

When Artificial Social Agents Try to Persuade People: The Role of Social Agency on the Occurrence of Psychological Reactance

Maaïke Roubroeks · Jaap Ham · Cees Midden

Accepted: 28 December 2010 / Published online: 8 January 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract In the near future, robotic agents might employ persuasion to influence people's behavior or attitudes, just as human agents do in many situations. People can comply with these requests, but, people can also experience psychological reactance, which may lead to the complete opposite of the proposed behavior. In this study we are interested in the social nature of psychological reactance. Social agency theory proposes that more social cues lead to a more social interaction. We argue that this also holds for psychological reactance. Therefore, we expect a positive relationship between the level of social agency of the source of a persuasive message and the amount of psychological reactance the message arouses. In an online experiment, participants read an advice on how to conserve energy when using a washing machine. The advice was either provided as text-only, as text accompanied by a still picture of a robotic agent, or as text accompanied by a short film clip of the same robotic agent. Confirming our expectations, results indicated that participants experienced more psychological reactance when the advice was accompanied by the still picture or when the advice was accompanied by the short film clip as compared to when the advice was provided as text-only. This indicates that stronger social agency of the messenger can lead to more psychological reactance. Furthermore, our results

confirmed earlier research about the effects of controlling language on psychological reactance. Implications are discussed.

Keywords Persuasive agents · Psychological reactance · Intentionality · Social influence · Energy conservation behavior

1 Introduction

In the near future, robotic agents might employ persuasion to influence people's behavior or attitudes, just as human agents do in many situations. That is, in many day-to-day situations human agents use persuasion to influence other people (e.g., in health campaigns, political campaigns, environmental campaigns), but the effects of persuasion may sometimes be quite different from what was intended. For instance, persuasion can be used through environmental campaigns in which people are persuaded to conserve energy, recycle and to keep the environment clean. Indeed, all the people you interact with are, most of the time, skillful persuaders. Your partner wants you to bring out the garbage, your child keeps whining for that expensive new toy and your parents really need your help in repairing the car. What all these acts of persuasion have in common is that people can experience *psychological reactance* in response to these persuasive appeals.

When people are the subject of persuasion, they are directed to perform a specific behavior or adapt an attitude. In response to persuasive messages, people can comply and adapt their behavior or attitude in the desired direction. However, people might also experience psychological reactance, which could lead to the complete opposite of the intended behavior. People become psychologically reactant

M. Roubroeks · J. Ham · C. Midden
Department of Human-Technology Interaction, Eindhoven
University of Technology, Den Dolech 2, 5600 MB Eindhoven,
The Netherlands

M. Roubroeks
e-mail: m.a.j.roubroeks@tue.nl

J. Ham (✉)
e-mail: j.r.c.ham@tue.nl

C. Midden (✉)
e-mail: c.j.h.midden@tue.nl

when they experience a persuasive message as a threat to their autonomy. Consequently, they feel the urge to defend their autonomy and do the opposite of whatever the persuasive message directed them to do [3]. So, when people are persuaded to exercise—and they experience psychological reactance as a response—they might do the opposite and watch TV instead. Although earlier research assessed reactance in human–human interactions, in which the persuasive social agent was always human, for an overview see [7], one can imagine that experiences of reactance will also occur when people are being persuaded by artificial social agents in human–computer interactions. In the current study, we will investigate whether an artificial social agent can trigger psychological reactance. In the current paper we argue that reactance effects occur when an artificial social agent is perceived to have stronger social agency. Therefore, we will present participants with artificial social agents with different levels of social agency.

1.1 Psychological Reactance

Psychological reactance has been investigated primarily in research instigated by the theory of psychological reactance [3]. According to this theory [3], persuasive attempts may fail if they are experienced as an intrusion on people’s autonomy. As a response to such attempts, people can experience a strong desire to restore this feeling of autonomy [3]. This is in line with self-determination theory [26] that states that one of the basic humans needs is the need for autonomy.

The theory of psychological reactance argues that the following elements are relevant to reactance (also see Fig. 1); (1) a perceived autonomy¹ (e.g., autonomy to choose coffee or tea at a conference), (2) a threat² to that perceived autonomy (e.g., a waitress persuading you to drink coffee), (3) the experience of psychological reactance (i.e., as indicated by feelings of anger and negative cognitions), and (4) a desire to restore the threatened autonomy (e.g., asking for tea instead of coffee). The restoration of autonomy can occur directly, by doing the forbidden act; or indirectly by an increase of attractiveness or liking for the eliminated option [5], by denying the existence of the threat to autonomy, by exercising a similar autonomy to gain a feeling of control, or by derogating the source, described in [8].

Recent work has investigated the effect of message formulation on psychological reactance, e.g. [6, 8, 9, 17, 19, 21–23, 25]. This research indicates that when language becomes more controlling, people are more prone to respond in

¹The original theory of psychological reactance [3] speaks of perceived *freedom* instead of perceived autonomy. However, we deem autonomy a more concrete concept than freedom, and we will use this term in the rest of the paper.

²In this article the term “threat” is always referred to as an intrusion to one’s autonomy to choose.

