downplay the seriousness of the health risk, question the accuracy of the threatening information or evidence presented in the message, and process the information in a biased fashion (e.g., De Wit, Das, & Vet, 2008; Kunda, 1987; Liberman & Chaiken, 1992). Although such defensive processing of threatening health information is likely to keep worries at a distance, it can prevent people from protecting their personal health. Self-affirmation theory (Steele, 1988)—a general theory about how people deal with self-threats—may help understand why people respond defensively to threatening health information.

According to self-affirmation theory (Steele, 1988), people are highly motivated to protect and maintain a global sense of self-integrity. Threatening information arouses this motive and people respond defensively to satisfy it (e.g., they downplay, minimize, or avoid threatening health information). The theory further assumes that people are primarily concerned with their *global* sense of self-integrity. Consequently, they can also restore self-integrity by drawing on alternative sources unrelated to the threat, such as reflecting on another important value (Steele, 1988). Thus, when threatened in one domain (e.g., health), people can restore their global self-integrity by affirming another important domain (e.g., intelligence). Such a "self-affirmation" restores global self-integrity, thereby reducing the need to respond defensively to the specific threat.

Although more and more studies show that theory-based manipulations of self-affirmation increase acceptance of threatening health information among people who normally tend to be defensive (e.g., Harris, Mayle, Mabbott, & Napper, 2007; Harris & Napper, 2005; Sherman, Nelson, & Steele, 2000), there is little direct empirical evidence regarding the cognitive processes that self-affirmation instigates (Sherman & Cohen, 2006). A better insight into these processes is pivotal to increasing our understanding of the processes that potentially mediate self-affirmation effects as well as their consequences for the design of health information. Most pressing is the issue of the origins and consequences of defensive processing at the implicit level: What makes people "turn off" from threatening health information, and is it true that self-affirmation makes people "turn on" to threatening health information again? Implicit processes are important to understand because they influence both how information is processed (Arpan, Rhodes, & Roskos-Ewoldsen, 2007) and subsequent behavior (Olson & Fazio, 2009). The present study provides a first pub-

Guido M. van Koningsbruggen, Department of Social and Organizational Psychology, Utrecht University, The Netherlands; Enny Das, Department of Communication Science, VU University Amsterdam, The Netherlands; David R. Roskos-Ewoldsen, School of Communication, The Ohio State University.

Correspondence concerning this article should be addressed to Guido M. van Koningsbruggen, Department of Social and Organizational Psychology, Utrecht University, PO Box 80140, 3508 TC Utrecht, The Netherlands. E-mail: g.m.vankoningsbruggen@uu.nl

lished test of the effect of self-affirmation on people's implicit responsiveness to threatening health information by using an implicit, unobtrusive measure of information accessibility that provides more insight into the cognitive processes instigated by self-affirmation.

Self-Affirmation and Responses to Health Information

Self-affirmation can increase attention to and acceptance of threatening health messages, increase perceptions of personal risk, strengthen intentions to take precautions, and promote behavior change regarding various health risks, such as smoking (Harris et al., 2007), excessive caffeine consumption (Reed & Aspinwall, 1998; Sherman et al., 2000), alcohol consumption (Harris & Napper, 2005), unsafe sex (Sherman et al., 2000), insufficient fruit and vegetable consumption (Epton & Harris, 2008), and Type 2 diabetes (van Koningsbruggen & Das, in press). These effects have remained stable over a period of 1 month (Harris & Napper, 2005).

