

Abstract

Leaders often have been shown to behave self-interestedly and toxically. Interestingly, followers do not always react to this toxic behavior. The current research argues that social identity is a crucial factor when investigating the boundary conditions for followers' reactions to toxic leaders. Followers who identify strongly with their group are more likely to react to toxic behaving leaders, as compared to low social identity. More specifically, we argue and show that followers act less cooperative and more revengeful toward a toxic leader if they identify with their group. Furthermore, this effect was mediated by negative emotions. Taken together, this study adds to the line of toxic leadership research by highlighting the importance of social identity when investigating boundary conditions for followers to react to toxic leaders.

Key words: boundary condition, cooperation, negative emotions, retribution, social identity, toxic leadership

The recent corporate scandals have increasingly showed that leaders often act in a self-interested and toxic manner (e.g. Leijonhufvud, 2009; Padilla, Hogan, & Kaiser, 2007; Einarsen, Aasland & Skogstad, 2007). Here we define toxic leadership, according to Lipman-Blumen's definition as "a global label for leaders who engage in numerous destructive behaviors and who exhibit certain dysfunctional personal characteristics. To count as toxic, these behaviors and qualities of character must inflict some reasonably serious and enduring harm on their followers and their organizations" (Lipman-Blumen, 2004, p. 18). Research indeed showed that leaders often act toxically thereby enhancing their self-interest and harming group members (Stouten, De Cremer & Van Dijk, 2005; De Cremer & Van Dijk, 2005). Although toxic leaders have been argued to be destructive to followers, it may well be that followers do not always react to their leader's behavior (Padilla et al., 2007). Indeed, Stouten and Tripp (2009) showed that toxic leaders were seen as entitled to act self-interestedly. Whereas the leader's toxic behavior was tolerated, other group members were held accountable for the same behavior. This would imply that followers sometimes do react negatively to toxic leader behavior and sometimes they do not. The present research builds upon this line of research by arguing that there are boundary conditions for group members' reactions to toxic leadership. Here, we argue that the group's social identity is an important aspect of whether group members will react toward their leader.

Social identity

In daily life, people identify with numerous social groups (social activities, work, or family). Ellemers, Kortekaas and Ouwerkerk (1999) argued that three components underlie this process of group identification: "A cognitive component (a cognitive awareness of one's membership in a social group-self-categorization), an evaluative component (a positive or

negative value connotation attached to this group membership-group self-esteem), and an emotional component (a sense of emotional involvement with the group-affective commitment)". Members from same social groups often evaluate and identify themselves in a similar way and have a clear definition of who they are and how they relate to and differ from other individuals and groups (Tajfel, 1981). The way groups are perceived is important for its members, because one's self is defined and evaluated in prototyped group categories (Ashforth & Mael, 1989). Thus, how a member feels about him or herself is directly impacted by the performance of its group. As a consequence, members of a group strive toward a positive social identity, thereby impacting their self-definition positively (Hogg & Turner, 1987; Hogg & Terry, 2000). The social identity of a group creates a collective concern for the group's welfare. Indeed, group members value their membership and are emotionally attached to the group (Tajfel & Turner, 1979; Ellemers et al., 1999). Research convincingly showed that group members who identify strongly with their group are more concerned for the group and are also committed to bring the group's cooperation to a higher level (Tyler & Blader, 2003). Any failure to follow through caring for the collective enterprise will be seen as unacceptable. Continuing this line of reasoning, one might expect that a group member who violates the group's cooperation could elicit negative reactions from the other group members such as the urge to retaliate and a lower desire to cooperate. Furthermore, given the fact that the leader of the group is often seen as the main authority and followers expect the leader to be a role model for legitimate behavior, leaders are expected to be held responsible if they fail not to do so. (Stouten, De Cremer & Van Dijk, 2005; Hogg, Martin & Van Knippenberg, 2003). More specifically, group members are expected to react more severely if the leader acts toxically in contrast to a follower.

