Taking Back Control: Findings on the Cognitive, Behavioral, and Motivational Consequences of Victim Sensitivity

Merle Buchholz[1] · Mario Gollwitzer[2] · Zoe Magraw-Mickelson[2] · Christopher Stolz[1] · Philipp Süssenbach[3]

Abstract

Being duped is an aversive experience which people are motivated to avoid. For this reason, especially people with a high fear of exploitation (i.e., people high in victim sensitivity; VS) often behave pre-emptively selfish and defensive in socially uncertain situations. Because the cognitive and motivational processes underlying such defensiveness have received little attention so far, we conducted two studies aiming to close this research gap. In Study 1 ($n = 84$), we used virtual reality technology to examine whether social distancing, hostile interpretations of an interaction partner's intentions and behavior, and legitimizing cognitions regarding own selfish reactions (as elements of a suspicious mindset) mediate the effect of VS on uncooperativeness. Results did not show the expected mediation, but VS was still related to hostile information processing and fear of exploitation. In Study 2 ($n = 273$), we extended these findings by showing that defensive reactions of people high in VS can be attenuated if a sense of control is reinforced. Together, the two studies crucially expand our knowledge of the defensive motivational system in victim-sensitive individuals.

Keywords Fear of exploitation · Victim sensitivity · Suspicious mindset · Virtual reality · Need for control

✉ Mario Gollwitzer
mario.gollwitzer@lmu.de

[1] Department of Psychology, Philipps-Universität Marburg, Marburg, Germany

[2] Department of Psychology, Ludwig-Maximilians-Universität München, Leopoldstrasse 13, 80802 Munich, Germany

[3] Department for Human Resources/Health/Social Sciences, Fachhochschule des Mittelstands, Bielefeld, Germany

Introduction

The perception of having been taken advantage of is painful and aversive (Vohs et al., 2007). Not only does being "duped" result in a disadvantageous outcome for the victimized person, but it also implies that one has been a fool for not having seen it coming. For this reason, the feeling of being duped is often accompanied by self-blame and other self-conscious emotions like shame, embarrassment, and regret. Consequently, people are strongly motivated to avoid feeling duped, and when they anticipate or expect exploitation, they often behave in a pre-emptively defensive way.

For example, such pre-emptive defensiveness or pre-emptive selfishness has been discussed as one explanation for social loafing, which describes the phenomenon that people tend to decrease effort when working in groups compared to when working alone (Latané et al., 1979). One reason why people loaf in groups appears to be that they expect others to do the same: They work less hard because they are afraid that their group members would otherwise "free ride" on their efforts. Stated differently, the feeling of being duped is so aversive that one would rather slack off and put up with a worse outcome than being the "sucker" who carries the rest of the group (Kerr, 1983).

Importantly, there are stable inter-individual differences in the motivation to avoid exploitation. While some people hate the idea of falling prey to other people's malicious intentions, others simply care less about potentially being duped. The personality trait capturing such a latent fear of exploitation is called "victim sensitivity" (VS). VS has been shown to predict uncooperative and even immoral behavior in socially uncertain interdependence situations, when the risk of being duped is (considered to be) high. For instance, Gerlach et al. (2012) found victim-sensitive people to be less forgiving and more revengeful following a relational transgression committed by a romantic partner or close friend. Similarly, Faccenda and colleagues (2009) showed that after being fouled, soccer players high in victim sensitivity are more inclined to engage in unsportsmanlike, transgressive behaviors themselves. Additionally, victim-sensitive individuals also tend to make more egoistic choices in social dilemma games and other enticing situations (Fetchenhauer & Huang, 2004; Gollwitzer & Rothmund, 2011; Gollwitzer et al., 2005, 2009; Maltese et al., 2016; Rothmund et al., 2011) and more likely engage in counterproductive work behavior when feeling treated unfairly in organizational contexts (Lavelle et al., 2018).

A theoretical explanation for these findings is provided by the Sensitivity to Mean Intentions (SeMI) Model (Gollwitzer & Rothmund, 2009; Gollwitzer et al., 2013). According to this model, victim-sensitive individuals become more easily suspicious and in turn often behave selfishly and uncooperatively. More specifically, people high in VS are assumed to be particularly sensitive toward contextual cues indicating untrustworthiness. When such cues are present, a suspicious mindset (consisting of an attributional bias regarding others' malevolence, legitimizing cognitions regarding one's own selfish behaviors, and an avoidance-related motivational state) is activated in victim-sensitive people, and in this state

of suspiciousness, there is less adherence to social norms and fairness principles. Thus, the suspicious mindset is theoretically conceptualized as the mechanism explaining the effect of VS on uncooperative behavior. In technical terms, the SeMI model describes a moderated mediation: When cues of untrustworthiness are present ("moderator"), VS activates a suspicious mindset ("mediator"), which then results in the withdrawal of trust and cooperation in social exchanges.

Empirical Evidence for the Suspicious Mindset

In situations that require trust, victim-sensitive individuals' strong fear of exploitation leads them to react defensively and uncooperatively in order to protect themselves from possible victimization. In other words, victim-sensitive people are strongly motivated to avoid being duped. While a number of studies have looked at the "total effect" of VS (in the presence of untrustworthiness cues) on uncooperative behavior, the precise cognitive and motivational processes underlying such defensiveness have received less attention. As a consequence, evidence for the suspicious mindset and its three components (i.e., hostile interpretations, legitimizing cognitions, avoidance motivation) has been so far indirect.

The few studies that investigated the suspicious mindset have mostly used scenario-based approaches. For example, Gerlach et al., (2012, Study 3) applied the SeMI model to research on interpersonal forgiveness and instructed their participants to imagine that a close friend had harmed them (for instance, by disclosing a secret). After imagining this scenario, participants were asked how they interpreted the friend's intentions, that is, whether they thought the friend's post-transgression behavior was indicative of ulterior motives (attributional bias), and whether or not unforgiving reactions on their behalf would be justified (legitimization). In line with the SeMI model's predictions, victim sensitivity was negatively related to participants' willingness to forgive, and this relationship was (partially) mediated by hostile interpretations and self-protective legitimizing cognitions.

Following these findings, Maltese et al., (2016; Study 1) used ambiguous scenarios to show that heightened expectancies of injustice can (in part) account for victim-sensitive people's uncooperative behavior in trust games. As expected by the authors, participants high in VS were more inclined to endorse and anticipate unjust outcomes when reading the scenarios (at least when a suspicious mindset had been activated before), and these anticipations mediated the negative effect of VS on trust and cooperation.

Lastly, Rothmund et al. (2011) were able to demonstrate that experiencing virtual aggression from a victim's perspective resulted in reduced trust expectations as well. In their study, participants first played a video game sequence in which they were either betrayed by a video game character or not. Afterward, they took part in a common goods dilemma game in the real world. Here, participants had to indicate (1) how much they wanted to contribute, and (2) how much they expected other players to invest in this situation. Again, a significant indirect effect was found: When confronted with virtual untrustworthiness, participants were more reluctant to cooperate

in the subsequent dilemma game, and this was mediated by heightened mistrust expectancies with regard to others' investments.

In sum, previous research has attempted to elucidate the suspicious mindset and its components. Although these studies have shed some light on the cognitive dynamics of people high in VS, they have focused mainly on one element (namely the attributional bias), oftentimes by using scenario-based approaches with low ecological validity. Therefore, more research is needed in this context, which looks for stronger and more direct support regarding all three elements of the suspicious mindset. This is what Study 1 of the present research aims to do. In Study 2, we go a step further by illuminating the motivational processes operating behind the self-protective concerns of victim-sensitive people.

Motivational Processes Underlying the Pre-Emptive Defensiveness

What exactly motivates people to avoid exploitation and to react pre-emptively selfish when the risk for being duped is considered to be high? One of the few considerations of this question is given by Vohs et al. (2007). As briefly discussed above, these authors argue that people are motivated to avoid exploitation because the feeling of being duped is an aversive emotional state related to self-blame and other negative self-conscious emotions. More specifically, Vohs et al. (2007) posit that feeling duped arises from the perception of having been taken advantage of in a social context after one *chose to trust* the other person. Therefore, the victim is partly to blame for its situation: The unfavorable outcome could have been avoided if a different decision had been made (i.e., not trusting the other party), and for this reason, the experience of victimization results in self-recrimination. Stated differently, being deceived by another person has threatening implications for the self; it suggests that one is socially incompetent or stupid. Because people are reluctant to adopt such negative self-views, they will hence be motivated to avoid being the "sucker." Thus, according to Vohs et al. (2007), individuals with a high fear of exploitation (i.e., people high in VS) are defensive because the anticipation of being exploited threatens their positive self-image.