Most studies focus on explicit persuasive outcomes (e.g., attitudes, intentions); less is known about the cognitive processes instigated by self-affirmation in response to threatening health information. To illustrate, Sherman et al. (2000) had coffee drinkers and noncoffee drinkers respond to information linking caffeine consumption to severe health problems. Typically, coffee drinkers respond defensively, with the result that they accept such information to a lesser extent than noncoffee drinkers (e.g., Kunda, 1987). Self-affirmation reduced defensive responses among high-relevance participants: Self-affirmed coffee drinkers were more likely to accept that caffeine consumption was linked to health problems and reported greater intentions to take precautions than coffee drinkers who had not self-affirmed. Using a similar paradigm, self-affirmation made high-relevance participants read risk-confirming information earlier than risk-neutral or risk-disconfirming information (Reed & Aspinwall, 1998). However, contrasting previous findings, self-affirmed high-relevance participants reported lower intentions to reduce their caffeine consumption than their nonaffirmed counterparts.

One recent study reported more direct evidence regarding the cognitive processes instigated by self-affirmation in response to threatening health information by drawing on dual process models of persuasion (e.g., Petty & Cacioppo, 1986). When self-affirmation reduces defensive processing, people should differentiate between weak and strong arguments in a message and recognize the merits and demerits of these arguments (Petty & Cacioppo, 1986). Consistent with this, self-affirmation made people more sensitive to the quality of the arguments in a health message (van Koningsbruggen & Das, 2009). These findings suggest that self-affirmation gives people the resources to deal with threatening health information, thus enabling them to process this information in a more open-minded fashion.

Most pressing is the absence of research on self-affirmation in the area of less controlled, automatic cognitive responses to threatening health information. Because many self-regulatory processes occur automatically and outside awareness (e.g., Bargh & Chartrand, 1999), it is important to test whether the proposed resource function of self-affirmation may also extend to this area, and whether it increases responsiveness to threatening health information at the implicit level. Theoretically, self-affirmation is presumed to diminish self-integrity concerns and stimulate integration

of the threat into the self-system (Sherman & Cohen, 2006). Thus, self-affirmation should make it easier to encode threatening aspects of health information because these aspects no longer pose a threat to people's self-integrity. Because information that is better encoded should be easier to retrieve from memory (i.e., more accessible; Carlston & Smith, 1996; Higgins, 1996), we measured the accessibility of threat-related cognitions with an unobtrusive lexical decision task in the present study. When self-affirmation promotes responsiveness to health information at the implicit level, it should increase the accessibility of threat-related cognitions.

In the present study, we tested this assumption employing the "caffeine consumption paradigm" that has been frequently used in previous research on defensive processing of health information (e.g., see Kunda, 1987; Liberman & Chaiken, 1992; Sherman et al., 2000). Moreover, because previous studies that used this paradigm provided contradictory evidence regarding the impact of self-affirmation on intentions (Reed & Aspinwall, 1998; Sherman et al., 2000), we explicitly assessed intentions to take precautions and perceptions of health message quality. In doing so, we meet the call for delineating how cognitive processes influence health self-regulatory efforts (Williams, Wasserman, & Lotto, 2003).

The Present Study

Following previous research (e.g., Sherman et al., 2000), we recruited high-relevance (coffee drinkers) and low-relevance (noncoffee drinkers) participants and presented them with health information linking caffeine consumption to severe health problems. Prior to reading this information, we manipulated self-affirmation by allowing participants to affirm a value that was either personally important (self-affirmation) or unimportant (no affirmation) to them. All participants then completed an unobtrusive lexical decision task to assess the accessibility of threat-related cognitions. Subsequently, they reported their perceptions of message quality and their intentions to take precautions. We predicted that among high-relevance participants (coffee drinkers), relative to low-relevance participants (noncoffee drinkers), self-affirmation would increase accessibility of threat-related cognitions, perceptions of message quality, and intentions to reduce caffeine consumption.

Method

Design and Participants

The hypotheses were tested in a 2 (relevance: coffee drinker vs. noncoffee drinker) × 2 (self-affirmation status: nonaffirmed vs. self-affirmed) between-participants factorial design. A total of 84 university students participated in the experiment, of which 20 were men and 64 were women with a mean age of 23.62 years ($SD = 3.66$). As compensation for their participation, participants could take part in a lottery in which they could win gift vouchers (€25, approximately \$35).