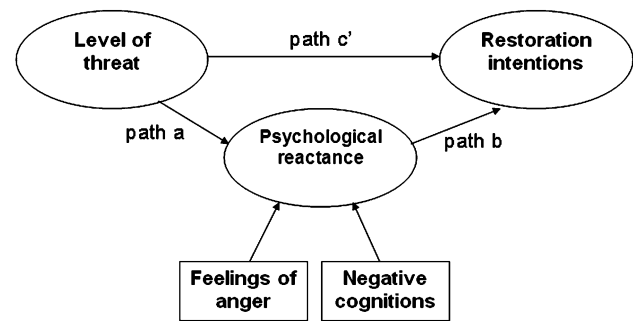


Fig. 1 Psychological reactance model

a reactant way. For example, research by Grandpre and colleagues [6] found that adolescents who were provided with explicit (i.e., containing clear intentions) messages against smoking showed more signs of psychological reactance and initiated more smoking behavior than when these adolescents who were provided with implicit messages (i.e., containing no clear intentions) against smoking. Intriguingly, this effect was also observed for pro-smoking messages; after receiving explicit messages in favor of smoking, participants became more psychologically reactant and engaged in less smoking behavior. These results suggest that persuasive actors should consider their message formulation when conveying a persuasive message.

A study by Miller and colleagues [17] provided insights in the underlying processes leading to psychological reactance. This research indicated that the use of concrete (descriptive) language that avoids any doubts about the meaning of the message, resulted in more compliant behavior. However, when a message contains controlling language (i.e., containing explicit directives) persuasion was diminished and psychological reactance was enhanced. Specifically, Miller and colleagues’ results [17] indicate that high-controlling language makes use of imperatives (i.e., commands and orders) and controlling terms such as “have to”, “must”, “should”, “ought to”, or “need to”. On the other hand, low-controlling language makes use of terms like (i.e., suggestions) “could”, “can”, “may”, “might”, or “could try to” [17]. High-controlling messages are directive and clearly show the persuaders intentions, whereas lowly controlling messages are experienced more as suggestions and as autonomy-supportive [17]. In short, Miller and colleagues [17] showed that when using a (low-controlling) concrete message, persuasion will be enhanced. But when the same message contains high-controlling language, compliance will decrease and such a message might even lead to psychological reactance.

1.2 Psychological Reactance as a Measurable Concept

Psychological reactance can be measured by assessing feelings of anger and negative cognitions. That is, in the original theory of psychological reactance, Brehm and Brehm [4]

explained that psychological reactance cannot be measured directly because it is a motivational state that is the consequence of a threat to autonomy. However, a recent study by Dillard and Shen [8] proposed a different conceptual perspective on psychological reactance to make it measurable. In their study, they present two concepts that could be used as indicators of psychological reactance; anger and negative cognitions. Further research showed that these two factors could be explained best by an intertwined model—the two indicators contributed to psychological reactance in an intermingling way [8, 21, 23]. That is, anger and negative cognitions were best explained as an inseparable construct. The implication of the intertwined model is that one should assess both anger and negative cognitions when measuring psychological reactance. Moreover, research indicated that there exists a relationship between the message that contains a threat-to-autonomy-of-choice and attitude, as well as a relationship between the message that contains a threat-to-autonomy-of-choice and behavioral intentions. That is, results indicated that participants who read a high-threat message had a more negative attitude towards the advocated behavior (e.g., responsible drinking) in the message, and had lower behavioral intentions in line with the advocated behavior (e.g., responsible drinking) than participants who read a low-threat message. In addition, it was found that this relationship was (almost) fully mediated by anger and negative cognitions [8, 21, 23].

1.3 Psychological Reactance and Social Influence

In the current article, we suggest that psychological reactance is a social phenomenon that is the outcome of an interaction between social agents. That is, we argue that psychological reactance occurs primarily when a social agent causes the threat to autonomy of choice. This is in line with Brehm [3], who stated that “...reactance will frequently occur in response to restrictions or threats thereof imposed by social entities, and that the general effect of reactance is to produce tendencies to oppose the actual or threatened restrictions” (p. 387). We believe that when the cause of a threat to one’s autonomy of choice is not another social being but a non-social entity, we argue, no psychological reactance might occur. For example, imagine that while driving your car you come to a point where you can either go right or left and you know that both ways are equally short and you do not have a preference to go one way or the other. But before you make your choice, the road to the right gets blocked by a tree (a non-social entity) that is struck by lightning. In this situation there is a threat to your choice to go to the right. This will most likely cause frustration, but not psychological reactance, in which a desire is experienced to restore that threat to your autonomy of choice. However, when you have the choice to go to the right or left, and a person (a social agent) tries to persuade you to go to the right,

you can experience psychological reactance and maybe even turn left. Thus, in the current research it is proposed that psychological reactance can be attributed to the influence of the social agency of the source of a persuasive message. Consistent with social agency theory [16], we use the term social agency to indicate the degree to which a social agent is perceived as a social entity. In other words, the degree to which a social agent is perceived as being capable of social behavior that resembles human-human interaction, see also [24]. Especially in situations in which humans interact with artificial social agents (e.g., a virtual agent, or a robot), we argue that this issue is relevant. We think that also in these kinds of interactions, the social agency of the source of a persuasive message is important for the occurrence of psychological reactance in response to that message. Support for this hypothesis, can be found in the *Computers are Social Actors paradigm* (CASA paradigm) [24], which states that people react to computers as if they were reacting to other people. The theory behind the CASA paradigm is based on a theory by Langer [12] about mindlessness behavior. When in a state of mindlessness, people act on automatic pilot and do not think consciously about their behavior. Attention is directed to information that is relevant for the current task, leaving no room for alternative information [12]. It is suggested that when people interact with computers, they mindlessly form a premature commitment to social scripts of human-human interaction experienced in the past [18]. In short, the CASA paradigm [24] suggests that social cues trigger social rules and people react in accordance with these social rules. We suggest that this also holds for the social rules of psychological reactance. Therefore, we expect that an artificial social agent can also trigger psychological reactance.