Challenging this theorizing, we argue that victim-sensitive people's defensiveness is not so much about avoiding self-image threats, but more about maintaining a *sense of control*. As a vast amount of research shows, people generally prefer to have (perceived) control over themselves, their environment, and the outcome of social situations (e.g., Fiske, 2003; Kay et al., 2008; Rothbaum et al., 1982; Skinner, 1995, 1996). Victim-sensitive individuals, however, might be *particularly* motivated to be in control: As Schmitt et al. (1995) demonstrated, a higher sensitivity to own unjust disadvantages was associated with a higher need for control. In our view, socially uncertain interdependence situations are particularly threatening to this need for control. In these situations, untrustworthiness cues act as a warning signal for vulnerability and victimization—an aversive outcome imposed by others. The individual should hence feel at the mercy of untrustworthy forces and, as a consequence, should be motivated to (re)establish a sense of control by preventing the expected exploitation (e.g., Rothbaum et al.,

1982; Skinner, 1996). Thus, we argue that the defensiveness shown by people high in VS reflects a strategy to cope with the anticipated loss of control over the situation.

Study 2 was specifically designed to test these two competing hypotheses (i.e., anticipated exploitation threatens victim-sensitive people's positive self-image vs. their need for control) directly against each other. Notably, if we were able to explain why people high in VS tend to react with self-protective and uncooperative tendencies in socially uncertain situations, then we could design strategies in order to attenuate these reactions. Understanding the motivation behind victim-sensitive individuals' defensiveness becomes especially relevant when considering that a high fear of exploitation crucially impairs social interactions.

The Present Research

The aim of the present research is to further illuminate the defensiveness mechanism underlying victim sensitivity. More specifically, we pursue two goals: In Study 1, we simultaneously investigate all three elements of the suspicious mindset by using immersive virtual reality technology. A virtual environment has the advantage that it allows to study participants' behavior in an ecologically realistic way while maintaining a high degree of experimental control (McCall & Blascovich, 2009). Thus, Study 1 goes far beyond what previous research on the suspicious mindset has achieved. In addition, we introduce social distancing as a behavioral measure of the mostly neglected third component of the mindset, avoidance motivation. Accordingly, two hypotheses are being tested in this study: First, we expect to replicate the established effect of VS on trust and cooperation. Second, we hypothesize that this relation is mediated by hostile interpretations, legitimizing cognitions, and participants' approach-avoidance behavior in a virtual world.

In Study 2, we go a step further by exploring why people high in VS actually show such pre-emptive defensiveness. As discussed above, two hypotheses are plausible in this context: On the one hand, it may be that victim-sensitive individuals are defensive because they are afraid that being duped threatens their self-image (Vohs et al., 2007). On the other hand, persons high in VS might be defensive because they want to maintain a sense of control over what is happening in their social environment.

Study 1

Study 1 was designed to investigate the suspicious mindset in a more immersive and direct way than has been done in past research. More specifically, the focus of Study 1 was on participants' avoidance motivation; a component of the suspicious mindset that has been neglected in previous studies. As a behavioral measure for this avoidance-related motivational state, social distancing was used. We expected that in a state of suspiciousness, individuals high in VS would (physically) distance themselves from others because in trust-relevant situations, any interaction partner represents a source of potential exploitation (Gollwitzer & Rothmund, 2009; Gollwitzer et al., 2013).

To activate such a state of suspiciousness in our participants, we confronted them with untrustworthiness cues in the beginning (see below). Then, they had to solve easy tasks in a fully immersive virtual world together with an ostensible partner while we tracked the distance between both players. The use of an immersive virtual environment allowed us to accurately, continuously, and unobtrusively measure participants' distancing behavior (Bailenson et al., 2003; Kane et al., 2012) and, at the same time, to standardize the behavior of the ostensible partner (all of the confederate's movements were pre-programmed). After completing the tasks in the virtual world, participants played a modified trust game (Berg et al., 1995). Victim sensitivity and other personality traits were measured several days or weeks prior to the laboratory appointment in a separate online questionnaire.

In sum, Study 1 tested the mediation described in the SeMI model (Gollwitzer et al., 2013). Thus, we predicted that after being confronted with untrustworthiness cues, people high in VS would keep more distance toward their partner's avatar (M1), would interpret the other person's intentions and behavior in a more distrustful, hostile way (M2), and would more strongly endorse legitimizing cognitions regarding their own uncooperative behavior (M3). We further assumed that these mechanisms would explain (i.e., mediate) the negative effect of VS on participants' cooperation in a trust game (DV).

Participants

Taking into account the efforts that come with the use of the VR laboratory, our goal was to recruit as many participants as feasible within our six-month time frame (November 2018 to May 2019). The study, which was ostensibly about "coordination in a virtual world," was advertised via university mailing lists, flyers, and an advertisement in a local newspaper. The study consisted of two parts: an online questionnaire (which was completed by 269 participants) and a subsequent laboratory appointment several days or weeks later (here, 92 subjects took part). Participants whose datasets could not be unequivocally matched were excluded. In addition, three participants had to be excluded because they were suspicious with regard to the confederate ($n=2$) or because of technical issues in the virtual reality laboratory ($n=1$). Thus, the final sample consisted of 84 participants (48% male, 52% female). The majority (93%) were undergraduate students from a wide variety of disciplines; ages ranged from 18 to 50 years ($M=23.6$, SD$=4.6$ years). Participants received monetary compensation for their time, the exact amount (4€-10€) was dependent upon their decisions in the trust game.

Materials and Measures

As mentioned above, the study consisted of an online questionnaire in which personality traits were measured, and a laboratory appointment in which we assessed social distancing, hostile interpretations, legitimizing cognitions, and trusting behavior.

Data from these different parts were matched via a personalized code that participants created on their own and that ensured (pseudo-)anonymity. For privacy reasons, this code was later deleted and replaced by a numerical code in the dataset.

Personality Traits

At the beginning of the online questionnaire, participants were informed that the study consisted of two parts and that it would take approximately one hour to complete. After giving informed consent, participants provided demographic information and then completed several trait measures, including victim sensitivity (VS), observer sensitivity (OS), general trust, and the Big Five. Victim and observer sensitivity were assessed with 10 items each (Justice Sensitivity Inventory; Schmitt et al., 2010), which were answered on a Likert scale from 0 ("totally disagree") to 5 ("totally agree"). Example items are: "It bothers me when others receive something that ought to be mine" (VS) and "I am upset when someone does not get a reward that he/she has earned" (OS). The scales have good internal consistency (in the present study: Cronbach's α's = 0.84 and 0.89, respectively). General trust ($\alpha = 0.81$) was measured with a German version of the six-item General Trust Scale by Yamagishi and Yamagishi (1994) and was answered on a six-point Likert scale as well. Finally, neuroticism ($\alpha = 0.80$), extraversion ($\alpha = 0.85$), and openness to experience ($\alpha = 0.58$) were measured with the BFI-10 (Rammstedt & John, 2007); here, response scales ranged from 1 to 5. Several other scales were assessed as well but because they are not relevant for the present research, they will not be referred to any further. At the end, participants had the possibility to sign up for the second part of the study, which was scheduled several days or weeks later. In addition, they were informed that the laboratory appointment would take place with a second participant.

Virtual Reality

In the laboratory, participants were informed that their partner (i.e., a female confederate) would be a couple minutes late but that they would start preparations without her. The confederate being unpunctual had two reasons: First, because it should raise doubts about her dependability and trustworthiness, and second, because participants should not see the other person in the "real" world (to avoid any positive or negative first impressions). For this reason, participants were also told not to talk to the other person once she arrived.