Drawing from social identity theory which argues that group members commit their identity to the group, toxic behavior of one of the members is likely to create negative reactions (e.g.

Stouten et al., 2005). A primary reaction that can be expected from group members is the emergence of negative emotions. Indeed, emotions influence people which face situations inconsistent with their needs and expectations, such as the behavior of a toxic leader (Bies & Tripp, 2002; Brown & Trevino, 2006; Gross, 1998). Further, negative emotions have been argued to be a catalyst of behavior (Frijda, 1986; Kidd & Utne, 1978; Shaver, 1985; Weiner, 1985). Research indeed showed that cooperation and retribution are driven by strong emotional reactions that emerge as part of the actions of a transgressor (e.g. Bies & Tripp, 1996; Stouten, De Cremer & Van Dijk, 2006). We thus expect that negative emotions will mediate for the effect of social identity on cooperation and retribution.

Method

Design

The experiment consisted of a 2 (social identity: low vs. high) x 2 (toxic behavior: leader vs. follower) between factorial design. Participants were randomly assigned to the social identity and toxic behavior conditions.

Participants

Participants were 79 high school students (39 females and 40 males) and participated voluntarily. Their mean age was 15.7 years ($SD = 1.02$).

Procedure

The experiment was introduced to participants as a study regarding non verbal communication. Participants were invited in a classroom in groups of 12. Upon arrival, they were instructed to complete a task in silence. The task, called 'Jenga-challenge', would succeed if a Jenga tower of at least 25 cm within a 7 minute deadline was built. A Jenga

tower consists of 54 wooden, rectangular blocks. Typically, a group gathers around the Jenga tower and removes one block at a time, placing it on top of the tower and thereby creating a higher construction than the initial tower. When the group succeeded, each participant would receive a valuable incentive, that is, a movie ticket. If the task failed unintentionally (group members who would be careful in their handling the tower), no one would receive a movie ticket. However, if someone threw over the tower deliberately (an obvious and unambiguous act such as pushing the tower) that person would receive two movie tickets and the other participants would receive nothing.

Manipulation of social identity

After the experimenter's introduction, the salience of social identity was manipulated threefold. First, participants in the high social identity condition received a white t-shirt they all got to wear (Lount & Philips, 2007). Second, they were addressed by the experimenter as 'team white', referring to the color of their t-shirt, whereas participants in the low social identity condition were addressed as 'you' (Haslam, 2004; Haslam, Ryan, Postmes, Spears, Jetten & Webley, 2006). Third, participants in the high social identity condition were informed that their group would compete with groups from other schools in the area (Hogg & Terry, 2000). Concretely, the experimenter wrote fictional task results of the other groups on a blackboard in the social identity condition, whereas nothing would be written on the blackboard in the low social identity condition.

Manipulation of toxic behavior

In the leader condition, the experimenter asked whether someone wanted to lead the experiment. It was the confederate who raised her hand and said she had a year of experience leading youth groups. Then, if no one disagreed, she was announced as the leader and received a document with all further instructions for the experiment. The confederate read the

instructions out loud and gathered all participants around the table to initiate the experiment. In the follower condition, the experimenter did not ask whether someone wanted to lead the experiment. Instead, he provided the participants with all instructions needed to complete the experiment. In the leader and follower conditions, the confederate always threw over the tower after participants had built three blocks, thereby deliberately failing the task and therefore receiving two movie tickets (whereas the others would receive nothing). Throwing over the tower is a clear toxic behavior as the benefit is motivated from self-interest (e.g. Padilla et al., 2007).

Dependent Measures

After completion of the Jenga tower task, the dependent measures were solicited. Items were assessed on a 7-point Likert scale, ranging from 1 (= *not at all*) to 7 (= *very much so*), except the manipulation check for leadership. First, to check for the social identity manipulation, participants were asked “to what extent they felt connected to the other participants” (Crocker & Luthanen, 1992; Ashforth & Mael; 1989). Then, to check for the leadership manipulation, participants were asked “whether there was, according to them, a leader in the group”. This item was assessed on a binary scale (0: no; 1: yes).

