In striving for peace of mind, people are often motivated to hold on to important personal values, social identities, beliefs, or goals. Accordingly, receivers may bias their processing and evaluation of information that threatens important self-concepts in order to arrive at a preferred conclusion (e.g., Kunda, 1990). For instance, when health information challenges one’s image of being a healthy, responsible person, people may derogate the information, minimize the seriousness of the risks, or engage in wishful thinking (e.g., Jemmott et al., 1986; Liberman & Chaiken, 1992). Several theories suggest that this defensiveness originates from concerns over self-worth (e.g., Aronson E., 1969; Steele, 1988; Tesser & Cornell, 1991). Self-affirmation theory (Steele, 1988) - a theory about how people deal with concerns over self-worth - has played a central role in this respect (for a review, see Sherman & Cohen, 2006).

According to self-affirmation theory (Sherman & Cohen, 2006; 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. 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 upon alternative sources unrelated to the threat, such as reflecting upon an other 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 when threatened.
Numerous studies across a wide variety of domains and research paradigms support the notion that self-affirmation reduces defensive responses (for a review, see Sherman & Cohen, 2006) and show that people become more open-minded toward threatening information after self-affirmation (e.g., Cohen et al., 2000; Correll et al., 2004; Harris & Napper, 2005; Sherman et al., 2000). Unfortunately, consistent empirical evidence regarding the cognitive processes that mediate self-affirmation effects remains lacking. Furthermore, little is known about potential factors that moderate the effect of self-affirmation on cognitive processing (Sherman & Cohen, 2006).

The present research examines how self-affirmation affects cognitive processing of threatening information and when self-affirmation results in extensive, careful information processing. In the following paragraphs, we discuss prior work on self-affirmation and the processing of self-threats, and identify level of self-threat as a potential moderating variable. With three experiments we empirically tested this assertion.

Self-Affirmation and Information Processing

Although self-affirmation has been posited to increase extensive, careful information processing (Steele, 1988), only a few studies have employed methods or measures that provide robust tests of how self-affirmation affects cognitive processing. The extent of information processing can be assessed at the explicit level by varying the quality of the arguments in a message and by examining the impact of these arguments on attitudes. When people differentiate between weak and strong arguments, it suggests that they have carefully attended to and thought about the (de)merits of the message (Petty & Cacioppo, 1986). On the implicit level, the extent of information processing can be inferred from the accessibility of topic-related cognitions, and extensive processing is implied when threat-related cognitions
are accessible after a relevant self-threat. Close examination of previous findings reveals that self-affirmation can affect cognitive processing in different ways.

Two recent studies examined the effect of self-affirmation on explicit information processing by manipulating the quality of the arguments in a message. Briñol, Petty, Gallardo, and DeMarree (2007) manipulated the quality of the arguments in an advertisement about a phone. Participants read and evaluated this advertisement and reported their attitudes toward the product. Self-affirmation prior to reading this message increased participants’ confidence in their current views and made them less sensitive to the quality of the arguments than their non-affirmed counterparts. Self-affirmation thus decreased the extent of information processing. Correll, Spencer, and Zanna (2004) manipulated the quality of the arguments in a discussion about a tuition raise. Participants against a tuition raise watched a videotaped discussion and evaluated the persuasiveness of the arguments. It was found that self-affirmed participants were more sensitive to the quality of the arguments than their non-affirmed counterparts, particularly if this issue was important for them. Here, self-affirmation increased the extent of information processing.

Other studies have shown how self-affirmation affects implicit cognitive processes in the face of self-relevant threats by examining the accessibility of topic-related cognitions. Koole, Smeets, Van Knippenberg, and Dijkstra (1999) provided participants with failure feedback information on an alleged IQ test and found that self-affirmation decreased the cognitive accessibility of threat-related cognitions. Likewise, self-affirmation has been shown to reduce the accessibility of death-related cognitions after a mortality salience manipulation (Schmeichel & Martens, 2005) and of social rejection thoughts in a threatening social situation (Schimel et al., 2004). In these studies, self-affirmation thus decreased the extent of information processing.
In the respective research domains, self-affirmation generally increased positive outcomes on the dependent variable of interest, as it decreased defensive responses to preference-inconsistent information (Correll et al., 2004) and reduced rumination about a frustrated goal (Koole et al., 1999). Focusing on the underlying cognitive processes, however, self-affirmation affected these processes in two different ways. In some instances, self-affirmation increased the extent of information processing; in other instances, self-affirmation decreased the extent of information processing. We propose that these divergent findings may be explained by variations in the level of self-threat. Learning that one has failed on an IQ test (Koole et al., 1999), or a confrontation with perhaps the biggest threat to self-integrity – death (Schmeichel & Martens, 2005), poses a greater threat to the self than a potential tuition raise for students (Correll et al., 2004). A tuition raise, in turn, poses a greater self-threat than a message about a cell phone (Briñol et al., 2007). Thus, advancing that the level of self-threat moderates the impact of self-affirmation on the processing of self-threatening information, self-affirmation may increase extensive information processing under conditions of moderate self-threat, whereas it may decrease extensive information processing under conditions of low and high self-threat.

On a theoretical level, our proposition about the moderating role of self-threat level in the cognitive processes underlying self-affirmation effects nicely fits within the assumptions of self-affirmation theory (Steele, 1988). In the next paragraph, we focus on the impact of self-affirmation in settings that are threatening to the self and propose that the hypothesized moderating role of self-threat level is sound with the basic premise of self-affirmation theory (i.e., that people are primarily concerned with global self-integrity). Furthermore, we discuss what this means for the processing of threatening informational messages.
Self-Affirmation, Level of Self-Threat, and the Processing of Threatening Messages

According to self-affirmation theory, self-affirmation makes self-resources cognitively accessible (or puts the self-system "on-line", cf. Steele et al., 1993) and enables people to view the threat within a “broader, larger view of the self” (Sherman & Cohen, 2006, p. 189). Put differently, self-affirmation stimulates integration of the threat into the self-system, that is, a larger cognitive network that consists of self-representations of a person’s motives, goals, needs, values and autobiographical experiences (cf. Koole & Kuhl, 2003). This integration into a larger view of the self enables people to cope with a threat without freezing on its implications for self-integrity (cf. Sherman & Cohen, 2006). As a consequence, self-affirmation increases extensive, careful processing of threatening information. It is important to keep in mind, however, that the ultimate goal of self-affirmation is not to increase extensive processing of a particular threat, but to restore the general integrity of the self (Steele, 1988). In this perspective, extensive information processing should be seen as a fortunate by-product of self-affirmation.