Next, participants received an information sheet, which informed them that they would first solve different coordination tasks in a virtual world and that they would play a (trust) game with their partner afterward. They were also given a brief summary of the trust game instructions to make the possibility of a later exploitation salient. After participants gave their informed consent, they put on a head mounted display (HMD) and were led to their starting position for the virtual tasks (while wearing the HMD, participants were unable to see any part of the "real" world). Next, the experimenter secretly informed the confederate via SMS or WhatsApp,

who arrived at the laboratory approximately two minutes later. The confederate signed the informed consent as well and then (ostensibly) put on a second HMD.[1]

Next, the experimenter explained that both players would be randomly assigned to an avatar (i.e., a human representation in the virtual world), but that they would only see the avatar of their counterpart. In fact, participants did not see any virtual representation of their own body. The confederate, however, was always represented by a female avatar with a neutral facial expression (see Fig. 1). A neutral appearance was chosen because, as previous research has demonstrated, victim-sensitive individuals perceive neutral faces as less trustworthy (Gollwitzer et al., 2012; Study 1). A neutral facial expression, albeit a subtle cue, should thus activate a suspicious mindset in individuals high in VS, especially if combined with other untrustworthiness cues (like the tardiness of the confederate and the possibility of later exploitation in the trust game). Next, the experimenter started the virtual environment. Here, participants saw a spacious plaza with a fountain, plants, and some shops and restaurants (see Fig. 1). In addition, a blue and a red cross on the ground indicated the starting points for the respective players (participant and confederate were facing each other when standing on their starting positions). We did not present participants with auditory input in the virtual world because they had to listen to the experimenter's instructions.

After familiarization with the virtual world in a warm-up task, participants' approach-avoidance behavior was assessed in three proxemic tasks. Here, participants were asked to (1) walk around their partner to read a combination of numbers and letters from their back (this task was validated in previous research, see Bailenson et al., 2003), (2) to avoid additional agents (i.e., virtual humans controlled by a computer algorithm) who walked through the virtual place, and (3) to approach their

---

[1] We used two fully functional HMDs for the study (nVisor ST50 and HTC Vive); however, the second HMD was only a requisite and not connected to the virtual world. Thus, while participants were led to believe that the other player controlled the avatar they were seeing, the confederate was not even able to see the virtual world. Instead, our research assistants had been trained to imitate the pre-programmed avatar (to make it more realistic, the avatar followed recorded movements of a real person). As a result, participants saw the avatar moving in the virtual world and, at the same time, heard the confederate's footsteps in the real world.

partner up to the point where the closeness still felt natural and comfortable for them (e.g., Sommer, 2002; Uzzell & Horne, 2006).

Importantly, the two players took turns in performing these tasks; thus, only the participant or avatar were moving at any given time (in Tasks 1 and 3 the participant went first, in Task 2 the confederate). In all proxemic tasks, the position and orientation of participant and avatar were continuously tracked and these data were used to compute the minimum distance for each task.[2] We chose minimum distance as a measure for social distancing because it has been successfully used in both traditional proxemic research and studies investigating interpersonal distance in virtual environments (e.g., Bailenson et al., 2001, 2003).

Trust Game

After leaving the virtual world, participants were escorted to another room in which they were seated in front of a computer screen. They were informed that they would now play a game with their partner and that the points they won in this game would be added to their 2€ compensation fee: For every 10 points they gained, they would be payed 1€ extra. The experimenter then left the room, ostensibly to tend to the other person who still waited in the virtual reality laboratory,[3] while participants familiarized themselves with the instructions of the game (adapted from Berg et al., 1995). The rules were explained as follows: Both players would make two decisions, one in the role of "sender" (i.e., trustor) and one in the role of "receiver" (i.e., trustee), but they would not be informed about the decisions of their partner until the end of the game. As senders, they would receive 20 points, which they could either keep (those points would be added to their payment) or send to the receiver. More specifically, participants could send 0, 5, 10, 15, or 20 points that were then quadrupled by the experimenter. The receiver, on the other hand, could then decide to either keep all or to share any amount of the quadrupled points with the sender.

After participants read these instructions, they first made their decision as senders (which was our focal DV in this study) and then as receiver. However, because they were not informed about the decisions of their alleged partner during the game, participants had to make the second decision for all possible outcomes (strategy method). After each decision, participants answered a short follow-up survey about their feelings and thoughts during the respective decision (e.g., "I am satisfied with my decision" and "I was afraid of being taken advantage of by the other player"; see Table 4).

---

[2] Participants' position was tracked from the front right of the HMD. We only used length and width coordinates for distance computations; thus, height differences between participants did not influence distance measures.

[3] Participants always left the virtual reality lab first (while the confederate still wore the HMD) to make sure that participant and confederate did not interact with each other outside of the experimental tasks.

SeMI Components

After the trust game, participants answered a series of questions about their perception of the other person and her behavior in the virtual world. Four of these items assessed hostile interpretations (e.g., "Do you think the other person is trustworthy?" (reversed), "Do you think the other person has hostile motives?"; Cronbachs $\alpha = 0.75$) and two items measured legitimizing cognitions (i.e., "Do you think it is justified not wanting to work with the other person?", "Suppose you were to work on a project with the other person (e.g., for a seminar). Do you find it justified to refuse?"; Cronbachs $\alpha = 0.67$). Ten additional items (e.g., "I had the impression that the other person kept some distance to me in the virtual world") were included for exploratory purposes. All of these follow-up items (see also Appendix A; Tables 5 and 6) were answered on six-point Likert scales. Finally, participants were probed for suspicion, fully debriefed, and received their compensation fee.[4]

Results and Discussion

To test our hypotheses, we correlated victim sensitivity (and other relevant personality traits) with our three mediator variables and our DV (i.e., trust behavior) in a first step. These correlations, as well as means and standard deviations of the respective variables, can be found in Table 1. Correlations among personality measures resembled earlier findings (e.g., Schmitt et al., 2005, 2010): For example, victim sensitivity was positively related to observer sensitivity and neuroticism, but uncorrelated with extraversion or openness.

As can be seen in Table 1, minimum distance,[5] that is, the shortest distance that participants maintained between themselves and the avatar of the other person, was uncorrelated with victim sensitivity ($r = 0.11$, $p = 0.31$). Interestingly, though, exploratory analyses of the follow-up items revealed that whereas participants high in VS showed no avoidance tendencies themselves, they still perceived *the other person* to be more distant and reserved. Thus, despite the fact that avatar movements were pre-programmed, victim-sensitive participants more strongly had the impression that the other person kept some distance to them in the virtual world ($r = 0.25$, $p = 0.02$). Unexpectedly, we also found observer sensitivity and openness to experience to be positively correlated with the social distance measure, which implies that participants high in OS and openness kept more distance to the avatar than participants low in these traits.

With regard to the other two suspicious mindset components, we found that VS was positively correlated with hostile perceptions of the other player ($r = 0.24$,

---

[4] The calculation of the fee was standardized but depended on the decisions in the trust game. Thus, participants received between 4€ and 10€.

[5] Looking at this first potential mediator, we found, rather unexpectedly, that the minimum distances measured in the different proxemic tasks were uncorrelated with each other (r's = -.04 to .19, all p's > .08). For this reason, we did not compute a mean across tasks as originally planned; instead, we focused on Task 1 (where participants had to read the number from the other person's back) because this task has been used and validated in previous research on interpersonal distance using virtual reality technology.

Table 1 Means, standard deviations, and correlations between measured variables (Study 1)

Variable	M (SD)	Correlations									
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Victim sensitivity	2.92 (0.83)	1									
(2) Observer sensitivity	2.96 (0.90)	.41**	1								
(3) General trust	2.96 (0.82)	−.18	.08	1							
(4) Neuroticism	3.17 (1.06)	.45**	.20	−.09	1						
(5) Extraversion	2.97 (1.10)	−.04	.14	.27*	−.38**	1					
(6) Openness	3.82 (0.94)	.04	.07	.14	.02	.13	1				
(7) Minimum distance	0.46 (0.14)	.11	.27*	.07	.13	.10	.24*	1			
(8) Hostile interpretation	2.14 (0.66)	.24*	−.15	−.34**	.10	−.03	.15	−.08	1		
(9) Legitimizing cognitions	2.24 (1.25)	.14	−.25*	−.34**	.09	−.07	−.05	−.03	.37**	1	
(10) Trust behavior	16.25 (4.96)	.09	−.02	.004	−.01	−.08	−.04	.19	−.09	−.06	1

$N = 84$. Response scales on personality scales ranged from 0–5 (Victim Sensitivity, Observer Sensitivity, General Trust) or from 1 to 5 (Big Five). Minimum Distance refers to the shortest distance participants kept to the avatar when reading the number on her back (in meters). Hostile Interpretation was measured with 4 items; Legitimizing Cognitions with two items (on a scale from 1 to 6). Trust Behavior refers to the sender decision in the trust game (measured on a scale from 0 to 20 points).