In addition, *Social Agency Theory* [16] and the *Social-Cue Hypothesis* [13] state that the more social cues (e.g., voice, presence of a face, face expressions) are available in an interaction, the more social the interaction becomes. That is, the interaction is experienced to a higher degree as a human-human interaction. Various studies confirm this theory. Additionally, when more social cues are available, people try to understand the relationship with the other actor better [13, 16]. One example that demonstrates this, is a study by Hone [11], which suggests that using an embodied agent (virtual agent & text-bulb) leads to a greater influence of the embodied agent on participants’ attitudes compared to participants who were provided with a text box only. In other words, the virtual agent had a greater influence on the participants, which could be caused by the experience of a more social interaction. Also, research indicates that the visual presence of a high- social agent (i.e., a fully embodied agent) led to a better performance than when this agent was not visually present [15]. Furthermore, research concludes that a more realistic interaction leads to a greater influence of the agent [2, 10] see also [24]. Therefore, we propose that

the more social the source of a message is (i.e., displaying more social cues), the more a persuasive message by that source can give rise to psychological reactance.

1.4 The Current Research

The current study assessed the influence of the social agency of the source of a persuasive message on psychological reactance in response to that message. To study this question, we investigated psychological reactance in response to a persuasive message coming from either an artificial agent depicted on a short film clip; a persuasive message coming from an artificial agent depicted by a still picture; or a persuasive message consisting of text only.

The persuasive message in this study consisted of an advice that persuaded people to conserve energy when using a washing machine. Threat to autonomy was induced by providing participants with either a high-threat advice (e.g., “The washing machine uses a lot of energy, so you really *have to* set the temperature to 40 °C instead of always setting it to 60 °C.”), a low-threat advice (e.g., “The washing machine uses a lot of energy, so you *could consider* to set the temperature to 40 °C now and then instead of 60 °C.”), or a no-threat advice (e.g., “The washing machine uses a lot of energy. *People could* save energy by setting the temperature to 40 °C instead of 60 °C.”). In all conditions we used concrete language to avoid any doubt about the meaning of the advice [17]. We expected that a high-threat advice would lead to more psychological reactance (i.e., negative cognitions and feelings of anger) than a low-threat advice, which in turn would lead to more psychological reactance than a no-threat advice (hypothesis 1). The key hypothesis of the current study, however, investigates whether the more social cues the source of a message expresses (i.e., the more social agency the source of a message has), the more a persuasive message by that source will give rise to psychological reactance. We manipulated social agency by providing participants either with an advice as text-only (low social agency condition) or an advice accompanied by a still picture of the message source (i.e., the artificial agent; medium social agency condition). This manipulation is based on earlier research by Hone [11] that indicated that people are influenced more by a text that is accompanied by a picture than by a text that is not accompanied by a picture. In addition, we added a condition in which even more agency might be ascribed to the social agent. That is, in a third condition we provide participants with an advice accompanied by a moving picture of the message source (high social agency condition). In this condition the mouth of the message source moved as if it was talking. This implies that the message is originated by the message source. Social Agency Theory [16] states that the higher the amount of social cues that can be perceived, the higher the level of social agency that people can ascribe,

meaning that people experience the social agent more as a social entity (i.e., human). Hence, we argue that the additional social cue of facial movement will heighten the level of social agency that people ascribe, which consequently can lead to an increase in psychological reactance. In other words, people experience the interaction more as a human-human interaction when the social agent is capable of moving than when it is a still picture, which thus increases social agency. Therefore, we expected that participants who were provided a persuasive message from a source depicted by a short film clip would experience more reactance than participants who were provided with a persuasive message from a source depicted by a still picture, who in turn would experience more reactance than participants who were provided with a persuasive message consisting of text only (hypothesis 2). Furthermore, we expected an interaction between threat level and social agency level, leading to the strongest effect of psychological reactance when high threat would co-occur with high social agency (hypothesis 3). Finally, in line with previous research [8, 21, 23], we expected a positive relationship between threat to autonomy and restoration intentions, and that this relationship would be mediated by psychological reactance (hypothesis 4).

2 Method

2.1 Participants and Design

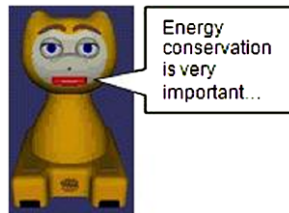
One hundred and thirty-eight participants (70 males, 68 females; age $M = 35.2$, $SD = 15.6$) were randomly assigned to one of nine experimental conditions of a 3 (threat level: no threat vs. low threat vs. high threat) \times 3 (social agency level: low social agency vs. medium social agency vs. high social agency) design. The dependent variable consisted of (the two types of) psychological reactance; a score indicating the amount of feelings of anger and a score indicating the amount of negative cognitions. Participants were mainly recruited on the Internet via a Web link. All participants were native Dutch speakers. The experiment lasted about 15 minutes, for which participants were paid 3 Euros (approximately \$3.75 U.S. at the time this study was conducted), and participants were provided with the opportunity to win an additional price of 100 Euros (approximately \$125 U.S.).

2.2 Materials

The specific text we used in this experiment to give participants advice (the advisory text) described energy conservation while using a washing machine. This advisory text either contained non-controlling language (e.g., “People could save energy when using the washing machine.”), low-controlling language (e.g., “You could save energy when

using the washing machine.”), or high-controlling language (e.g., “You have to save energy when using the washing machine.”). The language use in the advisory texts was based on previous studies about psychological reactance [8, 17, 19]. For some participants, the advisory text was not accompanied by a picture (low social agency). This advisory text was presented as text-only. For other participants, the advisory text was accompanied by a still picture (medium social agency; see Picture 1).