Furthermore, participants were asked whether the task succeeded, failed unintentionally or failed intentionally. If they answered that the task failed intentionally, they were asked to skip a few questions and continue at the relevant section for their task result. Participants who indicated that the task failed intentionally, which all of them did, they were asked to complete the following questions. First, to measure negative emotions, four anger-related emotions were measured: frustration, restlessness, anger, and irritation (see also Stouten et al., 2005; 2006). These items were combined in an average anger scale (cronbach’s Alpha = .62). Second, the desire to cooperate with the person who threw over the tower (and won the

reward) was measured using the item “I want to cooperate with the transgressor in the near future”. Finally, retribution was measured using the item “I want to get back at the transgressor”.

To check whether any of the participants were aware of the intentionality of the confederate, participants were asked whether they noticed anything suspicious about the setup of the game and if so what exactly. None of them did.

Results

Manipulation checks

A 2 (social identity) x 2 (toxic behavior) ANOVA on the social identity manipulation check revealed a marginal main effect for social identity, $F(1, 75) = 2.87, p < .09$, showing that participants in the high social identity condition felt more connected ($M = 4.63, SD = 1.01$) with the other participants than in the low social identity condition ($M = 4.17, SD = 1.15$).

A logistic regression analysis on the leadership item showed that the percentage of participants who attributed the leader to be the transgressor was a function of the toxic behavior condition, $X^2(1, N = 75) = 35.34, p < .01$. In the toxic leader condition 80 % of participants reported that there was a leader, whereas 10 % in the toxic follower condition did so.

Negative Emotions

A 2 (social identity) x 2 (toxic behavior) ANOVA on the negative emotion scale showed a significant interaction, $F(1, 75) = 11.55, p < .01$ (Table 1). Further analyses revealed that participants in the toxic leader condition experienced more negative emotions in the high rather than low social identity condition, $F(1, 75) = 4.80, p < .03$. In the follower condition

participants experienced more negative emotions in the low social identity than in the high social identity condition, $F(1, 75) = 7.21, p < .01$.

Cooperation

A 2 (social identity) x 2 (toxic behavior) ANOVA on the extent that participants were willing to cooperate with the toxic member in the near future revealed a main effect for leadership, $F(1, 73) = 4.16, p < .05$, and a significant interaction, $F(1, 73) = 4.84, p < .03$ (Table 1). Further analyses showed that in the leader condition participants in the high rather than low social identity condition expressed a lower desire to cooperate in the future, $F(1, 73) = 5.05, p < .05$, whereas no effect for social identity could be revealed in the follower condition, $F(1, 73) = .97, p < .33$.

Retribution

A 2 (social identity) x 2 (toxic behavior) ANOVA on the extent that participants wished to get back at the toxic member showed a significant interaction, $F(1, 73) = 10.10, p < .01$ (Table 1). Further analyses revealed that participants in the leader condition wanted to get back more at their leader in the high rather than low social identity condition, $F(1, 73) = 14.79, p < .01$. In the follower condition no differences could be revealed for social identity, $F(1, 73) = .84, p < .36$.

Mediational Analyses

To examine whether negative emotions mediated the interactive effect of social identity and toxic behavior on cooperation and retribution, four regression analyses need to be performed (Baron & Kenny, 1986). First, the effects of the independent variables (social identity, toxic behavior, social identity x toxic behavior) on the dependent variables (cooperation, retribution) have to be tested. Second, the effects of the independent variables on the

proposed mediator (negative emotions) have to be tested. Third, the mediating variable, negative emotions, has to significantly influence the dependent variables. Fourth, the effect of the independent variable has to be reduced if the mediating variable is added.