$^* p < .05$; $^{**} p < .01$ (two-tailed)

$p = 0.03$), but uncorrelated with legitimizing cognitions ($r = 0.14$, $p = 0.21$). Thus, participants high in VS interpreted the other person's intentions and behavior in a more negative, mistrustful way, but they did not consider uncooperativeness toward the other person as more justifiable. However, this may have not been necessary because participants high in VS showed no uncooperative behavior in our study anyway: Contrary to our hypothesis and to the results of prior research, we found no correlation between victim sensitivity and trust behavior ($r = 0.09$, $p = 0.40$). In fact, the decision to send points was unrelated to any of the personality traits measured, as well as to the suspicious mindset variables (see Table 1).

Mediation Analysis

Despite the insignificant total effect of VS on trust behavior, we conducted our multiple mediation analysis as planned. This was based on the reasoning that indirect effects can exist even though a total effect is not statistically significant (e.g., Hayes, 2018), for instance, because different indirect effects may suppress each other.

More specifically, we included all three mediator variables (i.e., hostile interpretations, legitimizing cognitions, social distancing) simultaneously into the model while controlling for observer sensitivity, general trust, neuroticism, and extraversion.[6] An overview of the results can be found in Fig. 2. Overall, only 10% of the variance in trust behavior was explained by our model. Contrary to our hypothesis, victim sensitivity had no (total or direct) effect on trust. With regard to the mediators, we found VS to predict hostile interpretations ($\beta = 0.27$, $p = 0.04$), but no significant effects on legitimizing cognitions or social distancing were found (p's $= 0.14$ and 0.78, respectively). In accordance with the correlation analyses, none of these mediator variables significantly influenced the amount of points sent in the trust game (although the positive effect of social distancing reached marginal significance). For this reason, it is not surprising that no indirect effect of VS on trust was found: Bootstrapping results showed that the 95% confidence intervals always contained zero. Thus, we were unable to replicate the established negative relationship between victim sensitivity and trust, and we also did find no support for the assumption that social distancing, hostile interpretations, and legitimizing cognitions (as elements of a suspicious mindset) mediated this hypothesized relation.

In sum, victim sensitivity was not (directly or indirectly) related to distrust in our findings. However, when taking a closer look at our DV it became clear that a ceiling effect (which reduces variation in trusting behavior) might have contributed to this lack of findings: While more than 50% of participants decided to send all of their 20 points to their partner, only one person sent nothing.[7] Importantly, exploratory

---

[6] Openness to experience was not included as a covariate because of the scale's low internal consistency.

[7] The median was 20, the mean was 16.25 (SD $= 4.962$), and the skewness index was -1.13 (SE $= 0.263$), suggesting a high negative skewness and a significant deviation from the normal distribution (Shapiro-Wilks' $W = 0.75$, $p < .001$). For this reason, we used the Box-Cox transformation (Box & Cox, 1964) as described in Study 2. This procedure estimates a parameter (lambda or "Box-Cox parameter") which yields a quasi-optimal fit of a variable against a normal distribution. We used Free Statistics Software (v1.2.1) to estimate this parameter (Wessa, 2016), which equaled $= 1.47$ in our case. The R code and
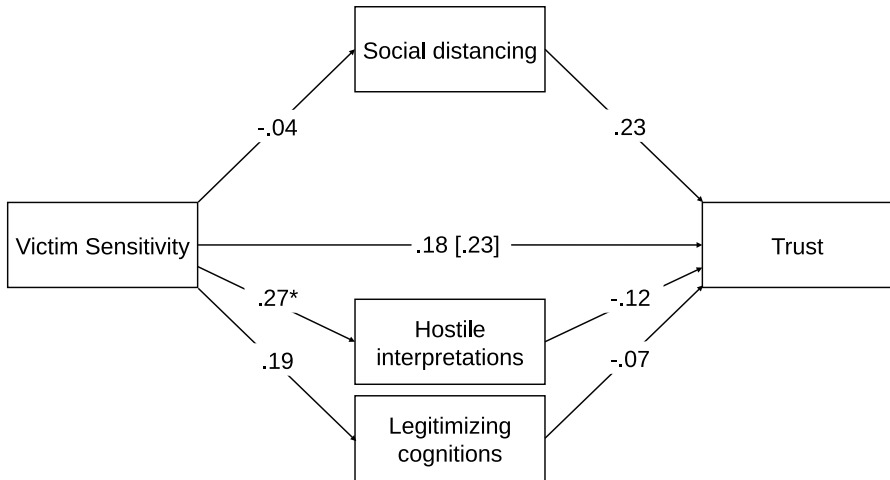
Fig. 2 Multiple mediation of the effect of VS on trust via social distancing, hostile interpretations, and legitimizing cognitions. Standardized regression coefficients are depicted. $^{*}p < .05$ (two-tailed)

analyses of the follow-up items revealed that high-VS participants were less satisfied with their decision ("I am satisfied with my decision"; $r = -0.29$, $p = 0.01$) and had a greater fear of exploitation in the trust game than low-VS participants ("I was afraid of being taken advantage of by the other player"; $r = 0.24$, $p = 0.03$). In other words, participants high in VS were afraid that the other person might take advantage of them, but they did not act upon this fear. One reason for this could be that the untrustworthiness cues that we provided participants with were not strong enough to trigger self-protective behavior. Another explanation is given by Dunning et al. (2014): These authors argue that humans sometimes show trust in others (especially strangers) because they believe that norms of respect force them to do so. Thus, in many situations, individuals do not really expect benevolence from a cooperation partner, but feel obliged to show trusting behavior nonetheless. It is therefore possible that victim-sensitive participants in our study sent points to their partner because this was what they felt they *should* do, although it was not what they really *wanted* to do.

Although we were unable to find support for our main hypotheses, the results of Study 1 still provided some important insights into the defensiveness mechanism underlying victim sensitivity. For instance, our results show that whereas VS did not predict social distancing, participants high in VS perceived the confederate to be physically distant—despite the fact that the partner's behavior was standardized. Together with the finding that victim sensitivity promoted hostile interpretations

---

Footnote 7 (continued)

the Box-Cox Normality Plot for our data is stored here: https://www.freestatistics.org/blog/index.php?v=date/2020/Oct/12/t1602496919kua7mdhjy2olhr5.htm/. Importantly, results were virtually identical when using the transformed trust variable. To simplify interpretation, we therefore report results with the DV in its original metric.

of the other person's behavior and intentions in general, this lends further support for the "inaccuracy hypothesis" (Gollwitzer et al., 2012, 2013): According to this notion, VS implies an asymmetrical sensitivity toward untrustworthiness cues, which results in less accurate predictions of other people's cooperativeness (i.e., a bias in social judgments). In addition, it is also noteworthy that VS predicted hostile interpretations even though our cues of untrustworthiness were rather subtle in this study.

In conclusion, we were able to demonstrate that victim sensitivity was related to more unfavorable perceptions of the other player as well as to fear of exploitation, but we did not find any effects on a behavioral level (i.e., participants high in VS showed no social distancing and no withdrawal of trust and cooperation). Hence, while these results still suggest that victim-sensitive individuals' information processing is characterized by a pronounced suspiciousness in socially uncertain situations, no conclusions about the underlying motivational processes could be drawn. Therefore, Study 2 was conducted to investigate why individuals high in VS are more likely than others to act pre-emptively defensive and how their self-protective strikes may be attenuated.

Study 2

Study 2 was designed to test the hypothesis that restoring a sense of control over what happens in one's social environment can effectively alleviate victim-sensitive individuals' tendency to distrust others. Similar to Study 1, participants played a trust game with an unknown fellow participant and participants' behavior as senders (or "trustors") was used as our focal DV. To manipulate the activation of a suspicious mindset, half of the participants were confronted with cues of untrustworthiness (i.e., a respective facial expression and a message; see below).

To restore a sense of control, we used an established control-affirmation procedure adapted from Kay et al. (2008), in which participants are prompted to recall a situation in which they had recently experienced a sense of control (see Methods section). Although this procedure has been successfully used in prior research, we figured that recalling a situation in which participants had control might not only affirm their momentary sense of control, but may also and inadvertently boost their self-regard. For instance, recalling a situation in which one has shown moral courage and protected another person from being bullied may effectively reinstall a sense of control, but may also (re)affirm one's moral virtues and self-regard (e.g., Monin & Miller, 2001). Therefore, in order to test whether restoring a sense of control is *specifically* responsible for alleviating the effect of suspiciousness on distrust among victim-sensitive individuals, we also implemented an experimental condition in which participants were self-affirmed (using a procedure adapted from Monin et al., 2008).