Self-affirmation theory further predicts that there are limits to the effects of self-affirmation. When the level of self-threat exceeds a given self-affirmation, self-affirmation may become ineffective in restoring the general integrity of the self (Steele, 1988, p. 291). In these conditions, the fortunate by-products of self-affirmation will also cease to occur, and self-affirmation will become ineffective in promoting extensive, careful information processing. Moreover, when self-resources are not of sufficient magnitude to cope with a threat, self-affirmation may actually reinforce concerns over self-integrity, rather than diminish them, and may thus backfire on information processing. This effect occurs because self-affirmation puts the entire self-system “on-line”, thus making the entire self-system vulnerable to a severe threat that it is unable to cope with. For high levels of self-threat, self-affirmation thus ironically adds to the problem by increasing concerns over self-integrity (see...
also Crocker, 2002; Crocker & Park, 2004). In these conditions, self-affirmation is likely to
decrease extensive information processing in the service of self-integrity maintenance (Koole,
2004; Nowak et al., 2000; Sedikides & Green, 2004).

The question whether this decrease in extensive processing for high levels of self-
threat is good or bad news remains a matter of perspective, and ultimately depends on the
research question (i.e., dependent variable) of interest. For instance, when self-affirmation
decreases extensive processing, people ruminate less about a failure experience (Koole et al.,
1999) and are less likely to derogate someone who has violated one’s worldview (Schmeichel
& Martens, 2005). However, when focusing exclusively on information processing, self-
affirmation may have an adverse effect when self-threat levels become severely high, as it
may decrease extensive, careful information processing.

**What Constitutes a High Self-Threat?**

Prior self-affirmation research has focused on a broad range of self-threats (e.g., see
Sherman & Cohen, 2006) that, in all probability, have ranged from relatively low or moderate
(e.g., a tuition increase) to very severe (e.g., death). However, previous self-affirmation
research did not explicitly assess whether self-threat level moderates the effects of self-
affirmation. To ensure a fair test of this proposition, the effects of self-threat level and self-
affirmation should be assessed within one research paradigm. Because health represents a
topic for which level of self-threat can be effectively varied, it provides a domain in which our
theoretical analysis can be properly tested. In addition, previous research in this domain
already found reliable effects of self-affirmation (e.g., Harris & Napper, 2005; Reed &
Aspinwall, 1998; Sherman et al., 2000; Van Koningsbruggen & Das, in press).

In the health domain, self-threat level is conceptualized as perceptions of vulnerability
to and severity of a health risk (e.g., Rogers, 1975, 1983; Rosenstock, 1974) . Importantly,
research has shown that once severity has reached a certain level, outcomes are solely determined by perceptions of vulnerability (Das et al., 2003; De Hoog et al., 2005; Kruglanski & Klar, 1985; Weinstein, 1988). Thus, however severe a health risk, it is unlikely that people will perceive a threat to the self unless they are personally vulnerable to the health risk.

Vulnerability thus constitutes the key determinant of level of self-threat in the health domain. Vulnerability can either be classified on *measurements* of perceived vulnerability or actual behavior (such as the amount of caffeine or alcohol consumption; e.g., Harris & Napper, 2005; Sherman et al., 2000), or *manipulated* by providing participants with false feedback regarding their vulnerability level (e.g., Das et al., 2003; De Hoog et al., 2005; Ditto & Lopez, 1992). The more people perceive themselves as vulnerable, the greater the threat to the self. Moreover, the different methods to assess vulnerability levels (i.e., measurement vs. manipulation) provide us with a unique opportunity to create different levels of self-threat.

In particular, a measurement of vulnerability (e.g., comparing non-smokers with smokers) is likely to pose a lower threat to the self than when vulnerability is manipulated by providing participants with false feedback. This effect occurs because people with unhealthy lifestyles (e.g., smokers) learn to develop defensive strategies to deal with this continually activated dissonance-provoking behavior (see Gibbons et al., 1997, for evidence relevant to this assumption). This “minimization” of health threats following previous experience with a threat is underscored by research on desensitization (cf. Wolpe, 1982), and inoculation (cf. McGuire, 1968). In line with this minimization-account, measurements of vulnerability may therefore reflect only moderate feelings of vulnerability (i.e., moderate self-threat). In contrast, when vulnerability is manipulated by providing participants with false feedback that people have no previous experience with, it becomes much more difficult to discount one’s vulnerability, thus producing heightened feelings of vulnerability (i.e., high self-threat; cf. Loewenstein et al., 2001). This suggests that a manipulation of vulnerability is likely to bring
forth a higher threat to the self than when vulnerability is not manipulated, but measured. The present research provides a first explicit test of this assumption.

Overview of the Present Studies

In this research we examined the effect of self-affirmation on information processing under varying levels of threat to the self (low, moderate, and high). We predicted that self-affirmation would increase extensive, careful information processing under moderate self-threat conditions, and decrease extensive, careful information processing under low and high self-threat conditions. The threatening information under investigation concerned a health topic, and we conceptualized self-threat level as participants’ vulnerability to a health risk.