Picture 1 The Medium Social Agency condition



The still picture showed a robotic agent (i.e., the virtual iCat), and this robotic agent had a neutral expression on its face. A speech-bulb connected the robotic agent with the words of the advisory text, implying that the robotic agent was uttering the words. And for the final third of participants, the advisory text was accompanied by a short film clip (high social agency). The short film clip showed the identical robotic agent (i.e., the virtual iCat) without emotional expressions, but now the mouth of the robotic agent made speech movements indicating that it was saying the words.³ (see procedure for further details).

To check the effect of our manipulation of threat level, participants were asked to answer four questions that measured perceived threat to autonomy. This measure was the Dutch translation of a measure of perceived threat to freedom developed by Dillard and Shen [8]. For each question, participants could indicate their agreement to a statement on a five-point Likert scale, ranging from (1) *Completely disagree* to (5) *Completely agree*. These four statements were: “The advice restricted me in my autonomy to choose how I wanted to do the laundry”, “The advice tried to manipulate me”, “The advice tried to make a decision for me”, and “The advice tried to pressure me”. A participant’s mean score on these four statements formed a reliable ($\alpha = 0.90$) measure for perceived threat to autonomy.

To measure feelings of anger (the one part of our measurement of psychological reactance), participants were asked to answer four questions that measured feelings of anger. This measure was the Dutch translation of a measure of anger developed by Dillard and Shen [8]. For each question, participants could indicate their agreement to a statement on a four-point Likert scale ranging from (1) *Not*

at all to (4) *Completely*. These four statements were: “I was irritated”, “I was angry”, “I was annoyed”, and “I was aggravated”. A participant’s mean score on these four statements formed a reliable ($\alpha = 0.83$) measure for feelings of anger.

To measure negative cognitions (the other part of our measurement of psychological reactance), participants were asked to perform a thought-listing task. This task was based on a measure that was used in Dillard and Shen [8] and translated in Dutch. In this task, participants were asked to type all the thoughts they had while reading the advice text, even if those thoughts had nothing to do with the advice text (the amount of thoughts the participants typed ranged from 1 to 10, with $M = 3.2$, $SD = 1.9$). There was no time limit, and participants could type as much as they want in a text box. After listing all thoughts, participants had to indicate whether a thought was either negative (by typing an “N” behind that thought description), positive (by typing a “P” behind that thought description), or neutral (by typing “Neu” behind that thought description). Examples of negative thoughts that participants reported were “I don’t HAVE TO do anything”, and “Unfriendly tone of speaking”. We followed the procedure of Dillard and Shen [8] for removing emotions to minimize overlap between affect and cognition in the reported thoughts. To calculate the “negative cognitions score”, the total amount of negative cognitions was divided by the total amount of all thoughts times 100; this resulted in the percentage of negative cognitions.

To measure restoration intentions, the Reactance Restoration Scale (RRS) [20] was administered. This measure consists of three questions and uses a seven-point continuum, with the following anchor points “motivated–unmotivated”, “determined–not determined”, “encouraged–not encouraged”, and “inspired–not inspired”. The following three items were assessed: “Right now, I am [...] to save energy when doing the laundry” (reversed scored; boomerang restoration), “Right now, I am [...] to be around others that save energy when doing the laundry” (reversed scored; vicarious boomerang restoration), and “Right now, I am [...] to do something totally energy consuming” (related boomerang restoration).⁴ The mean score of the boomerang restoration items formed a reliable ($\alpha = 0.92$) for the boomerang restoration score; the mean score of the vicarious boomerang restoration items formed a reliable ($\alpha = 0.97$) for the vicarious boomerang restoration score; and the mean score of the related boomerang restoration items formed a reliable ($\alpha = 0.97$) for the related boomerang restoration score.

³No actual sound was played, only the mouth moved.

⁴Principal Axis Factoring (PAF) with Oblimin rotation showed that the three constructs were clearly empirically separable from each other, with high factor loadings ranging from 0.78 to 0.97. This is in line with earlier work by Quick and Stephenson [20].

2.3 Procedure

Participants were invited to participate in an online experiment on energy conservation. They were told that the study was about behavior in the household. At the beginning of the experiment, participants were given general introductions and were asked for their consent to proceed with the experiment. After consent, participants were asked to read an advisory text about energy conservation. Participants in the no threat condition received the no threatening advisory text, participants in the low threat condition received the low-threatening advisory text, and participants in the high threat condition received the high-threatening advisory text. For some participants, the advisory text was either provided as text-only (low social agency); for others the advisory text was accompanied by a still picture of a robotic agent (medium social agency; see Picture 1); and for the third part of participants the advisory text was accompanied by a short film clip, in which the same robotic agent moved its mouth as if it was talking (high social agency).

After reading the advisory text, participants were asked to perform the thought-listing task, and answer the questions that measured feelings of anger, perceived threat to autonomy, and restoration intentions (see materials section).

Finally, participants were thanked for their participation and paid, and were reminded about the opportunity to win 100 Euros.

3 Results

3.1 Manipulation Check

First, we checked whether our manipulation of threat level was successful. To test this, the perceived threat of autonomy score was submitted to a Univariate Analysis of Variance (ANOVA), with threat level as the independent variable. Confirming our hypothesis, this analysis showed a significant effect of threat level, $F(2, 135) = 17.31$, $p < 0.01$. Pairwise comparison analyses indicated that participants in the high threat condition reported more perceived threat to autonomy ($M = 3.44$, $SD = 1.21$) than participants in the low threat condition ($M = 2.20$, $SD = 1.04$), or no threat condition ($M = 2.28$, $SD = 1.00$), both p 's < 0.01 . However, these analyses did not indicate that participants in the no threat condition reported a different perceived threat to autonomy from participants in the low threat condition, *n.s.* A possible explanation for this might be that the participants may have experienced even the no-threat advisory text as slightly threatening, because in this condition messages containing general advice were also included in the text.