Thus, Study 2 used a 2 (suspicious mindset, no suspiciousness) × 3 (control-affirmation, self-affirmation, no affirmation) between-subjects design with victim sensitivity as a continuous moderator variable. We predicted that when a suspicious

mindset has been activated, victim-sensitive individuals would be less trustful toward a fellow participant when no affirmation occurred and that this effect was alleviated specifically after restoring participants' sense of control. Thus, we hypothesized a suspiciousness×affirmation×victim sensitivity three-way interaction effect.

Participants

Undergraduate students from a mid-size German university were invited via email to participate in an online study about "social interactions, self-views, and emotions." In exchange for participation, five online book vouchers worth 50€ each were raffled among all participants. The invitation contained a link to the survey, which was active for two months. We aimed at recruiting as many participants as possible within this time frame. Detecting a small- to medium-size three-way interaction effect (i.e., $f^2 = 0.03$, as in Gollwitzer & Rothmund, 2011; Study 2), with a power of 80% would have required a sample size of 325 if $\alpha = 5\%$ (calculated with G*Power 3.1; Faul et al., 2009).[8]

At the end of the data collection phase, 321 participants had started the study; 291 completed it. Participants who needed less than 1 min ($n = 12$) or more than 2 h to complete the online survey ($n = 5$) were discarded from further analyses. One additional case was deleted because this participant gave an invalid response (i.e., 12) to the trust game measure (values could only range between 0 and 10, see below). Thus, the final sample consisted of 273 participants (173 female; 100 male). Ages ranged between 18 and 64 years ($M = 25$, $SD = 7.1$ years). Students were enrolled in a broad range of undergraduate programs, including psychology ($n = 18$). A sensitivity power analysis revealed that, with 273 valid cases, the population effect would have to be equal or larger than $f^2 = 0.036$ in order to be detected with a probability (i.e., power) of 80% on a significance level of 5%. We considered this to be at the upper limit of a plausible effect, given that especially ordinal interactions are often smaller than expected (e.g., Lakens & Caldwell, 2021). Therefore, we decided to increase our significance level to $\alpha = 10\%$, which we considered justifiable given that our hypothesis is directional. This increased our power to detect our pre-specified population effect of $f^2 = 0.03$ with 273 participants to 82%.

---

[8] To test the hypothesized three-way interaction effect, we specified a multiple regression model with 13 predictors (i.e., 4 unconditional effects, 2 covariates (Observer Sensitivity and General Trust), 5 two-way interactions, and 2 three-way interactions, see also Table 3). The two three-way interaction effects were the focal predictors in this model.

Materials and Measures

Trust Game

After participants gave their informed consent, they were introduced to the rules of the trust game. They were told that they would be paired with a second (female)[9] participant from another university in the south of Germany, with which we would be cooperating on a research project. Participants were informed that they would play two rounds with their partner: In Round 1, participants would be the "sender" (i.e., trustor) and their partners would be "receivers" (i.e., the trustee). In Round 2, roles would be switched. Unbeknownst to participants, there was no second round, and there also was no other participant. Participants received 10 lottery tickets, of which they could send any number (between 0 and 10) to their partner. The number of tickets sent would be tripled by the experimenter so that receivers would now have their own 10 tickets plus the tripled number of received tickets from their partner (i.e., the real participant). At that point, receivers would have to decide whether or not they want to split the total number of tickets equally between both players (this version of the trust game has been described by Kuwabara, 2005; for an application in research on the SeMI model, see Gollwitzer & Rothmund, 2011; Study 2). Participants were also told that the number of lottery tickets they had in the end would determine their chances of winning the 50€ book voucher: The more tickets, the higher the likelihood of winning one of these vouchers.

Suspiciousness Manipulation

Next, participants saw a picture that their respective partner has allegedly taken of herself during their experimental session. These pictures were selected from the Amsterdam Dynamic Facial Expression Set (ADFES; van der Schalk et al., 2011), a stimulus set containing 648 emotional (facial) expressions displayed by 22 models (for a validation, see Wingenbach et al., 2016). For the present purpose, two pictures were selected (see Fig. 3), one displaying a neutral facial expression (left panel) and the other displaying joy (right panel). More precisely, the neutral facial expression was used to induce a "suspicious mindset" among participants, whereas the joyful facial expression was used to prevent the induction of suspiciousness. Using similar stimuli, Gollwitzer et al., (2012; Study 1) showed that a neutral facial expression is indeed sufficient to induce a suspicious mindset among victim-sensitive individuals.

To make our manipulation even stronger, participants received a (fake) message from their partner, which she allegedly wrote to introduce herself. These messages were designed to induce a suspicious mindset ("*Hi, my name is Heike. I'm 25 and studying business administration in Tübingen. It's awesome that I have the chance to win something here.*") vs. to reduce the likelihood of suspicions being raised ("*Hi, my name is Sarah. I'm 25 and studying social work in*

---

[9] Partners were female because the majority of participants in our study was female. Our rationale for not varying the partner's sex was that doing so would have reduced the statistical power for detecting the suspiciousness affirmation victim sensitivity three-way interaction effect.

Fig. 3 Stimulus material used in Study 2 to induce (left panel) vs. not induce (right panel) suspiciousness. Stimuli were selected from the Amsterdam Dynamic Facial Expression Set (ADFES; van der Schalk et al., 2011). Reprinted by permission

Tübingen. It's awesome that we both have the chance to win something here."). Prior research showed that a message in which one's partner focuses on maximizing their own profit (as in the suspicious mindset condition) vs. on maximizing the team's collective profit (as in the no suspiciousness condition) can effectively amplify the effect of distrust on uncooperative responses (Parks et al., 1996).

Affirmation Manipulation

Next, participants were randomly assigned to either a control-affirmation, a self-affirmation, or a no-affirmation (control) condition. Participants in the *control-affirmation condition* were instructed to recall a positive event from the last couple months for which they were personally responsible, in which they experienced a "sense of control over what had happened" (adapted from Kay et al., 2008). Participants in the *self-affirmation condition* were prompted to recall a positive event from the last couple months in which they were able to display a virtue or express a value that was personally relevant for them, such as a sense of humor, social competence, and musical talent (taken from Monin et al., 2008). Participants in the *no-affirmation (control) condition* were asked to list all the things they do on a typical weekday. In all three conditions, participants were asked to use the free-response field provided to write a short text (100 words max). These texts were screened and coded for validity and credibility. All participants followed the instructions and wrote a principally credible text, which was in line with the respective prompt.

After writing these texts, participants were re-familiarized with the rules of the trust game. Then, they actually played the trust game and decided how many lottery tickets they would want to transfer to their partner. This measure was our dependent variable.

Personality Traits

Finally, participants completed a battery of personality trait measures, including the 10-item Victim Sensitivity Scale ($\alpha = 0.83$), the 10-item Observer Sensitivity Scale ($\alpha = 0.84$), both taken from Schmitt et al. (2010), and the 6-item General Trust Scale ($\alpha = 0.82$) taken from Yamagishi and Yamagishi (1994). Response scales on all measured variables ranged between 1 ("totally disagree") to 6 ("totally agree"). After responding to a number of control questions (e.g., "What do you think this study was about?"), participants were informed about their final amount of lottery tickets, thanked, and debriefed.

Results and Discussion

Looking at our central DV, participants' decisions in the trust game, many of them ($n = 119$, i.e., 44%) decided to transfer all of their lottery tickets (i.e., 10) to their partner. Only 7 participants (2.6%) decided to keep their 10 tickets for themselves. The median was 8, the mean was 7.57 (SD = 2.7), and the skewness index was -0.83 (SE = 0.15), suggesting a moderate to high negative skewness and a significant deviation from the normal distribution (Shapiro–Wilks' $W = 0.832$, $p < 0.01$). Therefore, we used a procedure proposed by Box and Cox (1964) to find the optimal transformation for our variable,[10] which, in our case, was

$$Y_{\text{trans}} = \frac{(Y + 0.5)^{\lambda} - 1}{\lambda}$$

with Y denoting the trust measure in its original metric, and Y_{trans} denoting the transformed variable (see also Fox & Weisberg, 2011; Velilla, 1993). Means, standard deviations, and correlations between the transformed trust variable and the three personality traits measured are displayed in Table 2.