Study 3.1 tested the assertion that a manipulation of vulnerability to a health risk induces a higher threat to the self than when vulnerability is measured. In addition, this study provided an initial test of our main hypothesis regarding the effect of self-affirmation on information processing. Studies 3.2 and 3.3 examined the cognitive processes following self-affirmation under different self-threat levels in more detail, by including a manipulation of argument quality in the health message, and by measuring participants’ cognitive responses (Study 3.3) and attitudes (Studies 3.2 & 3.3). Past research has shown that extensive, careful information processing is indicated by a significant differentiation between strong and weak arguments in a message, whereas less thorough processing is implied when this differentiation is absent on the dependent measures (Petty & Wegener, 1999). It was hypothesized that self-affirmation would increase sensitivity to the quality of the arguments in the message under moderate self-threat conditions. In contrast, we hypothesized that self-affirmation would decrease sensitivity to argument quality under low and high self-threat conditions.
Study 3.1

To test our main hypothesis that self-threat level moderates the impact of self-affirmation on information processing, we first had to establish three different levels of self-threat. In Study 3.1, we therefore modified the research paradigm most frequently used to study the effect of (measured) vulnerability on message processing, that is, the caffeine consumption paradigm (see Kunda, 1987; Liberman & Chaiken, 1992; Raghunathan & Trope, 2002; Reed & Aspinwall, 1998; Sherman et al., 2000). In this paradigm, participants read a health message that contains threatening information about a link between coffee consumption and severe health problems. To assess the effect of vulnerability on the dependent variables of interest, previous studies applied a “measurement” method, that is, participants were divided in two conditions (i.e., low and high vulnerability) based on their actual behavior (e.g., non-coffee drinkers vs. coffee drinkers). In the present study we adopted a similar approach, but we included two additional conditions in which vulnerability to the health risk was manipulated, rather than measured, by providing participants with false feedback regarding their vulnerability level prior to reading the same health information. This resulted in four conditions, which were crossed with self-affirmation (self-affirmation status: non-affirmed vs. self-affirmed).

Our key dependent measures were perceptions of vulnerability to the health risk (to measure level of self-threat) and the number of critical thoughts participants reported about the message (to explore cognitive processing). We expected that a manipulation of vulnerability induces a higher level of self-threat than measurements of vulnerability, and that the results should reveal three different levels of perceived vulnerability: measured low vulnerability participants should feel least vulnerable (i.e., low self-threat); manipulated high vulnerability participants should feel most vulnerable to the health risk (i.e., high self-threat), and measured high and manipulated low vulnerability participants should take an intermediate
position (i.e., moderate self-threat). Finally, we predicted that self-affirmation would increase the extent of information processing under conditions of moderate self-threat (cf. Correll et al., 2004), but would decrease the extent of information processing under conditions of low (cf. Briñol et al., 2007) and high self-threat.

**Method**

**Design and Participants**

Participants were 139 university students (61 males, 78 females; $M_{age} = 22.78$ years, $SD_{age} = 5.89$ years). In exchange, they received €3.50. Non-coffee drinkers were assigned to the measured low vulnerability condition, while coffee drinkers were randomly assigned to the remaining three conditions.

**Procedure and Materials**

Participants were informed that they would participate in two separate studies, one about values and the other about consumption patterns concerning several beverages. 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 Allport-Vernon-Lindzey study of Values (AVL; 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 were presented with several questions designed to identify their coffee drinking behavior (e.g., coffee drinker yes/no, if yes: what brands they consumed). To bolster our cover story, similar questions were asked about several other beverages. The computer then assigned non-coffee drinking participants to the measured low vulnerability condition (vulnerability-method condition 1; no feedback), while coffee drinking participants were randomly assigned to the remaining three conditions. One-third of
the coffee drinkers was assigned to the measured high vulnerability condition (vulnerability-method condition 2; no feedback), and two-third to the conditions in which vulnerability was manipulated by providing participants with false feedback (vulnerability-method conditions 3 & 4; low and high vulnerability feedback respectively). Participants in vulnerability-method conditions 1 and 2 continued with the self-affirmation manipulation; participants in vulnerability-method conditions 3 and 4 were first given false feedback regarding their vulnerability level prior to the self-affirmation manipulation.

While all participants learnt that coffee drinking could cause severe health problems, participants who received vulnerability feedback learnt that only some brands were associated with these health problems. They were told about a test that analyzed whether the ingredients of the brands they used could cause severe health problems. They were led to believe that the computer analyzed the details they provided earlier 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 low vulnerability feedback condition (vulnerability-method condition 3), they saw a square, filled with a bright green color, that displayed: “Your risk of developing health problems is quite low due to the coffee brands you use.” In the high vulnerability feedback condition (vulnerability-method condition 4), participants saw a square, filled with a bright red color, that displayed: “Your risk of developing health problems is quite high due to the coffee brands you use.”

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

Health message. Participants read a fictitious article, entitled “Coffee associated with intestinal cancer”, which was supposedly published in “The Journal of Health” (based on Liberman & Chaiken, 1992). The article described intestinal cancer, and research findings that confirmed the link between coffee consumption and this disease, implying that one should decrease coffee consumption. The message was held constant across conditions.

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 studies were related.

Dependent Variables

Manipulation checks. To confirm the success of the vulnerability manipulation and to compare the effects of a manipulation of vulnerability with measurements, participants rated four items that measured their perceived vulnerability to health problems related to coffee consumption on a 7-point scale (e.g., “The probability that I will experience health problems because of my coffee consumption is high”; 1 = strongly disagree, 7 = strongly agree; α = .93). Higher scores indicate higher perceptions of vulnerability.

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.

*Cognitive processing.* Participants were asked, in an open-ended format, to report any critical thoughts they had while reading the message (e.g., “The article raised an important issue, but more could have been said about how the research was conducted”; Range 0-4).¹

**Results**

*Manipulation Checks*

*Perceived vulnerability.* A one-way analysis of variance (ANOVA) with vulnerability-method condition as factor, revealed a significant effect of condition, $F(3, 135) = 18.36, p < .001, \eta_p^2 = .29$. As predicted, Tukey HSD post hoc analyses revealed three levels of perceived vulnerability. As can be seen in Figure 3.1, non-coffee drinkers felt least vulnerable to the health risk (measured low vulnerability condition; $M = 1.52, SD = 0.86$), while coffee drinkers who received the high vulnerability feedback felt most vulnerable to the health risk (manipulated high vulnerability condition; $M = 3.77, SD = 1.40$). Coffee drinkers who received no feedback (measured high vulnerability condition; $M = 2.66, SD = 1.43$) and coffee drinkers who received the low vulnerability feedback (manipulated low vulnerability condition; $M = 2.40, SD = 1.26$) felt equally vulnerable to the health risk. Consistent with our predictions the results thus showed three levels of self-threat. We therefore analyzed the dependent measure using a 3 (self-threat level: low vs. moderate vs. high) x 2 (self-affirmation status: non-affirmed vs. self-affirmed) ANOVA.²
Figure 3.1. Mean perceived vulnerability as a function of vulnerability-method condition (Study 3.1). Capped vertical bars denote 1 standard error.

