Health promotion campaigns often try to motivate people to change their unhealthy behaviors by communicating information about the risks associated with the behavior in question. For instance, health messages may inform people with unhealthy diets about their increased risk for developing diabetes, which can result in serious health problems. By increasing perceptions of threat, these campaigns attempt to motivate people to change their unhealthy diet. Unfortunately, reading about health risks does not necessarily motivate people who are at-risk for the targeted outcome to adapt healthy behaviors or stop risky behaviors. Research shows that at-risk people often engage in defensive processing of threatening health information, with the result that people targeted by the message are often the least persuaded (e.g., Croyle et al., 1997; Kunda, 1987; Liberman & Chaiken, 1992). Defensive processing of threatening health information thus presents an obstacle for health campaigns, and efforts are needed directed at making people more accepting of the truth, even if it is as harsh as the reality of a serious and relevant health risk. Several researchers adopted self-affirmation theory (Steele, 1988) as a framework to make people less defensive and more accepting of threatening health information (e.g., Harris & Napper, 2005; Reed & Aspinwall, 1998; Sherman et al., 2000).

From the perspective of self-affirmation theory (Steele, 1988), people who are at-risk for the targeted outcome respond defensively, because they are highly motivated to protect and maintain a global sense of self-integrity. By derogating, minimizing, or avoiding threatening health information people are able to restore and maintain their self-integrity. A crucial tenet of self-affirmation theory, however, is that people are only concerned with their

* This chapter is based on Van Koningsbruggen & Das (2008a).
global sense of self-integrity. Accordingly, Steele’s (1988) self-affirmation theory suggests that if people can restore their global sense of self-integrity by drawing upon alternative self-resources unrelated to the specific threat (i.e., “self-affirm”), they should be less likely to respond defensively and more likely to accept threatening health information.

Consistent with this prediction, several studies show that theory-based manipulations of self-affirmation reduce defensive responses to threatening health information among people who are at-risk for the targeted outcome (e.g., Armitage et al., 2008; Harris & Napper, 2005; Reed & Aspinwall, 1998; Sherman et al., 2000). This research demonstrated that self-affirmation positively affects measures such as message acceptance, personal risk perceptions, self-efficacy, response-efficacy, and intentions to take precautions. Moreover, self-affirmation has been shown to promote healthy behaviors (e.g., Epton & Harris, 2008; Jessop et al., in press; Sherman et al., 2000; Van Koningsbruggen & Das, in press). For instance, Sherman and colleagues (2000, Study 2) found that, among at-risk participants who had been watching a threatening AIDS educational video, self-affirmed participants purchased more condoms and took more AIDS educational brochures than did those who had not self-affirmed. Self-affirmation thus promoted AIDS-preventive behaviors among a target audience. Likewise, self-affirmation has been shown to promote screening behavior (Van Koningsbruggen & Das, in press) and health-promoting behavior (the consumption of fruit and vegetables; Epton & Harris, 2008).

The above reviewed evidence suggests that self-affirmation always promotes adaptive responses to threatening health information. However, within self-affirmation research there has so far been little attention to possible boundary conditions of the effects of self-affirmation (Sherman & Cohen, 2006). Nevertheless, there may be limits to the positive effects of self-affirmation. Recent research suggests that self-affirmation may backfire on the
acceptance of threatening health information when self-threat levels become severely high (Van Koningsbruggen & Das, 2008b).

The Moderating Role of Self-Threat Level

Research on self-affirmation in the health domain generally assumes that self-affirmation reduces defensive responses to threatening health information because self-affirmation increases extensive, careful processing of the information (e.g., Harris & Napper, 2005; Sherman et al., 2000). As a consequence, people are more likely to draw the most accurate conclusion from the information, including less-desired ones such as that they are at-risk for a harmful disease. Consistent with this reasoning, recent research demonstrated that self-affirmation made people more sensitive to the quality of the arguments in a threatening health message (Van Koningsbruggen & Das, 2008b), indicating that self-affirmation indeed increases extensive, careful information processing (Petty & Wegener, 1999). Importantly, however, is that this beneficial effect of self-affirmation seemed to be limited to conditions in which people experienced a moderate threat to the self. Under high self-threat conditions, van Koningsbruggen and Das (2008b) showed that self-affirmation decreased extensive, careful processing of threatening health information. These findings suggest that when alternative self-resources made salient by the self-affirmation manipulation are not of sufficient magnitude to cope with a threat, self-affirmation reinforces concerns over self-integrity rather than diminishes them, and subsequently backfires on information processing.