As expected, victim sensitivity was positively related to observer sensitivity and negatively related to general trust, which is why we used observer sensitivity and general trust as covariates in the regression model reported below. In addition, the transformed trust variable was negatively related to victim sensitivity and positively related to general trust across all experimental conditions.

We predicted that when a suspicious mindset has been activated, victim-sensitive individuals would be less trustful toward a fellow participant when no affirmation occurred, and that this effect would be alleviated specifically by a control-affirmation (vs. a self-affirmation) procedure. To test the hypothesized suspiciousness × affirmation × victim sensitivity three-way interaction effect, we contrast-coded the suspiciousness conditions (no suspiciousness = −1, suspicious mindset = +1) as well as the affirmation conditions so that Contrast 1 reflected an effect of control-affirmation

---

[10] This procedure estimates a parameter (lambda or "Box-Cox parameter") which yields a quasi-optimal fit of a variable against a normal distribution. We used Free Statistics Software (v1.2.1) to estimate this parameter (Wessa, 2016), which equaled = 1.38 in our case. The R code and the Box-Cox Normality Plot for our data is stored here: https://www.freestatistics.org/blog/index.php?v=date/2020/May/13/t1589377509uogl22wdbqyqyvf.htm/.

Table 2 Means, standard deviations, and correlations between measured variables (Study 2)

Variable	M (SD)	Correlations		
		(1)	(2)	(3)
(1) Trust (transformed)	12.63 (5.58)	1		
(2) Victim sensitivity	3.59 (0.85)	−.14*	1	
(3) Observer sensitivity	3.90 (0.86)	.10	.22**	1
(4) General trust	3.90 (0.85)	.18**	−.17**	.19**

$N = 273$. Response scales on personality scales ranged from 1 to 6
$^*p < .05$; $^{**}p < .01$ (two-tailed)

Table 3 Results from the moderated regression model (Study 2)

| Predictor	B	SE(B)	t
Constant term	12.852	0.339	37.956*
Victim sensitivity (VS)	−0.592	0.358	−1.654†
Suspiciousness manipulation	−0.428	0.339	−1.263
Observer sensitivity	0.516	0.349	1.477
General trust	0.812	0.345	2.358*
Contrast_1 (control-affirmation vs. baseline)	0.681	0.498	1.366
Contrast_2 (self-affirmation vs. baseline)	0.247	0.475	0.519
VS × suspiciousness	−0.561	0.343	−1.637
VS × contrast_1	−0.015	0.501	−0.030
VS × contrast_2	0.488	0.495	0.986
Suspiciousness × contrast_1	−0.096	0.502	−0.190
Suspiciousness × contrast_2	0.120	0.476	0.252
VS × suspiciousness × contrast_1	0.913	0.505	1.810†
VS × suspiciousness × contrast_2	−0.781	0.495	−1.577

$N = 273$

$^†p < .10$ $^*p < .05$ (two-tailed). Victim Sensitivity, Observer Sensitivity, and General Trust are z-standardized

vs. baseline (control-affirmation = +1, self-affirmation = 0, no-affirmation = −1) and Contrast 2 reflected an effect of self-affirmation vs. baseline (control-affirmation = 0, self-affirmation = +1, no-affirmation = −1). Victim sensitivity was standardized to facilitate the interpretation of regression weights. Observer sensitivity and general trust were included as covariates.

The estimated regression coefficients are reported in Table 3. The two focal predictors in this model are the three-way interaction terms in the two bottom rows of the table. Adding these two terms to the model increased the explained variance by 1.3% (from $R^2 = 0.093$ without the three-way interaction terms to $R^2 = 0.106$ including them). Inspecting their regression weights suggests that a control-affirmation procedure (but not a self-affirmation procedure) was able to make victim-sensitive individuals just as trustful as their victim-insensitive counterparts, as indicated by

Fig. 4 Predicted values visualizing the suspiciousness × affirmation interaction effect for participants low vs. high in victim sensitivity (± 1 *SD* around the sample mean; Study 2)
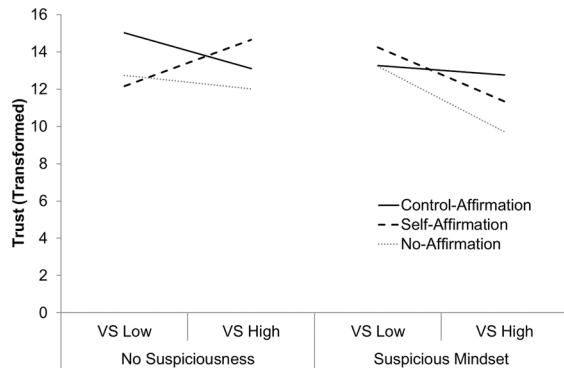
the VS × Suspiciousness × Contrast_1 interaction effect, $B = 0.913$, $SE(B) = 0.505$, $p = 0.072$. This interaction effect is also graphically displayed in Fig. 4. Simple slope analyses confirmed that while victim sensitivity was negatively related to trust in the suspicious mindset condition when no affirmation occurred ($B = -1.834$, $p = 0.015$), this effect disappeared in the control-affirmation condition ($B = -0.459$, $p = 0.596$). Thus, although participants high in VS tend to distrust others in socially uncertain situations, this pre-emptive defensiveness can actually be attenuated if a sense of control is (re-)established.

General Discussion

Being duped is a highly aversive, painful experience which people are motivated to avoid. For this reason, especially people with a high fear of exploitation (i.e., people high in victim sensitivity) tend to act pre-emptively selfish and hostile in situations in which the risk for being "suckered" is considered to be high. However, although past research has repeatedly investigated the relation between VS and uncooperativeness in socially uncertain situations (e.g., Faccenda et al., 2009; Fetchenhauer & Huang, 2004; Gollwitzer et al., 2005; Lavelle et al., 2018), the assumed mechanism (i.e., the suspicious mindset) has received far less attention (for exceptions, see Gerlach et al., 2012; Maltese et al., 2016; Rothmund et al., 2011). To close this research gap, we conducted two studies investigating the cognitive and motivational processes underlying such pre-emptive defensiveness. Together, the findings crucially expand our knowledge of the defensive motivational system in victim-sensitive individuals.

The Suspicious Mindset

To better understand how victim sensitivity translates into selfish and uncooperative behavior, we examined possible mediating processes in Study 1. More precisely, we used virtual reality technology to create an immersive environment in which participants interacted with another person to unobtrusively measure social distancing

(as a form of self-protective, avoidant behavior). Although our main focus was on this approach-avoidance behavior, we also assessed hostile interpretations of the other person's behavior and intentions, and legitimizing cognitions with regard to own uncooperative reactions in self-report (i.e., the other two components that are assumed to constitute the suspicious mindset). Finally, participants played a trust game with their alleged partner over real money.

Notably, the results did not support our hypotheses for the most part. First, we were unable to show the expected correlation between victim sensitivity and social distancing. More specifically, VS only predicted hostile and mistrustful perceptions of the other person, but not legitimizing cognitions or avoidance behavior. However, although participants high in VS showed no distancing tendencies themselves, they still perceived *the other person* to keep their distance, which was by design a purely subjective impression. Thus, our findings corroborate the notion that VS promotes a bias in social judgments: Victim-sensitive individuals focus strongly on untrustworthiness cues, which distorts their perceptions of their social environment (Gollwitzer et al., 2012, 2013).

Second, trust rates were overall very high, yet unrelated to any of the measured variables. In contrast to previous studies, victim-sensitive participants did not send fewer points to their partner in the trust game, even though they reported a higher fear of exploitation than victim-insensitive participants. Thus, we were unable to replicate the established effect of victim sensitivity on distrust, and, as a consequence, we were also unable to show that the suspicious mindset mediated this relation. However, there are a number of possible reasons that might explain why the expected direct and indirect effects of VS on trust were absent in our findings. For instance, the sample size may have simply been too small to ensure adequate statistical power for the detection of the hypothesized effects. Monte Carlo simulations suggest that several hundred participants are required if small to medium sized effects are tested in a (multiple) mediation model (Fritz & MacKinnon, 2007; Ma & Zeng, 2014; Thoemmes et al., 2010). For this reason, future studies should aim at recruiting a much larger sample than the one used in the present study.