Note. The measured low vulnerability condition differed significantly from the measured high ($p = .001$), manipulated low ($p = .020$), and manipulated high vulnerability conditions ($p < .001$). The manipulated high vulnerability condition differed significantly from the measured high ($p = .002$), and manipulated low vulnerability conditions ($p < .001$). The measured high vulnerability and manipulated low vulnerability conditions did not differ significantly from each other ($p = .83$).

Self-affirmation. A one-way ANOVA confirmed that self-affirmed participants endorsed the manipulated value more often ($M = 6.56, SD = 1.69$) than non-affirmed participants ($M = 2.71, SD = 1.72$), $F(1, 137) = 176.36, p < .001, \eta_p^2 = .56$.

Cognitive Processing

The number of critical thoughts listed by the participants was correlated with familiarity with the link ($r = -.25, p = .003$). Therefore, this measure was subjected to a 3 (self-threat level) x 2 (self-affirmation status) analysis of covariance (ANCOVA) with familiarity with the link as covariate. The analysis revealed a significant interaction effect between self-threat level and self-affirmation, $F(2, 131) = 5.32, p = .006, \eta_p^2 = .08$. Simple
How Self-Affirmation Affects Explicit Health Information Processing
effects analysis showed that self-affirmation increased the number of critical thoughts reported in the moderate self-threat condition, \( F(1, 131) = 5.44, p = .021, \eta_p^2 = .04 \), whereas it decreased the number of critical thoughts in the high self-threat condition, \( F(1, 131) = 4.24, p = .041, \eta_p^2 = .03 \). Self-affirmation had no effect under low self-threat conditions \( (F < 1) \).

Relevant means are displayed in Table 3.1.

Table 3.1. Means (SD's) of number of critical thoughts as a function of self-threat level and self-affirmation (Study 3.1)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Critical thoughts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Self-affirmation status</td>
</tr>
<tr>
<td></td>
<td>Non-affirmed</td>
</tr>
<tr>
<td>Low self-threat</td>
<td>0.68 (0.19)\text{a}</td>
</tr>
<tr>
<td>Moderate self-threat</td>
<td>0.53 (0.14)\text{a}</td>
</tr>
<tr>
<td>High self-threat</td>
<td>0.86 (0.21)\text{a}</td>
</tr>
</tbody>
</table>

*Note.* Means with different subscripts (in rows) differ significantly from each other \( (p < .05) \).

Discussion

The results of Study 3.1 show that participants who received false feedback regarding their high vulnerability level reported higher levels of perceived vulnerability to a health risk than participants who were classified as highly vulnerable based on measurements. Consistent with our hypothesis, the results revealed three levels of self-threat: low self-threat (measured low vulnerability condition), moderate self-threat (measured high and manipulated low vulnerability conditions), and high self-threat (manipulated high vulnerability condition).

These findings reflect that a manipulation of vulnerability induces a higher self-threat as it becomes more difficult to discount one’s vulnerability (cf. Loewenstein et al., 2001; McGuire, 1968; Wolpe, 1982). We also predicted that self-threat level would moderate the effect of self-affirmation on information processing. The findings on critical thoughts provide initial
support for this hypothesis. Self-affirmation increased the number of critical thoughts under moderate levels of self-threat, and decreased the number of critical thoughts under conditions of high self-threat. These findings are a first indication that self-affirmation increases the extent of information processing for moderate self-threats, but decreases extensive, careful information processing for high levels of self-threat.

In the present study, no self-affirmation effect was found under low self-threat levels. The absence of an argument quality manipulation and attitude measure may have limited our ability to detect this effect. Information processing can be assessed in a more robust way by varying the quality of the arguments in a message and by examining the impact of these arguments on attitudes (Petty & Cacioppo, 1986). In Studies 3.2 and 3.3, we added a manipulation of argument quality to the experimental design to provide a robust test of the effect of self-affirmation on information processing following low, moderate, and high threats to the self.

**Study 3.2**

Study 3.2 further tested the impact of self-affirmation on information processing under low, moderate, and high self-threat conditions. To obtain a more detailed picture of how self-affirmation affects information processing, we manipulated the quality of the arguments in the health message and measured participants’ attitudes. We assessed low and moderate self-threat levels using the measurement method (cf. Harris & Napper, 2005) and induced a high self-threat level by manipulating vulnerability (cf. Study 3.1). To ensure that the findings of Study 3.1 were not health-topic specific, we used a different health risk: Repetitive Strain Injury (RSI). It was hypothesized that self-affirmation would increase sensitivity to the quality of the arguments in the message under moderate self-threat conditions, and would decrease sensitivity to argument quality under low and high self-threat conditions. Thus, we
How Self-Affirmation Affects Explicit Health Information Processing

expected to obtain a three-way interaction between self-threat level, self-affirmation, and argument quality on attitudes.

Method

Design and Participants

The hypotheses were tested in a 3 (self-threat level: low vs. moderate vs. high) x 2 (self-affirmation status: non-affirmed vs. self-affirmed) x 2 (argument quality: strong vs. weak) between-participants factorial design. Participants were 197 university students (42 males, 155 females; $M_{age} = 23.92$ years, $SD_{age} = 4.69$ years). They were randomly assigned to experimental conditions. As compensation, participants could win gift vouchers.

Procedure and Materials

Participants were informed that they would participate in two separate studies: one about health and the other about values.