Procedure and Materials

Participants were informed that they would participate in several separate computerized studies: one about consumption patterns concerning several beverages, one about values, one assessing student opinions of scientific articles, and the other about word recognition.

Relevance. After the introduction, participants were presented with several questions designed to identify their coffee drinking behavior (e.g., coffee drinker = yes/no; if yes, whether they consumed regular or decaffeinated coffee, etc.). To bolster our cover story, we asked similar questions about several other beverages. The sample consisted of 47 coffee drinkers and 37 noncoffee drinkers. Participants then continued with the "value study" that contained the manipulation of self-affirmation.

Self-affirmation manipulation. The manipulation of self-affirmation was based on a well-established procedure (e.g., Harris & Napper, 2005). Participants were randomly assigned to either the self-affirmed status condition ($n = 40$) or the nonaffirmed status condition ($n = 44$). They were first presented with the six values of the Allport–Vernon–Lindzey Study of Values (Allport, Vernon, & Lindzey, 1960). The values listed were science, business, art, social, politics, and religion. In the self-affirmed status condition, participants were asked to choose their most important value and to write about why it was important to them and to describe a specific occasion when it had been particularly important. In the nonaffirmed status condition, they were asked to choose their least important value and to write about why the value might be important to the average student.

Health message. After the manipulation of self-affirmation, participants read a fictitious article titled "The Effects of Caffeine on Health," which was supposedly published in the *Journal of Medicine* (based on Block & Williams, 2002). The first part of the article presented several examples of products containing caffeine. It was stressed that the greatest part of people's caffeine intake could be related to their coffee consumption. The second part described possible health problems related to caffeine consumption (e.g., insomnia, restlessness, high blood pressure, cardiovascular disease). Moreover, it was stated that research suggested that people who consume high doses of caffeine, for example, coffee drinkers, are more likely to suffer a heart attack or a stroke. After participants read the health message, they were asked to participate in the study on word recognition (i.e., the lexical decision task). Participants then completed the perceived message quality and intention measures, were probed for suspicions about the purpose of the study, and then were extensively debriefed. None of the participants guessed any aspect of the true purpose of the study, and none of them reported a suspicion that the studies were related.

Dependent Variables

Accessibility of threat-related cognitions. Participants were informed that a sequence of words would appear individually in the center of the screen. They were asked to decide as quickly and accurately as possible whether the presented word was an existing Dutch word or a nonword by pushing either the *A* (for existing words) or the *L* (for nonwords) key on the keyboard. Each trial first presented a fixation point in the center of the screen that was replaced by the stimulus after 1 s. The next trial was initiated 500 ms after participants pressed one of the keys. In total, participants responded to 40 words, consisting of 20 existing Dutch words and 20 nonwords. In reality, we were interested in reaction times (RT) to existing target words between conditions. Among the existing words, 5 were threat-related target words from the health message participants had just read, pertaining to the adverse health problems of caffeine consumption (blood pressure, heart attack, heart

disease, restless, sick; all single words in Dutch), the remaining items reflected control words (e.g., apartment, mechanic, motorway). The threat-related words were selected on the basis of a pilot study in which participants ($n = 41$, not taking part in the actual experiment) rated on a 7-point scale (1 = *not at all*, 7 = *very much*) to what extent these and 10 other words pertained directly to health problems related to caffeine consumption. The 5 words selected were the ones rated highest on this measure ($M = 5.15$, $SD = 1.99$). All words were matched on word length across word type categories, randomly presented, and preceded by four practice trials.[1]

Perceived message quality. To assess participants' perceptions of message quality, they responded to the statement, "The evidence linking caffeine consumption and health problems is reliable" (1 = *strongly disagree*, 7 = *strongly agree*). Higher scores indicate higher perceptions of message quality.