The above implies that most self-affirmation studies in the health domain may have (inadvertently) involved relatively moderate levels of self-threat. In the health domain, self-threat level is conceptualized as people’s level of vulnerability to a health risk. Previous self-affirmation research classified participants’ vulnerability on their actual behavior (e.g., Harris & Napper, 2005; Reed & Aspinwall, 1998; Sherman et al., 2000). For instance, Sherman and
colleagues (2000, Study 1) compared coffee drinkers’ and non-coffee drinkers’ responses to threatening health information linking coffee consumption and severe health problems. Because the health risk information was only relevant to coffee drinkers, it was logically assumed that these participants experienced a self-threat whereas non-coffee drinkers did not. In contrast with this approach, Van Koningsbruggen and Das (2008b, Study 3) manipulated vulnerability by providing participants with false feedback regarding their vulnerability level. In this study, participants completed a fictitious health test and were either told that they were not very vulnerable to a health risk (moderate self-threat condition), or that they were highly vulnerable to a health risk (high self-threat condition). An initial study that compared both methods within one design indeed confirmed that a manipulation of vulnerability brings forth a higher threat to the self than when vulnerability is based on actual behavior (Van Koningsbruggen & Das, 2008b, Study 1). More specifically, classifying participants’ vulnerability on actual behavior (e.g., non-coffee drinkers vs. coffee drinkers) induced low and moderate self-threat levels, whereas a manipulation of vulnerability induced moderate and high self-threat levels.

Keeping in mind that self-affirmation has been shown to backfire on health information processing under high self-threat conditions (Van Koningsbruggen & Das, 2008b), there may be boundary conditions to the positive effects of self-affirmation on responses to threatening health information, under which self-affirmation is ineffective or even backfires. Less careful consideration of threatening health information may render people targeted by the outcome less likely to evaluate arguments supporting the health information on their merits and demerits. Consequently, this may prevent them from drawing the most accurate conclusion that follows from the health information, potentially decreasing the likelihood that people adaptively respond to this personally important information. Whereas previous research showed that self-affirmation decreases extensive, careful
processing of threatening health information under high self-threat conditions (Van Koningsbruggen & Das, 2008b), the present study investigated whether self-affirmation would similarly backfire on persuasive outcomes following a severe threat to the self.

**Study 5.1**

In this study, we experimentally manipulated self-threat level by providing participants with false feedback regarding their vulnerability to a health risk (RSI). Previous research has shown that a manipulation of vulnerability induces moderate and high self-threat conditions (Van Koningsbruggen & Das, 2008b). Prior to reading the threatening health information, we manipulated self-affirmation. We then measured participants’ attitudes toward the health message, intentions to take precautions, and behavior. Although previous self-affirmation studies in the health domain have shown positive effects of self-affirmation on such persuasive outcomes (e.g., Harris & Napper, 2005; Jessop et al., in press; Sherman et al., 2000), we predicted that this would be limited to conditions in which people experience a moderate threat to the self. Therefore, it was hypothesized that self-affirmation would increase persuasion under moderate self-threat conditions, and would decrease persuasion under high self-threat conditions.