Self-threat level. After a short introduction about RSI, the computer randomly assigned participants to the measurement condition (i.e., conditions of low and moderate self-threat) or to the condition in which self-threat level was manipulated (i.e., condition of high self-threat). Participants assigned to the measurement condition responded to four items that assessed participants’ perceived vulnerability to RSI (cf. Study 3.1, $a = .87$) and were divided into a low ($n = 67$) and moderate self-threat condition ($n = 59$) on the basis of their responses on the vulnerability measure. Similar to Study 3.1, participants assigned to the high self-threat condition were provided with false feedback on a fictitious health test to manipulate their vulnerability to RSI. Participants were led to believe that this test measured their risk of developing RSI-related health problems in the future. Participants completed a bogus personality 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
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.** Participants were asked to complete a short survey that was supposedly unrelated to the RSI-study. This survey consisted of 10 questions that contained the manipulation of self-affirmation. In the self-affirmed status condition, the questions induced affirmative responses about participants’ social value orientation. In the non-affirmed status condition, participants responded to questions in which this opportunity was not given (cf. Reed & Aspinwall, 1998).

**Argument quality.** Following the self-affirmation manipulation, participants read a health message supposedly published in “The Journal of Health.” The message described the consequences of RSI and prevention training as a means to reduce the negative health consequences, and was supported by five weak or five strong arguments. These arguments were selected from a pilot study (cf. Petty & Cacioppo, 1986). Examples are: “Research has demonstrated that knowledge of RSI-prevention strategies improved physical and psychological well-being” (strong argument), and “The institute that developed the RSI-prevention training would not put the training on the market unless it was effective” (weak argument).

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 studies were related.
Dependent Variables

Manipulation checks. To determine whether we successfully created a high self-threat condition, participants rated similar items as in Study 3.1 to measure perceptions of vulnerability (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 = .77). Similar to Study 3.1, we expected that a manipulation of vulnerability would induce higher perceptions of vulnerability than when vulnerability was not manipulated, but measured. To assess the effectiveness of the argument quality manipulation in the health message, three items on a 7-point semantic differential scale measured participants’ perceptions of how good, supportive, and useful the arguments in the message were (α = .89). The manipulation check for self-affirmation was the same as in Study 3.1.

Attitudes. Participants rated four items (α = .84) that measured their attitude toward the health message on a 7-point semantic differential scale (e.g., how bad-good and useless-useful participants thought this message was), with higher scores indicating a more favorable attitude.4

Results

Manipulation Checks

Self-threat level. A one-way ANOVA with method (measurement vs. manipulation) as factor confirmed that participants in the condition in which vulnerability was manipulated (i.e., high self-threat condition) felt significantly more vulnerable to the health risk (M = 3.87, SD = 1.22) than participants in the measurement (i.e., low and moderate self-threat) conditions (M = 2.86, SD = 1.22), F(1, 195) = 31.55, p < .001, ηp² = .14.
**Self-affirmation.** A one-way ANOVA confirmed that self-affirmed participants more often endorsed the manipulated value \((M = 7.59, SD = 1.72)\) than non-affirmed participants \((M = 3.71, SD = 1.62)\), \(F(1, 195) = 265.85, p < .001, \eta^2_p = .58\).

**Argument quality.** A one-way ANOVA revealed the expected main effect for argument quality, \(F(1, 195) = 52.98, p < .001, \eta^2_p = .21\). Argument quality was perceived as higher in the strong \((M = 4.75, SD = 0.88)\) than in the weak arguments conditions \((M = 3.64, SD = 1.20)\).

**Attitudes**

A 3 (self-threat level) x 2 (self-affirmation status) x 2 (argument quality) ANOVA on attitudes toward the health message revealed a main effect of argument quality, \(F(1, 185) = 15.82, p < .001, \eta^2_p = .08\). This main effect was qualified by a significant three-way interaction between self-threat level, self-affirmation, and argument quality, \(F(2, 185) = 6.79, p = .001, \eta^2_p = .07\). As expected, simple effects analysis showed that in the moderate self-threat condition, self-affirmed participants differentiated between strong and weak arguments, \(F(1, 185) = 14.10, p < .001, \eta^2_p = .07\), whereas non-affirmed participants did not, \(F < 1\). In contrast, in the high self-threat condition, self-affirmed participants did not differentiate between strong and weak arguments, \(F(1, 185) = 1.82, p = .18\), whereas non-affirmed participants (near-significantly) did, \(F(1, 185) = 3.48, p = .064, \eta^2_p = .02\). In the low self-threat condition, non-affirmed participants differentiated between arguments, \(F(1, 185) = 15.71, p < .001, \eta^2_p = .08\), whereas self-affirmed participants did not, \(F < 1\). Thus, self-affirmation increased extensive information processing in the moderate self-threat condition, and decreased extensive processing in the low and high self-threat conditions. Relevant means are displayed in Table 3.2.
How Self-Affirmation Affects Explicit Health Information Processing

Table 3.2. Means (SD’s) of attitudes as a function of self-threat level, self-affirmation, and argument quality (Study 3.2)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Argument quality</td>
</tr>
<tr>
<td></td>
<td>Weak</td>
</tr>
<tr>
<td>Low self-threat</td>
<td></td>
</tr>
<tr>
<td>Non-affirmed</td>
<td>3.93 (1.05)\textsubscript{a}</td>
</tr>
<tr>
<td>Self-affirmed</td>
<td>4.20 (0.84)\textsubscript{a}</td>
</tr>
<tr>
<td>Moderate self-threat</td>
<td></td>
</tr>
<tr>
<td>Non-affirmed</td>
<td>4.44 (1.07)\textsubscript{a}</td>
</tr>
<tr>
<td>Self-affirmed</td>
<td>3.98 (1.07)\textsubscript{a}</td>
</tr>
<tr>
<td>High self-threat</td>
<td></td>
</tr>
<tr>
<td>Non-affirmed</td>
<td>4.56 (1.08)\textsubscript{a}</td>
</tr>
<tr>
<td>Self-affirmed</td>
<td>4.92 (0.64)\textsubscript{a}</td>
</tr>
</tbody>
</table>

Note. Means with different subscripts (in rows) differ significantly from each other ($p < .05$).