Intentions to reduce caffeine consumption. Participants rated two statements that examined their intentions to reduce their caffeine consumption on a 7-point scale (based on Block & Williams, 2002; "I intend to cut down my caffeine consumption" and "I am convinced that I will reduce the amount of caffeine I consume"; 1 = *strongly disagree*, 7 = *strongly agree*; $r = .79$, $p < .001$). Higher scores indicate greater intentions to reduce caffeine consumption.

Results

Accessibility of Threat-Related Cognitions

The computer recorded both the response (i.e., word or nonword) and RT (in ms) per presented item for each participant. Incorrect responses were excluded from the analysis (2.86% across the experimental trials, which were evenly distributed across conditions). The harmonic means of RT (Ratcliff, 1993) for threat-related words and control words were subjected to a 2 (relevance: coffee drinker vs. noncoffee drinker) × 2 (self-affirmation status: nonaffirmed vs. self-affirmed) × 2 (word type: threat-related vs. control) mixed design analysis of variance (ANOVA) with repeated measures on the third factor. The analysis revealed a main effect of word type, $F(1, 80) = 17.49$, $p < .001$, $\eta_p^2 = .18$. Participants responded faster to threat-related words ($M = 775$ ms) than to control words ($M = 837$ ms). In addition, the analysis revealed the predicted interaction effect between relevance, self-affirmation, and word type, $F(1, 80) = 7.79$, $p = .007$, $\eta_p^2 = .09$. Simple effects analysis indicated that among the coffee drinkers, self-affirmed participants responded faster to threat-related words than nonaffirmed participants, $F(1, 80) = 4.69$, $p = .033$, $\eta_p^2 = .06$. No effect of self-affirmation was found among the noncoffee drinkers, $F < 1$. Also, no self-affirmation effects occurred on RT for control words: for coffee drinkers, $F < 1$; for noncoffee drinkers, $F(1, 80) = 1.19$, *ns*. Relevant means are displayed in Table 1.

---

[1] It should be noted that the nonthreat words were not equivalent in frequency of use to those of the threat words. Because our key comparisons are within-word between-conditions, this issue does not affect the results presented in this article.

Table 1

Mean Harmonic Mean Reaction Time (ms) as a Function of Relevance, Self-Affirmation Status Condition, and Word Type

Word type	Low relevance (noncoffee drinkers)		High relevance (coffee drinkers)	
	Nonaffirmed	Self-affirmed	Nonaffirmed	Self-affirmed
Threat related	753$_a$	761$_{a,b}$	841$_b$	744$_a$
Control	849$_a$	786$_a$	857$_a$	854$_a$

Note. Within rows, means with different subscripts differ significantly ($p < .05$).

Perceived Message Quality

A 2 (relevance) × 2 (self-affirmation status) ANOVA on perceived message quality revealed a significant interaction between relevance and self-affirmation, $F(1, 80) = 7.04$, $p = .01$, $\eta_p^2 = .08$. Simple effects analysis revealed that among the coffee drinkers, self-affirmed participants reported higher perceptions of message quality ($M = 4.56$, $SD = 0.64$) than nonaffirmed participants ($M = 3.90$, $SD = 0.91$), $F(1, 80) = 5.60$, $p = .02$, $\eta_p^2 = .07$. No effect of self-affirmation was found among the noncoffee drinkers ($M_{\text{self-affirmed}} = 3.69$, $SD_{\text{self-affirmed}} = 0.95$; $M_{\text{nonaffirmed}} = 4.17$, $SD_{\text{nonaffirmed}} = 1.20$), $F(1, 80) = 2.15$, *ns*.