Second, we checked whether our manipulation of social agency led to differences on a measure of experienced threat.

That is, to be able to ascribed effects of the manipulation of agency to effects of social agency only, it is important that this manipulation does not give rise to differences in experienced threat. To test this, the perceived threat of autonomy score was submitted to a Univariate Analysis of Variance (ANOVA), with social agency level as the independent variable. Confirming our expectations, this analysis showed that the manipulation of social agency did not lead to differences in the amount of experienced threat, *n.s.* This result suggests that participants did not experience more threat when exposed to a text accompanied by a social agent compared to a text alone, indicating that the agent did not cause any threat by itself.

Third, we investigated whether there were any gender differences in the current study. To test this, we submitted the negative cognitions score and feelings of anger score to a 3 (threat level: no threat vs. medium threat vs. high threat) \times 3 (social agency level: low social agency vs. medium social agency vs. high social agency) \times 2 (gender: male vs. female) Multivariate Analysis of Variance (MANOVA). Results indicate that gender did not have a significant effect on our two types of psychological reactance measures (i.e., negative cognitions and feelings of anger) nor did results indicate any interaction effects with gender, all effects *n.s.* However, results did indicate a significant Threat level \times Social agency level \times Gender interaction, $F(1, 4) = 2, 91$, $p < 0.05$, but in this analysis the amount of participants per cell was too small (some cells containing less than 3 participants) to draw firm conclusions about this result.

3.2 Threat

Our first hypothesis states that participants who received a high-threat advisory text would report more feelings of psychological reactance than participants who received a low-threat advisory text, who in turn would report more feelings of psychological reactance than participants who received a no-threat advisory text. The negative cognitions score and the feelings of anger score were submitted to a 3 (threat level: no threat vs. low threat vs. high threat) \times 3 (social agency level: low social agency vs. medium social agency vs. high social agency) Multivariate Analysis of Variance (MANOVA). We treated the two types of psychological reactance as a repeated measures factor. In line with previous research, e.g., [8], we expected that participants in the high-threat condition would have a higher psychological reactance score than participants in the low-threat condition, and participants in the low-threat condition would have a higher psychological reactance score than participants in the no-threat condition. Confirming our expectations, results showed that there was a significant main effect of threat level, $F(2, 129) = 13.81$, $p < 0.01$ (see Table 1 for the mean scores and standard deviations).

Table 1 Mean scores psychological reactance score (and standard deviations between brackets) in all conditions for the complete MANOVA design

Threat	Social agency		
	Low social agency	Medium social agency	High social agency
No threat	8.5 (11.4)	13.9 (13.9)	13.3 (17.9)
Low threat	9.7 (13.7)	15.9 (16.4)	22.4 (16.1)
High threat	24.0 (14.6)	31.0 (18.7)	31.6 (6.9)

Table 2 Means scores on the negative cognitions score and feelings of anger score (and standard deviations between brackets) for the threat manipulation

Psychological reactance	Threat		
	No threat	Low threat	High threat
Negative cognitions	21.5 (26.7) _a	32.3 (32.2) _a	54.0 (28.8) _b
Anger	1.4 (0.5) _a	1.4 (0.5) _a	2.1 (0.7) _b

Note: Mean scores in rows that do not share the same subscript differ significantly, $p < 0.05$

The pairwise comparisons for the main effect of threat level, corrected using a Bonferroni adjustments, indicated that participants in the high threat condition reported more feelings of psychological reactance ($M = 28.0$, $SD = 14.6$) than either participants in the low-threat condition ($M = 16.9$, $SD = 16.2$), $p < 0.001$, or participants in the no-threat condition ($M = 11.5$, $SD = 13.6$), $p < 0.001$. However, there was no significant difference between the no-threat condition and the low-threat condition, *n.s.* A contrast analysis demonstrated a linear trend of threat level, suggesting that participants in the high-threat condition experienced more psychological reactance ($M = 28.9$, $SD = 13.6$) than participants in the low threat condition ($M = 16.0$, $SD = 15.5$), who again reported more feelings of psychological reactance than participants in the no-threat condition ($M = 11.9$, $SD = 14.5$), $F(1, 136) = 25.86$, $p < 0.01$. Furthermore, the effect of threat level was qualified by an interaction of threat level X type of psychological reactance, $F(2, 129) = 13.1$, $p < 0.01$, indicating that the effect of threat level is different on the measure of feelings of anger than on the measure of negative cognitions. Closer examination of the separate effects on feelings of anger and negative cognitions showed that the effect of threat level was significant for both measures, both p 's < 0.01 , but, there was a stronger effect of threat level on the measure of feelings of anger, $F(2, 129) = 19.23$, partial $\eta^2 = 0.23$, than of threat level on the measure of negative cognitions, $F(2, 129) = 13.47$, partial $\eta^2 = 0.17$ (see Table 2). Although the correlation between negative cognitions and feelings of anger was quite high ($r = 0.61$, $p < 0.001$) we still found a different effect on the two types of psychological reactance measures. That is, it seems that the effects are not completely inseparable. This result was a bit surprising, because previous re-

search by Dillard and Shen [8] suggests that the effects of their distinct effects cannot be disentangled. However, these authors do note in their discussion that this does not always have to be the case, because "... *cognition and affect are phenomena of rapid change*" [8, p. 160].