Discussion

The present findings demonstrate that self-threat level moderates the effect of self-affirmation on information processing. On attitudes, a three-way interaction between self-threat level, self-affirmation, and argument quality was found. In the moderate self-threat condition, self-affirmation increased sensitivity to the quality of the arguments in the health message; participants had more favorable attitudes when the message was supported by strong arguments rather than weak arguments. In contrast, in the low and high self-threat conditions, self-affirmation decreased participants’ sensitivity to argument quality in the health message; no effects of argument quality were observed on attitudes. Overall, the present findings extend Study 3.1 by providing a robust test that self-affirmation increases the extent of information processing under moderate self-threat conditions (cf. Correll et al., 2004), and decreases the extent of information processing under low (cf. Briñol et al., 2007) and high self-threat conditions. These findings attest to the limits of self-affirmation by showing that self-affirmation may backfire on information processing under these conditions.
In the subsequent study, we sought to extend our test of the limits of self-affirmation by focusing on the most relevant conditions, that is, the moderate and high self-threat conditions (cf. McQueen & Klein, 2006; Sherman et al., 2000, Study 2). One drawback of Study 3.2 was that the moderate and high self-threat conditions varied not only with respect to perceptions of vulnerability, but also with respect to the method (measurement vs. manipulation). To increase confidence that the findings were indeed due to differences in vulnerability (self-threat) level, the objective of Study 3.3 was to replicate the effect of self-affirmation on an explicit processing measure while holding the manipulation of self-threat constant.

**Study 3.3**

In this study self-threat level was experimentally manipulated by providing participants with false feedback regarding their vulnerability to a health risk. The condition of low self-threat was dropped from the design. Similar to Study 3.2, we included a manipulation of argument quality and measured participants’ attitudes. Moreover, to obtain a more detailed picture of how self-affirmation affects information processing, we also measured participants’ cognitive responses. Extending previous studies (e.g., Briñol et al., 2007; Correll et al., 2004; Sherman et al., 2000), we thus employed both methods (i.e., argument quality manipulation) and measures (i.e., thought-listing task) in order to obtain a detailed picture of information processing. To extend the utility and generality of our findings, we used a different health topic: stress. Finding converging evidence for the moderating role of self-threat level using different methods to elicit the level of self-threat, and across different health risks, would assure the validity of the results we obtained in Study 3.2. It was hypothesized that self-affirmation would increase sensitivity to the quality of the arguments in the message under moderate self-threat conditions, and would decrease
sensitivity to argument quality under high self-threat conditions. Thus, similar to Study 3.2, we expected a three-way interaction between self-threat level, self-affirmation, and argument quality on cognitive responses and attitudes. Finally, we assessed mood in Study 3.3 to test whether the effects of self-affirmation may be explained by changes in participants’ mood. Although self-affirmation typically has no effect on mood (see Sherman & Cohen, 2006; Steele, 1988), ruling mood out as an alternative explanation for the observed effects is important as it could plausibly affect information processing in similar ways (e.g., Raghunathan & Trope, 2002).

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) x 2 (argument quality: strong vs. weak) between-participants factorial design. Participants were 135 university students (35 males, 100 females; $M_{age} = 22.19$ years, $SD_{age} = 2.60$ years). They were randomly assigned to experimental conditions. As compensation, participants could win cash prizes.

Procedure and Materials

Participants were informed that they would participate in several separate studies: one about health, and the other about values. Similar to Study 3.1, participants first ranked the six values of the AVL study of Values (Allport et al., 1960) and then 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 (cf. Study 3.2), measured their risk of developing stress-related health problems in the future. 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 stress-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 stress-related health problems is quite high.” Participants then completed the self-affirmation manipulation.

*Self-affirmation.* The manipulation of self-affirmation was similar to the one used in Study 3.1.

*Argument quality.* Following the self-affirmation manipulation, participants read the same strong or weak health message used in Study 3.2 (of course, the word “RSI” was replaced by “stress”).

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.* The manipulation checks for self-threat level ($\alpha = .94$), self-affirmation and argument quality ($\alpha = .81$) were the same as in Study 3.2.

*Mood.* Participants rated nine items that measured their mood on a 7-point semantic differential scale (e.g., how sad-happy, unpleasant-pleasant, pessimistic-optimistic they felt; $\alpha = .90$).

*Cognitive processing.* Participants completed a thought-listing task to assess cognitive processing of the health message. They were asked to write down the thoughts they had about the content of this message. Following Petty and Cacioppo (1986) and Sherman and
colleagues (2000), two independent judges categorized these thoughts in positive (e.g., “This article raised an important issue, people should read this message”; Range 0-4; Kappa = .72) and negative issue-relevant thoughts (e.g., “The article should have contained more detailed information and empirical evidence”; Range 0-4; Kappa = .81). This categorization was based on the following premises: (1) overall valence of thoughts is the most extensively investigated content dimension (e.g., see Eagly & Chaiken, 1993), and (2) the polarity of the categories is a good predictor of post-message attitude change (see reviews by Eagly & Chaiken, 1993; Petty & Cacioppo, 1986). A thought-valence index was created based on criteria specified in previous studies (Killeya & Johnson, 1998; Sherman et al., 2000) using the following formula: (positive issue-relevant thoughts + 1) / (total issue-relevant thoughts + 1). Index-scores could range between 0 and 1. Scores that approach 0 indicate predominantly negative thoughts, and scores that approach 1 signal predominantly positive thoughts.

Attitudes. Participants rated similar items used in Study 3.2 to measure their attitudes toward the health message (α = .86).

Results

Manipulation Checks

Self-threat level. Four participants were removed from the analyses because their scores fell outside the normal range on this measure (SD > 2.50, evenly distributed across conditions). A one-way ANOVA confirmed that participants in the high self-threat condition felt significantly more vulnerable to the health risk (M = 4.31, SD = 1.43) than participants in the moderate self-threat condition (M = 3.33, SD = 1.42), F(1, 133) = 15.97, p < .001, ηp² = .11.
**Self-affirmation.** A one-way ANOVA confirmed that self-affirmed participants endorsed the manipulated value more often ($M = 7.21, SD = 1.61$) than non-affirmed participants ($M = 2.97, SD = 1.73$), $F(1, 133) = 216.37, p < .001, \eta_p^2 = .62$.