Intentions to Reduce Caffeine Consumption

A 2 (relevance) × 2 (self-affirmation status) ANOVA on intentions to reduce caffeine consumption revealed a significant interaction between relevance and self-affirmation, $F(1, 80) = 6.03$, $p = .016$, $\eta_p^2 = .07$. Simple effects analysis revealed that among the coffee drinkers, self-affirmed participants reported greater intentions to reduce their caffeine consumption ($M = 3.02$, $SD = 1.73$) than nonaffirmed participants ($M = 2.10$, $SD = 0.97$), $F(1, 80) = 5.42$, $p = .022$, $\eta p^2 = .06$. No effect of self-affirmation was found among the noncoffee drinkers ($M_{\text{self-affirmed}} = 1.85$, $SD_{\text{self-affirmed}} = 0.55$; $M_{\text{nonaffirmed}} = 2.42$, $SD_{\text{nonaffirmed}} = 1.38$), $F(1, 80) = 1.54$, *ns*.

Relationship Between Implicit and Explicit Measures

Following Baron and Kenny (1986), we tested whether the difference in RT to threat-related words mediated the impact of self-affirmation on perceived message quality or intentions among high-relevance participants (controlling for overall speed of responding; cf. Rhodes, Roskos-Ewoldsen, Edison, & Bradford, 2008). Self-affirmation predicted RT to threat-related words, $B = -48.09$, $t(44) = -2.66$, $p = .011$, and also perceived message quality, $B = 0.34$, $t(44) = 2.97$, $p = .005$, and intentions, $B = 0.47$, $t(44) = 2.14$, $p = .038$. However, RT to threat-related words did not significantly predict perceived message quality ($p = .38$) or intentions ($p = .19$). Thus, no further analyses were conducted because not all conditions for mediation were met.[2]

Finally, we performed partial correlation analyses between RT to threat-related words and the perceived message quality and intention measures among high-relevance participants. As can be seen in Table 2, RT to threat-related words was negatively related to perceived message quality among self-affirmed coffee drinkers. Thus, participants with faster response latencies to threat-related words reported higher perceptions of quality of the message linking caffeine consumption and health problems. This finding, in conjunction with the results on the lexical decision task, suggests that increased accessibility of threat-related cognitions is related to less defensive responding to threatening health information.

Discussion

The present study provided a first published test of the impact of self-affirmation on people's implicit responsiveness to threatening health information. High-relevance (coffee drinkers) and low-relevance (noncoffee drinkers) participants were either self-affirmed or not prior to reading information linking caffeine consumption to severe health problems. The results revealed a highly consistent pattern across implicit and explicit measures. As predicted, self-affirmation increased high-relevance participants' accessibility of threat-related cognitions, positively affected perceptions of message quality, and increased intentions to take precautions.

Consistent with previous research, we showed that self-affirmation reduces defensiveness toward threatening health information, and promotes adaptive behavioral intentions among people for whom the information is highly relevant (e.g., Harris & Napper, 2005; Sherman et al., 2000). In addition, the present findings are the first to demonstrate that the beneficial effects of self-affirmation are not limited to changes in explicit variables related to persuasive outcomes but extend to the realm of less controlled, automatic cognitive responses to threatening health information. Herewith, this study provides more insight into the cognitive processes instigated by self-affirmation. At this point, little is known about implicit processes related to threatening information (Glaser & Banaji, 1999; Roskos-Ewoldsen, Yu, & Rhodes, 2004). Demonstrating a link between the accessibility of threatening information and perceptions of message quality when people are self-affirmed sheds further light on the dynamic relationship between threat and message processing.

Although there is some evidence that self-affirmation increases extensive, careful processing of threatening health information on a conscious, explicit level (van Koningsbruggen & Das, 2009), the present study suggests that when self-affirmation diminishes concerns about self-integrity, threatening health information becomes easier to retrieve from memory. This also fits nicely with and supplements previous research showing that self-affirmation makes people read risk-confirming information earlier (Reed & Aspinwall, 1998). Because information that is more accessible will be more likely encoded and recalled (Carlston & Smith, 1996; Higgins, 1996), the present results are consistent with the idea that self-affirmation stimulates integration of the threat into the self-system (Sherman & Cohen, 2006). Consequently, because the

---

[2] We also tested whether perceived message quality mediated the effect of self-affirmation on intentions among high-relevance participants. Self-affirmation predicted perceived message quality ($B = 0.34$, $t = 2.97$, $p = .005$), but message quality did not predict intentions ($p = .20$). Consistent with the notion of compatibility of levels of measurement (Ajzen, 1988), general beliefs about message quality did not predict specific behavioral intentions to reduce caffeine consumption.