3.3 Social Agency

In our second hypothesis, we predicted that high social agency would lead to more psychological reactance than medium social agency, which in turn would lead to more psychological reactance than low social agency. The MANOVA analysis reported above was investigated for the main effect of social agency level. Confirming our hypothesis, results demonstrated that there was indeed a main effect of social agency level on psychological reactance, $F(2, 129) = 3.66$, $p = 0.03$ (see Table 1 for mean scores and standard deviations). The pairwise comparisons for the main effect of social agency level, corrected using a Bonferroni adjustments, indicated that the participants in the low social agency condition experienced less psychological reactance (text-only; $M = 14.1$, $SD = 14.7$) than participants in the high social agency condition (text-short film clip; $M = 22.9$, $SD = 15.7$), $p < 0.05$. However, there were no significant differences between the medium social agency condition and the high social agency condition, *n.s.*, nor between the low social agency condition and the medium social agency condition, *n.s.* A contrast analysis demonstrated a linear trend of social agency level, suggesting that participants in the high social agency condition experienced more psychological reactance ($M = 22.4$, $SD = 13.8$) than participants in the medium social agency condition ($M = 20.2$, $SD = 16.5$), who again reported more

Table 3 Means scores on the negative cognitions score and the feelings of anger score (and standard deviations between brackets) for the social agency manipulation

Psychological reactance	Social agency		
	Low social agency	Medium social agency	High social agency
Negative cognitions	26.6 (29.1) _a	35.1 (33.8) _b	44.1 (31.1) _b
Anger	1.5 (0.6) _a	1.6 (0.7) _a	1.7 (0.6) _a

Note: Mean scores in rows that do not share the same subscript differ significantly, $p < 0.05$

feelings of psychological reactance than participants in the low social agency condition ($M = 14.1$, $SD = 13.3$), $F(1, 136) = 6.64$, $p = 0.01$. Furthermore, the effect of social agency was—just like the effect of threat on the two types of psychological reactance measures—also qualified by an interaction of social agency level X type of psychological reactance $F(2, 129) = 3.66$, $p < 0.05$. Closer examination of this interaction indicated that social agency has an effect on negative cognitions ($F(2, 129) = 3.66$, $p < 0.05$), but not on feelings of anger, *n.s.* More specifically, participants in the low social agency experienced less negative cognitions ($M = 26.6$, $SD = 29.1$), $p = 0.05$, than either the participants in the medium social agency ($M = 35.1$), or the high social agency condition ($M = 44.1$), $p = < 0.05$. Again, there were no significant differences between the medium social agency condition and the high social agency condition, *n.s.* (see Table 3).

3.4 Threat X Social Agency

Furthermore, we expected that there would be a significant interaction effect of Threat level X Social agency level, with the highest psychological reactance when a high threat co-occurred with a high social agency. However, in the MANOVA analysis reported above, the expected interaction effect between threat level and social agency level was not found, *n.s.* A possible explanation for this null effect could be that our manipulation of social agency was not powerful enough to elicit an additional effect on psychological reactance.

3.5 Restoration

In accordance with previous research about psychological reactance, we expected a positive relationship between perceived threat of autonomy and restoration intentions, and that this relationship would be mediated by psychological reactance (see Fig. 1). We performed three separate analyses for each restoration construct (i.e., boomerang, vicarious boomerang, and related boomerang). However, we found no relationship between threat and boomerang restoration, *n.s.*, and no relationship between threat and related boomerang

restoration, *n.s.* Therefore, we only investigated the mediation effect of psychological reactance on the relationship between threat and vicarious boomerang restoration.

To examine this, we followed the steps of mediation analysis developed by Baron and Kenny [1]. First, a linear regression analysis with threat level as the independent variable and vicarious boomerang restoration as the dependent variable showed a positive relationship between threat and vicarious boomerang restoration ($B = 0.35$, $SD = 0.17$), $t(137) = 2.08$, $p < 0.05$, $R^2 = 0.03$ (path *c*). Step 2 of the analysis is to check for a positive relationship between threat and psychological reactance (i.e., the combined score of feelings of anger and negative cognitions). Results showed that this was indeed the case ($B = 8.15$, $SD = 1.60$), $t(137) = 5.09$, $p < 0.001$, $R^2 = 0.16$ (path *a*). Subsequently, we checked whether the mediator affected the outcome. Results suggest that when entering vicarious boomerang restoration as the dependent variable, and entering psychological reactance and threat as the two independent variables, the outcome of psychological reactance on vicarious boomerang restoration was affected ($B = 0.02$, $SD = 0.01$), $t(137) = 2.47$, $p < 0.05$, $R^2 = 0.07$ (path *b*). This analysis further showed that the effect of threat on vicarious boomerang restoration became non-significant ($B = 0.17$, $SD = 0.18$), $t(137) = 0.96$, $p > 0.1$, $R^2 = 0.07$ (path *c'*). Because the effect of threat on vicarious boomerang restoration became almost zero (i.e., $B = 0.17$), this suggests a full mediation of psychological reactance on the relationship between threat and restoration. A Sobel test confirmed this suggestion, $z = 2.20$, $p < 0.05$. In short, it seems that psychological reactance served as a mediator in the relationship between threat to autonomy and vicarious boomerang restoration.