**Argument quality.** A one-way ANOVA confirmed that argument quality was perceived as higher in the strong arguments condition ($M = 4.98, SD = 0.79$) than in the weak arguments condition ($M = 4.17, SD = 1.09$), $F(1, 133) = 24.15, p < .001, \eta_p^2 = .15$.

**Mood**

A 2 (self-threat level) x 2 (self-affirmation status) x 2 (argument quality) ANOVA on mood revealed no significant effects (all $ps > .21$).

**Cognitive Processing**

A 2 (self-threat level) x 2 (self-affirmation status) x 2 (argument quality) ANOVA on the thought-valence index revealed a significant main effect for argument quality, $F(1, 127) = 17.04, p < .001, \eta_p^2 = .12$. This main effect was qualified by the expected significant three-way interaction between self-threat level, self-affirmation, and argument quality, $F(1, 127) = 3.84, p = .05, \eta_p^2 = .03$. Simple effects analysis revealed that in the moderate self-threat condition, self-affirmed participants differentiated between arguments, $F(1, 127) = 13.56, p < .001, \eta_p^2 = .10$, whereas non-affirmed participants did not, $F < 1$. In contrast, in the high self-threat condition, self-affirmed participants did not differentiate between strong and weak arguments, $F < 2, p > .20$, whereas non-affirmed participants did, $F(1, 127) = 5.91, p = .016, \eta_p^2 = .04$. Relevant means are displayed in Table 3.3.

**Attitudes**

A 2 (self-threat level) x 2 (self-affirmation status) x 2 (argument quality) ANOVA on attitudes toward the health message revealed a significant main effect of argument quality, $F(1, 127) = 8.26, p = .005, \eta_p^2 = .06$. This main effect was qualified by the expected three-way interaction between self-threat level, self-affirmation, and argument quality, $F(1, 127) =$
How Self-Affirmation Affects Explicit Health Information Processing

8.95, \( p = .003, \eta_p^2 = .07 \). Simple effects analysis revealed that in the moderate self-threat condition, self-affirmed participants differentiated between arguments, \( F(1, 127) = 10.37, p = .002, \eta_p^2 = .08 \), whereas non-affirmed participants did not, \( F < 1 \). In contrast, in the high self-threat condition, self-affirmed participants did not differentiate between strong and weak arguments, \( F < 1 \), whereas non-affirmed participants did, \( F(1, 127) = 7.12, p = .009, \eta_p^2 = .05 \). Relevant means are displayed in Table 3.3.

Table 3.3. Means (SD’s) of thought-valence and attitudes as a function of self-threat level, self-affirmation, and argument quality (Study 3.3)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Thought-valence</th>
<th>Attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Argument quality</td>
<td>Argument quality</td>
</tr>
<tr>
<td>Moderate self-threat</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Non-affirmed</td>
<td>0.66 (0.27)a</td>
<td>0.73 (0.25)a</td>
</tr>
<tr>
<td>Self-affirmed</td>
<td>0.52 (0.21)a</td>
<td>0.82 (0.21)b</td>
</tr>
<tr>
<td>High self-threat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-affirmed</td>
<td>0.62 (0.29)a</td>
<td>0.82 (0.19)b</td>
</tr>
<tr>
<td>Self-affirmed</td>
<td>0.68 (0.26)a</td>
<td>0.79 (0.25)a</td>
</tr>
</tbody>
</table>

Note. Means with different subscripts (in rows, for each dependent variable) differ significantly from each other \( (p < .05) \).

Mediation

We conducted a series of regression analyses to examine whether the effects of argument quality on attitudes observed among self-affirmed participants in the moderate self-threat condition and non-affirmed participants in the high self-threat condition were mediated by cognitive responses (cf. Baron & Kenny, 1986).

Regression analyses were first performed for self-affirmed participants in the moderate self-threat condition. First, attitudes were regressed on argument quality. Argument quality significantly predicted attitudes \( (B = .50, p = .012) \). Second, thought valence was regressed on
argument quality. Argument quality significantly predicted thought valence ($B = .15$, $p < .001$). Third, attitudes were regressed on argument quality and thought valence. Thought valence significantly predicted attitudes ($B = 3.25$, $p < .001$) and the effect of argument quality became non-significant ($B = .003$, $p = .99$). A subsequent Sobel-test of mediation proved significant ($Z = 3.03$, $p = .002$), suggesting full mediation.

We performed similar regression analyses for non-affirmed participants in the high self-threat condition. First, attitudes were regressed on argument quality. Argument quality significantly predicted attitudes ($B = .42$, $p = .028$). Second, thought valence was regressed on argument quality. Argument quality significantly predicted thought valence ($B = .10$, $p = .019$). Third, attitudes were regressed on argument quality and thought valence. Thought valence significantly predicted attitudes ($B = 2.82$, $p < .001$) and the effect of argument quality became non-significant ($B = .13$, $p = .40$). A subsequent Sobel-test of mediation proved significant ($Z = 2.18$, $p = .029$), again suggesting full mediation.

Discussion

The results of Study 3.3 replicate and extend those of Study 3.2 by showing that self-affirmation may promote or impede extensive processing of threatening information, depending on the level of self-threat. Under moderate self-threat conditions, self-affirmed participants differentiated between strong and weak arguments whereas non-affirmed participants did not, consistent with the predicted increase in the extent of information processing. In contrast, under high self-threat conditions, self-affirmed participants did not differentiate between arguments whereas non-affirmed participants did, consistent with the predicted decrease in the extent of information processing. These findings underscore that self-affirmation decreases extensive, careful information processing for severe threats to the self. Since self-affirmation did not affect participants’ mood, it seems unlikely that mood
accounts for the observed effects on information processing (cf. Briñol et al., 2007; Cohen et al., 2000; Schmeichel & Martens, 2005; Sherman et al., 2000).