Table 2

Partial Correlations Between Reaction Times to Threat-Related Words and Perceived Message Quality and Intentions for High-Relevance Participants

Condition/variable	r	p
	RT–threat-related words[a]	
Nonaffirmed status		
Perceived message quality	.15	.57
Intentions	−.13	.60
Self-affirmed status		
Perceived message quality	−.40	.046
Intentions	−.01	.97

[a] Controlling for overall speed of responding.

threatening aspects of the health information no longer pose a threat to people's self-integrity, it should reduce the need, for example, to derogate the health information. Consistent with this, we found that the more accessible threat-related cognitions were for self-affirmed coffee drinkers, the more enhanced were their perceptions of quality of the message linking caffeine consumption and health problems.

Although the accessibility of threat-related cognitions did not mediate the effects of self-affirmation on perceptions of message quality and intentions among high-relevance participants, we did find a significant relationship between the accessibility of threat-related cognitions and perceived message quality among self-affirmed, high-relevance participants. This finding suggests that implicit responsiveness to threatening health information is related to decreased defensive responding at the explicit level. In addition, the finding is in line with previous research showing that global cognitions may be related to specific cognitions, particularly when they are highly accessible (Fazio, 1995). In the present study, it is likely that we could not establish mediation because we measured construct accessibility rather than attitude accessibility (Fazio, 1995; see also, e.g., Houston & Fazio, 1989), the former being a less likely candidate in predicting specific attitudes. Furthermore, some features of the present study seem to have limited our ability to detect these effects; we assessed perceived message quality with only one item, and our data may have limited statistical power to detect mediation (Fritz & MacKinnon, 2007). Future studies should take these limitations into account when further examining the relation between accessibility of threat-related cognitions and variables involved in the behavior change process. This would also provide a better test of whether the change in implicit responsiveness observed in the present study is indicative of a mechanism underlying the obtained self-affirmation effects on the explicit measures, or indicative of a consequence of more open-minded information processing fostered by self-affirmation.

Related to this issue is the possibility that self-affirmed participants may have spent more time reading the health information and, as a result of greater priming, the threatening aspects of the health information became easier to retrieve from memory. Although the present study did not record how long participants spent reading the health information, recent research suggests that self-affirmation can increase extensive, careful information processing (van Koningsbruggen & Das, 2009). Self-affirmed participants' careful scrutiny of the message should influence judgments of

message quality, suggesting that more systematic information processing at the explicit level fostered by self-affirmation goes hand-in-hand with greater responsiveness at the implicit level.

The present findings also add to the growing evidence that self-affirmation can promote intentions to take precautions. Although we did not assess whether participants actually translated their intentions into actions, recent findings suggest that self-affirmation can promote behavior change. One study that examined screening behavior demonstrated that self-affirmation increased risk test taking among people for whom the health information was highly relevant (van Koningsbruggen & Das, in press), and another study showed that self-affirmation increased fruit and vegetable consumption (Epton & Harris, 2008). However, future studies are needed to extend these findings to other behaviors. In addition, although the accessibility of the threatening information was not related to behavioral intentions, this is consistent with a broad array of research demonstrating that construct and attitude accessibility play a greater role in the prediction of spontaneous behaviors rather than deliberative behavior or behavioral intentions (Olson & Fazio, 2009). Testing the impact of self-affirmation on spontaneous health behaviors constitutes an important agenda for future studies.