**Method**

**Design and Participants**

The hypotheses were tested in a 2 (self-threat level: moderate vs. high) x 2 (self-affirmation status: non-affirmed vs. self-affirmed) between-participants factorial design. Participants were 72 university students (20 males, 52 females; $M_{age} = 22.76$ years, $SD_{age} = 3.11$ years). They were randomly assigned to experimental conditions. As compensation, participants could win cash prizes.
Procedure and Materials

Participants were brought to the lab to participate in several unrelated studies on diverse psychological topics. This study was presented as consisting of two unrelated studies: one about health, and the other about values. The first part of the experiment was aimed at identifying participants’ most and least important values to manipulate self-affirmation. Participants were presented with the six values of the AVL study of Values (Allport et al., 1960) which they had to rank according to their personal importance. The values listed were: science, business, art, social, politics, and religion. After participants ranked the values, they proceeded with a fictitious health test that contained the manipulation of self-threat level.

Self-threat level. Participants were led to believe that this test, consisting of a bogus personality test, measured their risk of developing RSI-related health problems in the future (cf. Van Koningsbruggen & Das, 2008b). Participants completed the test, were led to believe that the computer analyzed their responses and were instructed to go to the next screen to see the result of this analysis. On this screen, participants received false feedback regarding their vulnerability to develop health problems. In the moderate self-threat condition, they saw a square, filled with a bright green color, that displayed: “Your risk of developing RSI-related health problems is quite low.” In the high self-threat condition, participants saw a square, filled with a bright red color that displayed: “Your risk of developing RSI-related health problems is quite high.” Participants then completed the self-affirmation manipulation.

Self-affirmation. The manipulation of self-affirmation was based on a well-established procedure (cf. Koole et al., 1999; Sherman et al., 2000; Tesser & Cornell, 1991), whereby participants receive a 10-item AVL-subscale (Allport et al., 1960) either matching their previously indicated most important value (self-affirmed status condition) or least important value (non-affirmed status condition). Each of the ten items consisted of a statement with two possible answers. One of the answers reflected participant’s most important value (self-
How Self-Affirmation Affects Health Persuasion

affirmed status condition) or least important value (non-affirmed status condition). The other answer was filler.

*Health message.* Following the self-affirmation manipulation, participants read a health message supposedly published in “The Journal of Health.” The message described the negative health consequences of RSI and prevention training as a means to reduce these negative health consequences (this message was similar to the strong version of the health message used in Van Koningsbruggen & Das, 2008b, Study 2).

After reading the health message, participants completed the dependent measures. They were then probed for suspicion 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 about the vulnerability feedback or that the two studies were related.

*Dependent Variables*

*Manipulation checks.* To confirm the success of the self-threat level manipulation, participants rated four items that measured their perceived vulnerability to RSI-related health problems on a 7-point scale (e.g., “The chance that I will develop RSI-related health problems is high” and “Due to my sensitivity to RSI I am more prone to RSI-related health problems”; 1 = strongly disagree, 7 = strongly agree; α = .74). Higher scores indicate higher perceptions of vulnerability (i.e., self-threat).

The validity of the self-affirmation manipulation was assessed by counting the number of times participants endorsed the manipulated value. A score of 1 was given when participants chose the manipulated value (thus affirming their least or most important value). A score of 0 was given when participants chose the filler answer. Thus, total scores on the affirmation task ranged from 0 to 10, with higher scores indicating higher affirmation. We
expected that self-affirmed participants would endorse the manipulated value more often than non-affirmed participants.

**Attitudes.** Participants rated one item that measured their attitude toward the health message on a 7-point semantic differential scale (how useless-useful participants thought this message was), with higher scores indicating a more favorable attitude.

**Intentions.** Participants rated two items that measured their intentions to take precautions on a 7-point scale (e.g., “How likely is it that you are going to do stretching exercises to reduce your risk at RSI when you do a lot of computer work?”; 1 = very unlikely, 7 = very likely; \( r = .80, p < .001 \)). Higher scores indicate greater intentions to take precautions.

**Behavior.** Participants could request a leaflet with more information about RSI and the prevention of RSI; this item had a “yes” (coded as 1) or “no” (codes as 0) option.