4 Discussion

The current research investigated the influence of the social agency of the source of a persuasive message on psychological reactance in response to that message. We expected that the more social the source of a message is (i.e., displaying more social cues), the more a persuasive message by that source will give rise to psychological reactance. In an online

study, participants read an advice about how to save energy when doing the laundry. This advice was either provided as text-only, as text accompanied by a still picture of a robotic agent, or as text accompanied by a short film clip of the same robotic agent. Psychological reactance was measured by assessing negative cognitions and feelings of anger. Confirming our expectations, results indicated that participants who read an advice that was accompanied by a short film clip showing a moving robotic agent experienced more psychological reactance than participants who read an advice that was accompanied by a still picture showing the same robotic agent, who in turn experienced more psychological reactance than participants that read an advice that was text-only. These results indicate that stronger social agency of an artificial agent can lead to more psychological reactance in response to the persuasive message of that artificial agent.

So, intriguingly, even though our manipulation of social agency was minimal (text-only vs. text with still picture vs. text with short film clip), results indicate a significant difference of social agency on reported psychological reactance. Especially the difference between the medium and high social agency condition was minimal: the only difference between these conditions was that in the high social agency condition the agent moved its mouth as if it was talking. Furthermore, we used a social agent that was non-human. Perhaps a stronger social agency manipulation with even clearer social cues (e.g., facial expressions, voice, more human-like agent, gestures) would result in even higher amounts of psychological reactance.

Moreover, we found that the effect of social agency was only pronounced on one of the two reactance measure—negative cognitions—but not on the measure of feelings of anger. A possible explanation for this result could be that our manipulation of social agency was not strong enough to cause detectable effects on the feelings of anger measure.

Furthermore, results of the current study suggest that participants who received a high-threat advice about energy conservation showed more evidence of experiencing psychological reactance than participants who received low-threat advice or participants who received no-threat advice. Different from the effect of social agency on the psychological reactance measures, which was only found on the negative cognitions measure, the effect of threat to autonomy of the language in the advice was found both on the measure for negative cognitions and the measure for feelings of anger. This finding is in correspondence with earlier research [8, 17, 19, 21, 23] that indicates that the use of threatening language (threatening people's autonomy) and controlling language (using explicit directives) will diminish persuasion and enhance psychological reactance. Moreover, this finding was especially pronounced on the reactance measure of feelings of anger, but also present on the reactance measure of negative cognitions. This difference

is also found in a recent other study, which also showed stronger effects related to psychological reactance on measures of anger than on measures of negative cognitions [19].

The manipulation of the amount of control (high-controlling vs. low-controlling vs. non-controlling), that is, threat to autonomy, used in the current study did not seem to have an influence on this relationship. Specifically, we did not find the expected interaction between the amount of threat and the level of social agency leading to the strongest reactance when high threat co-occurred with high social agency. Interestingly, a recent study did find some evidence that psychological reactance was aroused especially when a persuasive agent was combined with explicitness [14].

Finally, in confirmation of our fourth hypothesis, results indicated that the use of more controlling language led to a higher amount of reported restoration intentions, and this relationship was partially mediated by psychological reactance. However, this mediation effect of psychological reactance on the relationship between threat and restoration was only found for the vicarious boomerang restoration. Future research should investigate in which situation what type of restoration is triggered. This finding is in correspondence with previous research about psychological reactance [6], for an overview see [7–9, 17, 19, 21–23].

4.1 Implications and Future Directions

Overall, these findings have several implications for both the design of persuasive messages, as for the usage and design of an artificial agent that delivers these messages. If we would like to develop persuasive technology to stimulate people to diet, to stop smoking, or in the current research, to conserve energy, it is important to consider the possibility of psychological reactance. The current study replicates earlier research e.g., [6, 9, 17] in showing that psychological reactance can be avoided by formulating the message in such a way that it is not experienced as a directive. The current study adds to earlier research that the message deliverer has to be taken into account. That is, the current results suggest that when the image that is sketched of the communicator of a persuasive message consists of more social cues, more psychological reactance will occur in response to that message. Another possible explanation for the effect of social agency on psychological reactance could be that the presence of a robotic agent caused greater reactance than messages without a robot due to vividness effects. Quick and Stephenson [19] showed that the vividness of language use (i.e., words that lead to better visualization of consequences of behavior) can be perceived as threatening. For example, Quick and Stephenson used messages containing the following language use: "... *exercising saves you from dying of a massive heart attack, clogged arteries, morbid obesity, and*

being stressed out of your mind” (p. 474). However, as reported in the Results section, analyses suggest that our artificial social agent was not seen as more threatening than the text-only condition. Still, we did not directly measure whether the artificial social agent was seen as more vivid than the text-only condition, and therefore no strong conclusions can be drawn. Future research can further investigate the interesting issue of vividness effects. Thereby, the current research starts the investigation of the role of social agency for psychological reactance.

The current study extends psychological reactance theory [3, 4] by linking psychological reactance to the concept of social agency. To our knowledge, the current study is the first to make this connection. Thereby, it opens up many possibilities for future research. For example, other and perhaps stronger manipulations of social agency and threat of messages might be investigated. That is, perhaps people might easily experience psychological reactance when a human social agent threatens their autonomy to choose, but people might even more quickly experience psychological reactance to even the slightest controlling message coming from a non-human social agent. To answer that question future research could investigate what happens when also a human messenger is put into the comparison. Future research might also investigate our claim that psychological reactance predominantly occurs when the threat comes from a social agent and not when the threat comes from a non-social entity (e.g., a tree blocking the road). That is, although in our text-only condition social agency may have been low, future research could compare the effect of a persuasive message coming from a social agent to a limitation of people’s autonomy coming from a completely non-social source. Finally, future research could examine whether the occurrence of psychological reactance perseveres or even accumulates when people interact with an artificial social agent in a task that is more interactive in nature.