In conclusion, this study provided a first test of the impact of self-affirmation on implicit responsiveness to threatening health information. Although several studies have demonstrated that self-affirmation can reduce defensive responses at the explicit level, insight into the cognitive processes instigated by self-affirmation tended to be limited (Sherman & Cohen, 2006). The present findings represent a step forward in this area: Self-affirmation promotes accessibility of threatening health information.

References

Ajzen, I. (1988). *Attitudes, personality, and behavior*. Chicago: Dorsey.

Allport, G. W., Vernon, P. E., & Lindzey, G. (1960). *Study of values* (3rd ed.). Boston: Houghton Mifflin.

Arpan, L., Rhodes, N., & Roskos-Ewoldsen, D. R. (2007). Attitude accessibility: Theory, methods and future directions. In D. R. Roskos-Ewoldsen & J. L. Monahan (Eds.), *Communication and social cognition: Theories and methods*. Mahwah, NJ: Erlbaum.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54,* 462–479.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Block, L. G., & Williams, P. (2002). Undoing the effects of seizing and freezing: Decreasing defensive processing of personally relevant messages. *Journal of Applied Social Psychology, 32,* 803–833.

Carlston, D. E., & Smith, E. R. (1996). Principles of mental representation. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 184–210). New York: Guilford Press.

De Wit, J. B. F., Das, E., & Vet, R. (2008). What works best: Objective statistics or a personal testimonial? An assessment of the persuasive effects of different types of message evidence on risk perception. *Health Psychology, 27,* 110–115.

Epton, T., & Harris, P. R. (2008). Self-affirmation promotes health behavior change. *Health Psychology, 27,* 746–752.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Hillsdale, NJ: Erlbaum.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18,* 233–239.

Glaser, J., & Banaji, M. R. (1999). When fair is foul and foul is fair: Reverse priming in automatic evaluation. *Journal of Personality and Social Psychology, 77,* 669–687.

Harris, P. R., Mayle, K., Mabbott, L., & Napper, L. (2007). Self-affirmation reduces smokers' defensiveness to graphic on-pack cigarette warning labels. *Health Psychology, 26,* 437–446.

Harris, P. R., & Napper, L. (2005). Self-affirmation and the biased processing of threatening health-risk information. *Personality and Social Psychology Bulletin, 31,* 1250–1263.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Houston, D. A., & Fazio, R. H. (1989). Biased processing as a function of attitude accessibility: Making objective judgments subjectively. *Social Cognition, 7,* 51–66.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53,* 636–647.

Liberman, A., & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin, 18,* 669–679.

Olson, M. A., & Fazio, R. H. (2009). Implicit and explicit measures of attitudes: The perspective of the mode model. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes: Insights from the new implicit measures* (pp. 19–63). New York: Psychology Press.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). San Diego, CA: Academic Press.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–532.

Reed, M. B., & Aspinwall, L. G. (1998). Self-affirmation reduces biased processing of health-risk information. *Motivation and Emotion, 22,* 99–132.

Rhodes, N., Roskos-Ewoldsen, D. R., Edison, A., & Bradford, M. B. (2008). Attitude and norm accessibility affect processing of anti-smoking messages. *Health Psychology, 27,* 224–232.

Roskos-Ewoldsen, D. R., Yu, H. J., & Rhodes, N. (2004). Fear appeal messages affect accessibility of attitudes toward the threat and adaptive behaviors. *Communication Monographs, 71,* 49–69.

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., 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.

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). New York: Academic Press.

van Koningsbruggen, G. M., & Das, E. (2009). *When self-affirmation decreases extensive processing of threatening information: The moderating role of self-threat level.* Manuscript in preparation.

van Koningsbruggen, G. M., & Das, E. (in press). Don't derogate this message! Self-affirmation promotes online Type 2 diabetes risk test taking. *Psychology and Health.*

Williams, P. G., Wasserman, M. S., & Lotto, A. J. (2003). Individual differences in self-assessed health: An information-processing investigation of health and illness cognition. *Health Psychology, 22,* 3–11.