**Results**

**Manipulation Checks**

**Self-threat level.** A one-way ANOVA confirmed that participants in the high self-threat condition felt significantly more vulnerable to the health risk \( (M = 3.87, SD = 1.41) \) than participants in the moderate self-threat condition \( (M = 2.41, SD = 1.18) \), \( F(1, 70) = 22.86, p < .001, \eta_p^2 = .25 \).\(^1\)

**Self-affirmation.** A one-way ANOVA confirmed that self-affirmed participants endorsed the manipulated value more often \( (M = 8.00, SD = 1.73) \) than non-affirmed participants \( (M = 3.62, SD = 1.55) \), \( F(1, 70) = 127.90, p < .001, \eta_p^2 = .65 \).
Attitudes

A 2 (self-threat level) x 2 (self-affirmation status) ANOVA on attitudes toward the health message revealed a significant main effect of self-threat level, $F(1, 68) = 7.49$, $p = .008$, $\eta_p^2 = .10$. This main effect was qualified by the expected two-way interaction between self-threat level and self-affirmation, $F(1, 68) = 20.23$, $p < .001$, $\eta_p^2 = .23$. Simple effects analysis revealed that in the moderate self-threat condition, self-affirmed participants reported more positive attitudes toward the health message ($M = 5.30$, $SD = 1.08$) than non-affirmed participants ($M = 3.83$, $SD = 1.04$), $F(1, 68) = 19.22$, $p < .001$, $\eta_p^2 = .22$. In contrast, in the high self-threat condition, self-affirmed participants reported less positive attitudes toward the health message ($M = 3.53$, $SD = 1.25$) than non-affirmed participants ($M = 4.26$, $SD = 0.73$), $F(1, 68) = 4.21$, $p = .044$, $\eta_p^2 = .06$.

Intentions

A 2 (self-threat level) x 2 (self-affirmation status) ANOVA on intentions to take precautions revealed a significant two-way interaction effect between self-threat level and self-affirmation, $F(1, 68) = 15.53$, $p < .001$, $\eta_p^2 = .19$. The nature of this interaction is displayed in Figure 5.1. Simple effects analysis revealed that in the moderate self-threat condition, self-affirmed participants reported higher intentions to take precautions ($M = 4.15$, $SD = 1.59$) than non-affirmed participants ($M = 2.81$, $SD = 1.33$), $F(1, 68) = 8.11$, $p = .006$, $\eta_p^2 = .11$. In contrast, in the high self-threat condition, self-affirmed participants reported lower intentions to take precautions ($M = 2.87$, $SD = 1.56$) than non-affirmed participants ($M = 4.24$, $SD = 1.32$), $F(1, 68) = 7.46$, $p = .008$, $\eta_p^2 = .10$. 
Under moderate self-threat conditions, five of the self-affirmed participants \((n = 20)\) requested a leaflet, whereas zero of the non-affirmed participants \((n = 18)\) did. Under high self-threat conditions, one of the self-affirmed participants \((n = 15)\) requested a leaflet, whereas seven of the non-affirmed participants \((n = 19)\) did. Because of empty and small cells, we tested the impact of self-affirmation on behavior under different self-threat levels using cross-tabulations instead of logistic regression analysis (which is unstable when there are empty or small cells; Stevens, 1992). Under moderate self-threat conditions, significantly more participants in the self-affirmed status condition requested a leaflet than did those in the non-affirmed status condition, Likelihood Ratio = 7.10, \(p = .008\), Cramer’s V = .37. Under high self-threat conditions, significantly less participants in the self-affirmed status condition requested a leaflet than did those in the non-affirmed status condition, Likelihood Ratio = 4.74, \(p = .029\), Cramer’s V = .35.
Discussion

The present study provided a first test of the impact of self-affirmation on people’s responses to threatening health information following a severe threat to the self. Whereas previous research demonstrated that self-affirmation decreases extensive, careful processing of threatening health information under high self-threat conditions (Van Koningsbruggen & Das, 2008b), the present findings suggest that this decreased processing may backfire on persuasive outcomes. Under moderate self-threat conditions, self-affirmed participants reported more positive attitudes toward the health message, expressed higher intentions to take precautions, and requested more leaflets than non-affirmed participants, consistent with the predicted increase in persuasion. In contrast, under high self-threat conditions, self-affirmed participants reported less positive attitudes, expressed lower intentions, and requested less leaflets than non-affirmed participants, consistent with the predicted decrease in persuasion.