In addition, the SeMI model and previous research (e.g., Gollwitzer & Rothmund, 2011; Rothmund et al., 2011; Süssenbach & Gollwitzer, 2015) suggest that VS predicts uncooperative and distrustful behavior only if there are sufficient indications that an interaction partner might be untrustworthy. The untrustworthiness cues that participants encountered in Study 1, though, were relatively subtle. In addition, participants interacted with the other person for quite some time before playing the trust game, and, as previous research demonstrates, prior contact with a partner can increase trustworthiness perceptions (Glaeser et al., 2000). That said, the fact that victim-sensitive participants reported more fear of exploitation and less satisfaction with their own decision in the trust game corroborates the notion that our manipulation was (at least, in part) successful: Participants high in VS *were* more fearful and suspicious of their partner's intentions, although they did not act on it.

In conclusion, victim sensitivity was exclusively related to an attributional bias regarding others' malevolence in Study 1, which confirms and complements past

findings on the suspicious mindset in victim-sensitive individuals (Gerlach et al., 2012; Maltese et al., 2016; Rothmund et al., 2011). However, although VS predicted more unfavorable perceptions of the other player, we were not able to demonstrate that this bias actually mediated behavioral trust and cooperation.

Motivational Basis of Defensiveness

To examine why exactly victim-sensitive people tend to react defensively in socially uncertain situations and how such pre-emptively selfish behavior may be alleviated, a second study was conducted. More precisely, Study 2 was designed to test whether victim-sensitive individuals are defensive because (1) they are afraid that being exploited threatens their self-image (Vohs et al., 2007) or because (2) they want to maintain a sense of control over their social environment. To compare these two possible motivations directly against each other, we employed an experimental design in which participants were either self-affirmed or control-affirmed (or not affirmed at all) after the activation of a suspicious mindset (vs. after no induction of suspiciousness), and then instructed to play a trust game with an ostensible partner.

Our results suggest that a control affirmation was able to reduce the negative effect of VS on trust behavior when a suspicious mindset was evoked, although the effect was small and only significant on a 10% level. Affirming a positive self-image, on the other hand, did not significantly alleviate the negative effect of VS on trust behavior (see Fig. 4). These findings suggest that affirming a sense of control tends to be more relevant for high-VS individuals than affirming a positive self-image. In this context, it should be noted that, in contrast to Study 1, we activated a relatively strong suspicious mindset in Study 2 to test our hypotheses as strictly and conservatively as possible. The fact that affirming a sense of control made victim-sensitive individuals just as trustful as their victim-insensitive counterparts suggests that restoring control may be an effective strategy to mitigate suspiciousness. Following up on these insights is, in our view, an important avenue for future research.

Limitations

The present research further corroborates the SeMI model's predictions and expands our knowledge about the pre-emptive defensiveness shown by victim-sensitive individuals, but, of course, there are also some limitations. First and foremost, the sample size of Study 1 was rather small, and the effect found in Study 2 (i.e., the VS × Suspiciousness × Contrast_1 interaction effect, which was significant on a 10% level, see Table 3) was also smaller than expected. For this reason, the suspicious mindset and its components should be investigated in more highly powered studies in the future, especially if more complex designs are applied. This will allow to draw conclusive inferences about the underlying defensiveness mechanism.

Another constraint lies in the fact that the minimum distance measures were uncorrelated across proxemic tasks in Study 1, although all of these tasks were designed to assess participants' distancing behavior. However, they differed in their demands: While Task 1 (reading the number from the other person's back) and Task 3 (moving toward the other person and stopping at a comfortable distance) specifically asked participants to approach their partner, Task 2 focused on the avoidance of additional virtual humans (therefore increasing the possibility that participants were too distracted to pay attention to the confederate or to their physical proximity). Here, we focused on the minimum distance assessed in the first proxemic task, which should represent the most valid measure for social distancing in our study.[11] Nevertheless, future research may want to use more than one measure of approach-avoidance behavior (or even additional operationalizations of avoidance motivation such as the avoidance of eye contact).

Conclusion

A large body of empirical findings shows that a high fear of exploitation (i.e., high victim sensitivity) results in uncooperative and hostile behaviors in socially uncertain situations. Importantly, this pre-emptive defensiveness implies a focus on others and their behavior: Instead of worrying about themselves and the impression they make, people high in VS are more concerned about what others are up to. In other words, defensiveness is an inherently social phenomenon, not a merely self-centered phenomenon. Investigating this social sensitivity is important not only from a theoretical but also from an applied perspective: If we understand which cognitive and motivational processes are at work in victim-sensitive individuals, we may be able to develop effective interventions capable of preventing or attenuating adverse effects of VS on trust and cooperation. The present research suggests that boosting a sense of control helps victim-sensitive people to overcome their habitual suspiciousness. Stated differently, providing people with a sense of control might prove to be an effective tool to improve social interactions and to break vicious cycles of non-cooperation.

Appendix A

See Tables 4, 5 and 6

---

[11] Results vary slightly as a function of the distance measure used. For example, the minimum distances assessed in Tasks 2 and 3 were uncorrelated with Observer Sensitivity and Openness. In addition, the direct effect of social distancing on trust in the mediation analysis was marginally significant (although unexpectedly positive) for the minimum distance of Task 1, but far from significant for the distances of Tasks 2 and 3.

Table 4 Exploratory Follow-Up Items Trust Game Decisions

Item
I am satisfied with my decision
I was afraid of being taken advantage of by the other player
I could not resist the temptation to keep as much money as possible for myself
I think my counterpart would have acted exactly like me

Table 5 Exploratory follow-up items behavior in the virtual world

Item
I had the impression that the other person kept some distance to me in the virtual world
The other person seemed to be aware of where I was in the virtual world
The other person did not want to get too close to me in the virtual world
The other person acted in the virtual world the way I expected them to
It seemed as if the other person was executing the tasks as well as they could
In the virtual world, the other person did not seem to care where I was
The distance the other person kept in the virtual world influenced my behavior
The other person did not bother to work with me to solve the tasks in the virtual world
The other person came too close to me in the virtual world
The behavior of the other person had no influence on my behavior in the virtual world

Table 6 Items SeMI components

Item
Would you like to get to know the other person better? [+]
How likeable is the other person? [+]
How well would you get along with the other person? [+]
Do you think one can trust the other person? (reversed) [*]
Do you think the other person has hostile motives? [*]
Do you think the other person is trustworthy? (reversed) [*]
Suppose you were to work on a project with the other person (e.g., for a seminar). Do you find it justified to refuse?[†]
Do you think it is justified not wanting to work with the other person?[†]
Do you think the other person is a difficult cooperation partner? [*]

[*]These items assessed hostile interpretations

[†]These items measured legitimizing cognitions

[+]These items were filler items

No. GO 1674/5–1 and SU 902/2–1) as well as by the Research Training Group "Breaking Expectations" awarded to Philipps-Universität Marburg by the German Research Foundation (project number: 290878970-GRK2271, projects 4 and 6).

Declarations

Conflict of interest We have no conflicts of interest to disclose.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

References

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence, 10*(6), 583–598. https://doi.org/10.1162/105474601753272844

Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin, 29*(7), 819–833. https://doi.org/10.1177/0146167203029007002

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior, 10*(1), 122–142. https://doi.org/10.1006/game.1995.1027

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B, 26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Dunning, D., Anderson, J. E., Schlösser, T., Ehlebracht, D., & Fetchenhauer, D. (2014). Trust at zero acquaintance: More a matter of respect than expectation of reward. *Journal of Personality and Social Psychology, 107*(1), 122–141. https://doi.org/10.1037/a0036673

Faccenda, L., Pantaléon, N., & Reynes, E. (2009). Significant predictors of soccer players' moral functioning from components of contextual injustice, sensitivity to injustice and moral atmosphere. *Social Justice Research, 22*, 399–415. https://doi.org/10.1007/s11211-009-0105-4

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fetchenhauer, D., & Huang, X. (2004). Justice sensitivity and distributive decisions in experimental games. *Personality and Individual Differences, 36*(5), 1015–1029. https://doi.org/10.1016/S0191-8869(03)00197-1

Fiske, S. T. (2003). Five core social motives, plus or minus five. In S. J. Spencer, S. Fein, M. P. Zanna, & J. M. Olson (Eds.), *Motivated social perception: The Ontario Symposium* (Vol. 9, pp. 233–246). Erlbaum.