To come back to the psychological reactance you might experience when someone wants to persuade you: The current research indicates that it makes a difference which communicator brings the message. If—somewhere in the near future—this messenger happens to be a robotic agent or a robot, the current research indicates that you are prone to experience more psychological reactance (indicated by feelings of anger and negative cognitions) when this robotic agent or robot has a higher social agency.

Acknowledgements We would like to thank all the prior and current members of the Persuasive Technology Lab Group, and in particular Suzanne Vossen, Ad van Berlo, Matijn van der Made, Eric Postma, and Wilco Moerman for their insights and contributions to this research.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *J Pers Soc Psychol* 51:1173–1182
2. Blascovich J (2002) A theoretical model of social influence for increasing the utility of collaborative virtual environments. In: Proceedings of the 4th international conference on collaborative virtual environments, Bonn, Germany, 2000. Association for Computing Machinery, New York, pp 25–30
3. Brehm JW (1966) A theory of psychological reactance. Academic Press, New York
4. Brehm SS, Brehm JW (1981) Psychological reactance: a theory of freedom and control. Academic Press, New York
5. Brehm JW, Stires LK, Sensenig J, Shaban J (1966) The attractiveness of an eliminated choice alternative. *J Exp Soc Psychol* 2:301–313
6. Buller DB, Borland R, Burgoon M (1998) Impact of behavioral intention on effectiveness of message features: evidence from the family sun safety project. *Hum Commun Res* 24(3):433–453
7. Burgoon M, Alvaro E, Grandpre J, Voulodakis M (2002) Revisiting the theory of psychological reactance: Communicating threats to attitudinal freedom. In: Dillard J, Pfau M (eds) The persuasion handbook: theory and practice. Sage Publications, Thousand Oaks
8. Dillard JP, Shen L (2005) On the nature of reactance and its role in persuasive health communication. *Commun Monogr* 72:144–168
9. Grandpre J, Alvaro EM, Burgoon M, Miller CH, Hall JR (2003) Adolescent reactance and anti-smoking campaigns: a theoretical approach. *Health Commun* 15(3):349–366
10. Guadagno RE, Blascovich J, Bailenson JN, McCall C (2007) Virtual humans and persuasion: the effect of agency and behavioral realism. *Media Psychol* 10:1–22
11. Hone K (2006) Empathic agents to reduce user frustration: the effects of varying agent characteristics. *Interact Comput* 18:227–245
12. Langer EJ (1992) Matters of mind: mindfulness/mindlessness in perspective. *Conscious Cogn* 1:289–305
13. Louwse MM, Graesser AC, Lu S, Mitchell HH (2005) Social cues in animated conversational agents. *Appl Cogn Psychol* 19:693–704
14. Liu S, Helfenstein S, Wahlstedt A (2008) Social psychology of persuasion applied to human-agent interaction. *Hum Technol* 4(2):123–143
15. Lusk MM, Atkinson RK (2007) Animated pedagogical agents: does their degree of embodiment impact learning from static animated worked examples? *Appl Cogn Psychol* 21:747–764
16. Mayer RE, Sobko K, Mautone PD (2003) Social cues in multimedia learning: role of speaker’s voice. *J Educ Psychol* 95:419–425
17. Miller CH, Lane LT, Deatrick LM, Young AM, Potts KA (2007) Psychological reactance and promotional health messages: the effects of controlling language, lexical concreteness, and the restoration of freedom. *Hum Commun Res* 33:219–240
18. Nass C, Moon Y (2000) Machines and mindlessness: social responses to computers. *J Soc Issues* 56(1):81–103
19. Quick BL, Considine JR (2008) Examining the use of forceful language when designing exercise persuasive messages for adults: a test of conceptualizing reactance arousal as a two-step process. *Health Commun* 23:483–491
20. Quick BL, Stephenson MT (2007a) The reactance restoration scale (RRS): a measure of direct and indirect restoration. *Commun Res Rep* 24(2):131–138
21. Quick BL, Stephenson MT (2007b) Further evidence that psychological reactance can be modeled as a combination of anger and negative cognitions. *Commun Res* 34(3):255–276

22. Quick BL, Stephenson MT (2008) examining the role of trait reactance and sensation seeking on perceived threat, state reactance, and reactance restoration. *Hum Commun Res* 34:448–476
23. Rains SA, Turner MM (2007) Psychological reactance and persuasive health communication: a test and extension of the intertwined model. *Hum Commun Res* 33:241–269
24. Reeves B, Nass C (2002) *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press, New York
25. Reinhart AM, Marshall HM, Feeley TH, Tutzauer F (2007) The persuasive effects of message framing in organ donation: the mediating role of psychological reactance. *Commun Monogr* 74(2):229–255
26. Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development and well-being. *Am Psychol* 55(1):68–78

Maaïke Roubroeks is a Ph.D. Student at the University of Technology in Eindhoven, in the Department of Human-Technology Interaction. She received her master's degree in Experimental Health Psychology from Maastricht University. Her current research interests are persuasive technology, social psychology, implicit social cognition, and automatic social behavior.

Jaap Ham is an Assistant Professor at the University of Technology in Eindhoven, in the Department of Human-Technology Interaction. He got his Ph.D. at the University of Nijmegen specializing in the social-cognitive processes of person perception. His current research interests are persuasive technology, ambient intelligence, unconscious influencing and social robotics.

Cees Midden is a Professor at the University of Technology in Eindhoven, and chair of the Human-Technology Interaction group. Present research interests include persuasive technology, risk perception, and trust.