Consistent with previous research, we show that self-affirmation can reduce defensive responses to threatening health information and promote adaptive behavioral intentions and healthy behaviors among people for whom the information is highly relevant (e.g., Harris & Napper, 2005; Jessop et al., in press; Sherman et al., 2000). However, the present findings are the first to demonstrate that these beneficial effects of self-affirmation are limited to conditions in which people experience a moderate threat to the self. In these conditions, self-affirmation stimulates integration of the threat into the self-system, enabling people at-risk for the targeted outcome to focus on the informational value of the threatening health information instead of its implications for self-integrity (cf. Sherman & Cohen, 2006). When severely threatened, however, self-affirmation seems to reinforce concerns over self-integrity rather than diminish them (cf. Van Koningsbruggen & Das, 2008b). When the level of self-threat exceeds a given self-affirmation, self-affirmation seems to prevent people from drawing the
most accurate conclusion that follows from the health risk information, thereby impeding
adaptive responses to this personally important information.

Previous research suggested that incorporating a self-affirmation in a health promotion
campaign, for instance, by encouraging receivers of health promotion messages to reflect on a
personally important value could increase the effectiveness of these campaigns (e.g., Sherman
et al., 2000). However, the present findings underscore the importance of carefully attending
to self-threat level before deciding to adopt such a strategy. When people feel moderately
threatened, self-affirmation may indeed have positive effects on message processing and
persuasion, thereby making health promotion efforts more effective. However, when people,
for instance, receive personal feedback regarding their vulnerability to a health risk, they are
likely to feel highly vulnerable to this risk. Under these conditions, self-affirmation seems to
be a risky strategy that may decrease careful processing of important health information and
backfire on persuasion, thereby making health promotion efforts less effective.

As with any study, this study has some limitations that should be addressed in future
research. A first potential limitation concerns our behavioral measure. Although we replicated
earlier findings that self-affirmed participants more often requested additional health
information when moderately threatened (e.g., Jessop et al., in press; Sherman et al., 2000), it
does not provide evidence that these participants subsequently engaged in exercises to prevent
RSI when doing a lot of computer work (cf. Epton & Harris, 2008; Harris et al., 2007). Thus,
we do not know for sure whether self-affirmed participants actually translated this pre-
behavioral step into action. However, recent findings suggest that self-affirmation can indeed
promote actual behavior change (Epton & Harris, 2008; Van Koningsbruggen & Das, in
press). Moreover, our findings suggest that self-affirmation may backfire on actual behavior
change under high self-threat conditions. Thus, future studies are necessary to further
investigate the impact of self-affirmation on behavior change under varying levels of self-threat.

Second, the present study comprised an initial investigation of the impact of self-affirmation on persuasive outcomes under a high self-threat condition. Because of time constraints, we used a single-item measure of attitudes, which may have affected the reliability of this measure. However, the item included in the present study has been used in previous research in which it has shown convergent validity with other comparable items (Van Koningsbruggen & Das, 2008b). Moreover, the results obtained under moderate self-threat conditions, fully replicated the effects of self-affirmation on health persuasion found in previous studies (e.g., Harris & Napper, 2005; Jessop et al., in press; Sherman et al., 2000). Nonetheless, a replication of the present study using a multiple-item measure would be desirable.

In conclusion, this study provided a first test of the impact of self-affirmation on responses to threatening health information following a severe threat to the self. While previous self-affirmation research in the health domain mainly demonstrated positive effects of self-affirmation, the present findings stress the importance of carefully attending to the level of self-threat people experience when faced with threatening health information in determining whether self-affirmation will have beneficial effects or not.