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression* (2nd ed.). Sage.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological Science, 18*(3), 233–239. https://doi.org/10.1111/j.1467-9280.2007.01882.x

Gerlach, T. M., Allemand, M., Agroskin, D., & Denissen, J. J. A. (2012). Justice sensitivity and forgiveness in close interpersonal relationships: The mediating role of mistrustful, legitimizing, and pro-relationship cognitions. *Journal of Personality, 80*(5), 1373–1413. https://doi.org/10.1111/j.1467-6494.2012.00762.x

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics, 115*(3), 811–846. https://doi.org/10.1162/003355300554926

Gollwitzer, M., & Rothmund, T. (2009). When the need to trust results in unethical behavior: The Sensitivity to Mean Intentions (SeMI) model. In D. De Cremer (Ed.), *Psychological perspectives on ethical behavior and decision making* (pp. 135–152). Information Age Publishing.

Gollwitzer, M., & Rothmund, T. (2011). What exactly are victim-sensitive persons sensitive to? *Journal of Research in Personality, 45*(5), 448–455. https://doi.org/10.1016/j.jrp.2011.05.003

Gollwitzer, M., Schmitt, M., Schalke, R., Maes, J., & Baer, A. (2005). Asymmetrical effects of justice sensitivity perspectives on prosocial and antisocial behavior. *Social Justice Research, 18*(2), 183–201. https://doi.org/10.1007/s11211-005-7368-1

Gollwitzer, M., Rothmund, T., Pfeiffer, A., & Ensenbach, C. (2009). Why and when justice sensitivity leads to pro- and antisocial behavior. *Journal of Research in Personality, 43*(6), 999–1005. https://doi.org/10.1016/j.jrp.2009.07.003

Gollwitzer, M., Rothmund, T., Alt, B., & Jekel, M. (2012). Victim sensitivity and the accuracy of social judgments. *Personality and Social Psychology Bulletin, 38*(8), 975–984. https://doi.org/10.1177/0146167212440887

Gollwitzer, M., Rothmund, T., & Süssenbach, P. (2013). The Sensitivity to Mean Intentions (SeMI) model: Basic assumptions, recent findings, and potential avenues for future research. *Social and Personality Psychology Compass, 7*(7), 415–426. https://doi.org/10.1111/spc3.12041

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). The Guilford Press.

Kane, H. S., McCall, C., Collins, N. L., & Blascovich, J. (2012). Mere presence is not enough: Responsive support in a virtual world. *Journal of Experimental Social Psychology, 48*(1), 37–44. https://doi.org/10.1016/j.jesp.2011.07.001

Kay, A. C., Gaucher, D., Napier, J. L., Callan, M. J., & Laurin, K. (2008). God and the government: Testing a compensatory control mechanism for the support of external systems. *Journal of Personality and Social Psychology, 95*(1), 18–35. https://doi.org/10.1037/0022-3514.95.1.18

Kerr, N. L. (1983). Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology, 45*(4), 819–828. https://doi.org/10.1037/0022-3514.45.4.819

Kuwabara, K. (2005). Nothing to fear but fear itself: Fear of fear, fear of greed and gender effects in two-person asymmetric social dilemmas. *Social Forces, 84*(2), 1257–1272. https://doi.org/10.1353/sof.2006.0020

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science, 4*(1), 1–14. https://doi.org/10.1177/2515245920951503

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37*(6), 822–832. https://doi.org/10.1037/0022-3514.37.6.822

Lavelle, J. J., Harris, C. M., Rupp, D. E., Herda, D. N., Young, R. F., Hargrove, M. B., Thornton-Lugo, M. A., & McMahan, G. C. (2018). Multifoci effects of injustice on counterproductive work behaviors and the moderating roles of symbolization and victim sensitivity. *Journal of Organizational Behavior, 39*(8), 1022–1039. https://doi.org/10.1002/job.2280

Ma, Z.-W., & Zeng, W.-N. (2014). A multiple mediator model: Power analysis based on Monte Carlo simulation. *American Journal of Applied Psychology, 3*(3), 72–79. https://doi.org/10.11648/j.ajap.20140303.15

Maltese, S., Baumert, A., Schmitt, M. J., & MacLeod, C. (2016). How victim sensitivity leads to uncooperative behavior via expectancies of injustice. *Frontiers in Psychology, 6*, 2059. https://doi.org/10.3389/fpsyg.2015.02059

McCall, C., & Blascovich, J. (2009). How, when, and why to use digital experimental virtual environments to study social behavior. *Social and Personality Psychology Compass, 3*(5), 744–758. https://doi.org/10.1111/j.1751-9004.2009.00195.x

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81*(1), 33–43. https://doi.org/10.1037/0022-3514.81.1.33

Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*(1), 76–93. https://doi.org/10.1037/0022-3514.95.1.76

Parks, C. D., Henager, R. F., & Scamahorn, S. D. (1996). Trust and reactions to messages of intent in social dilemmas. *Journal of Conflict Resolution, 40*(1), 133–150. https://doi.org/10.1177/0022002796040001007

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Rothbaum, F., Weisz, J. R., & Snyder, S. S. (1982). Changing the world and changing the self: A two-process model of perceived control. *Journal of Personality and Social Psychology, 42*(1), 5–37. https://doi.org/10.1037/0022-3514.42.1.5

Rothmund, T., Gollwitzer, M., & Klimmt, C. (2011). Of virtual victims and victimized virtues: Differential effects of experienced aggression in video games on social cooperation. *Personality and Social Psychology Bulletin, 37*(1), 107–119. https://doi.org/10.1177/0146167210391103

Schmitt, M. J., Neumann, R., & Montada, L. (1995). Dispositional sensitivity to befallen injustice. *Social Justice Research, 8*(4), 385–407. https://doi.org/10.1007/BF02334713

Schmitt, M., Gollwitzer, M., Maes, J., & Arbach, D. (2005). Justice sensitivity: Assessment and location in the personality space. *European Journal of Psychological Assessment, 21*(3), 202–211. https://doi.org/10.1027/1015-5759.21.3.202

Schmitt, M., Baumert, A., Gollwitzer, M., & Maes, J. (2010). The Justice Sensitivity Inventory: Factorial validity, location in the personality facet space, demographic pattern, and normative data. *Social Justice Research, 23*, 211–238. https://doi.org/10.1007/s11211-010-0115-2

Skinner, E. A. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology, 71*(3), 549–570. https://doi.org/10.1037/0022-3514.71.3.549

Skinner, E. A. (1995). *Perceived control, motivation, and coping.* Sage Publications.

Sommer, R. (2002). Personal space in a digital age. In R. B. Bechtel & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 647–660). John Wiley & Sons Inc.

Süssenbach, P., & Gollwitzer, M. (2015). Us(ed): The role of victim sensitivity in potentially exploitative intergroup relationships. *Group Processes & Intergroup Relations*, *18*(2), 241–255. https://doi.org/10.1177/1368430214556700

Thoemmes, F., MacKinnon, D. P., & Reiser, M. R. (2010). Power analysis for complex mediational designs using Monte Carlo methods. *Structural Equation Modeling, 17*(3), 510–534. https://doi.org/10.1080/10705511.2010.489379

Uzzell, D., & Horne, N. (2006). The influence of biological sex, sexuality and gender role on interpersonal distance. *British Journal of Social Psychology, 45*(3), 579–597. https://doi.org/10.1348/014466605X58384

van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expressions Set (ADFES). *Emotion, 11*(4), 907–920. https://doi.org/10.1037/a0023853

Velilla, S. (1993). A note on the multivariate Box-Cox transformation to normality. *Statistics and Probability Letters, 17*, 259–263. https://doi.org/10.1016/0167-7152(93)90200-3

Vohs, K. D., Baumeister, R. F., & Chin, J. (2007). Feeling duped: Emotional, motivational, and cognitive aspects of being exploited by others. *Review of General Psychology, 11*(2), 127–141. https://doi.org/10.1037/1089-2680.11.2.127

Wessa, P. (2016). *Box-Cox Normality Plot (v1.1.12) in Free Statistics Software (v1.2.1).* [Software]. http://www.wessa.net/rwasp_boxcoxnorm.wasp/

Wingenbach, T. S. H., Ashwin, C., & Brosnan, M. (2016). Validation of the Amsterdam Dynamic Facial Expression Set – Bath Intensity Variations (ADFES-BIV): A set of videos expressing low, intermediate, and high intensity emotions. *PLoS ONE, 11*(1), e0147112. https://doi.org/10.1371/journal.pone.0147112

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion, 18*(2), 129–166. https://doi.org/10.1007/BF02249397

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.