**General Discussion**

In the present research, we examined how self-affirmation affects cognitive processing of threatening information. We hypothesized that the effect of self-affirmation on information processing is likely to vary as a function of the level of self-threat. Specifically, we proposed that under moderate self-threat conditions, self-affirmation would increase the extent of information processing, whereas it would decrease the extent of information processing under low and high self-threat conditions. We tested this hypothesis within the domain of health messages. Conceptualizing self-threat level as participants’ vulnerability to a health risk, and varying the health risk across studies, all three studies confirmed this prediction.

Study 3.1 demonstrated that low, moderate, and high self-threat levels are effectively created by employing different operationalizations of vulnerability (i.e., based on measurement vs. experimental manipulation) and provided initial support that self-threat level moderates the impact of self-affirmation on information processing. Study 3.2 provided a more robust test of the effect of self-affirmation on information processing by including a manipulation of argument quality in the health message. Under a moderate self-threat level, self-affirmed participants showed an increase in their sensitivity to the quality of the arguments in the message, whereas in the low and high self-threat conditions, self-affirmed participants did not differentiate between strong and weak arguments. Study 3.3 replicated these findings using an experimental manipulation of self-threat level across conditions. In addition, Study 3.3 ruled mood out as an alternative explanation as self-affirmation did not affect mood in this study.
Our findings have important implications for self-affirmation theory. Recently, different scholars have stressed the importance of examining potential moderating mechanisms to enhance our understanding of the effects of self-affirmation (Sherman & Cohen, 2006). The present research has met this call by demonstrating that level of self-threat moderates the effect of self-affirmation on information processing. Previous work has suggested that self-affirmation generally increases extensive, careful processing of information that poses a threat to self-integrity (e.g., Cohen et al., 2000; Correll et al., 2004; Harris & Napper, 2005; Sherman et al., 2000). However, other research findings indicated that self-affirmation can also decrease processing of threatening information (Koole et al., 1999; Schimel et al., 2004; Schmeichel & Martens, 2005). The present research reconciles this apparent contradiction in the literature, by showing that the effect of self-affirmation on cognitive processing will vary depending on the level of self-threat.

The moderating role of self-threat level observed in the present research is consistent with self-affirmation theory, and also nicely fits the premise that the self continually interacts with the environment (Kuhl, 2000; Kuhl & Koole, 2004; Nowak et al., 2000). The present findings suggest that self-affirmation promotes an information-processing mode that is most adaptive in a given context. When self-affirmation signals to the person that the self-system is able to cope with a threat, it increases extensive, careful information processing. In these conditions, self-affirmation stimulates integration of the threat into the self-system, thus enabling people to cope with a threat without freezing on its implications for self-integrity (cf. Sherman & Cohen, 2006). Conversely, when self-affirmation signals that the self-system is not able to cope with a threat, it decreases the extent of information processing. In these conditions, self-affirmation ironically reinforces concerns over self-integrity. The increased discrepancy resulting from this unfortunate situation brings people back to where they started: on restoring the general integrity of the self. It is important to bear in mind that these opposite
effects of self-affirmation both are a result of people’s strong desire to maintain global self-integrity, which is the main function of self-affirmation (Steele, 1988).

Finally, the current research underscores the pivotal role of level of self-threat in responses to threatening health information. In line with previous research (Das et al., 2003; De Hoog et al., 2005; Kruglanski & Klar, 1985; Weinstein, 1988), processing of health information varied mainly as a function of perceived vulnerability -regardless of whether the health risk pertained to intestinal cancer, RSI, stress, or toxic beauty products. These findings replicate previous studies by showing that when a health risk reaches a certain level of severity, responses mainly vary as a function of perceptions of vulnerability (Kruglanski & Klar, 1985; Weinstein, 1988). Importantly, the present findings suggest an inverted U-shaped relationship between vulnerability and information processing, with strongest defensive processing occurring for moderate threats. We tentatively propose that, under high self-threat conditions, reality constraints override simple rejection of the threatening information, thereby motivating intensive, careful information processing. This proposition is consistent with the idea that motivated reasoning is constrained by reality and rules of inference (Kunda, 1987; Pyszczynski & Greenberg, 1987). Of course, future research is needed to empirically test this account.

**Potential Limitations and Future Directions**

When looking at the motives underlying the effects of self-affirmation, recent research suggests that - when not particularly threatened - self-affirmation prior to exposure to information may increase people’s confidence in their current views, thereby obliterating the motivation to process the information (Briñol et al., 2007). Under moderate self-threat conditions, self-affirmation presumably restored global self-integrity and increased the motivation to extensively, and carefully process the information (e.g., Correll et al., 2004). In
the present study, we suggested that self-affirmation might reinforce concerns over self-integrity under high self-threat conditions, thus increasing feelings of psychological discomfort rather than decreasing them (Galinsky, Stone, & Cooper, 2000). The present research is an initial investigation that did not directly test this account; therefore, future research is needed to test potentially mediating processes.

In addition, our research ruled mood out as an alternative explanation for the observed effects of self-affirmation, consistent with most previous studies (see Sherman & Cohen, 2006; Steele, 1988), showing that mood does not account for the self-affirmation findings. Nevertheless, both self-affirmation and a positive mood seem to operate as a resource that people use to deal with threatening information, and similar effects on persuasion have been demonstrated (e.g., see Das & Fennis, 2008; Raghunathan & Trope, 2002). Furthermore, mood may play an important role in self-regulatory processes (e.g., Aspinwall, 1998; Raghunathan & Trope, 2002) and in the integration of new information into the self-system (Baumann & Kuhl, 2002, 2003; Bolte, Goschke, & Kuhl, 2003). Thus, disentangling the effects of self-affirmation and mood constitutes an important agenda for future studies.

Concluding Remarks

Whereas previous research mainly found positive effects of self-affirmation, our findings show that when facing a severe threat, self-affirmation decreases extensive, careful information processing. Should we consider this as bad or as good news? This issue may be best illustrated by a quote of Friedrich Wilhelm Nietzsche (1865): “If you wish to strive for peace of soul and pleasure, then believe; if you wish to be a devotee of truth, then inquire”. Our findings reflect good news when we want people to strive for peace of soul and pleasure, but can be considered as bad news when we want people to be devotees of truth.