Cooperation

First, a regression analyses of the independent variables on the desire to cooperate in the near future showed a significant interaction effect, $\beta = -1.36$, $p < .03$, mirroring the ANOVA results. Next, a regression analyses of the independent variables on the mediator negative emotions showed a significant effect for the interaction, $\beta = 1.82$, $p < .01$, reflecting the ANOVA results. Further, a regression analysis of the independent variables and the mediator on cooperation revealed a significant effect for the mediator, $\beta = -.31$, $p < .02$, and showed that the interaction effect between social identity and leadership was no longer significant, $\beta = -.82$, $p < .21$. A Sobel (1982) test showed that the indirect effect was significant ($z = 2.13$, $p < .03$) suggesting that negative emotions mediated the interactive effect of social identity and toxic behavior.

Retribution

Similarly, to test whether negative emotions mediated the interactive effect of social identity and toxic behavior on retribution, first, a regression analyses of the independent variables on retribution revealed a significant interaction effect, $\beta = 2.88$, $p < .01$, mirroring the ANOVA results. Further, a regression analysis of the independent variables and the mediator on retribution showed a significant effect for the mediator, $\beta = .90$, $p < .01$, and revealed that the interaction effect between social identity and leadership was no longer significant, $\beta = 1.23$, $p < .14$. The Sobel test showed that this reduction was significant ($z = 3.00$, $p < .01$) suggesting that negative emotions mediated the interactive effect of social identity and toxic behavior.

Discussion

Even though previous research showed that the toxic behavior of leaders elicits strong negative reactions (Stouten et al., 2005; 2006), it may well be that group members not always react to this (see Stouten & Tripp, 2009). The present research put forward that there are boundary conditions on people's reactions to a toxic leader. We argued and showed that the social identity of a group provides such a condition. As predicted, the findings of this experiment showed that group members react more severely toward a toxic leader if social identity is high rather than low. More specifically, it could be shown that followers from a group whose leader acted toxically and identified with their group reacted more negatively, wished to cooperate less, and were more likely to take revenge. Moreover, this was less the case if a group member acted in a toxic manner. Furthermore, emotional reactions were found to underlie the interaction effect between social identity and toxic behavior on cooperation and retribution. The most important findings will be discussed in following paragraphs.

First, the results illustrate that social identity is a boundary condition for group members to react to toxic behavior. More importantly, group members who identify with their group reacted more strongly to a toxic leader than a toxic group member. That is, our findings showed that if a leader acts toxically, group members who have a high social identity are less likely to cooperate and more inclined to take revenge. The performance of one's group is important for one's self-definition if one identifies with the group. In such a situation, a group leader elicits high expectations in order to act responsibly (Hogg, 2001). A leader failing to follow through by taking advantage of the situation may therefore harm group member's self-definition. If one of the other group members acted toxically reactions were not a function of social identity for the extent that participants wished to cooperate or take revenge on the transgressor. Interestingly, it did matter for people's negative emotional

reactions. That is, group members react more negatively if social identity was low rather than high. This suggests that if another group member acts toxically, this behavior is condemned less if there is a connection between group members. Hence, group members are willing to tolerate the action in this situation more. These findings might be explained by in-group favoritism, as social identity increases the affective connection between members. Social identity could then elicit forgiveness amongst group members, thereby condoning the toxic behavior (Hewstone, Rubin & Willis, 2002). We encourage future research to examine in more detail the effect of in-group favoritism on the reaction of group members toward toxic behaving followers.

Second, negative emotions were found to underlie the effect of social identity on the follower reaction. The mediational analyses draws attention to the important role that emotions have on the expression of retribution and desire to cooperate in the future. Indeed, the present research showed that group members are emotionally tolerant dependent on their social identity in the group. Research showed that the perceived fairness of an event largely influences one's emotional reactions (Mikula, Scherer & Athenstaedt, 1998). That is, the extent that group members identified with their group possibly affected how fair the behavior of the transgressor was and this determined their emotional response (e.g. Weiss, Suckow, & Cropanzano, 1999). Such emotional reactions frequently elicit a number of perceptual reactions as well as behavioral intentions toward transgressors (see also Goldberg, Lerner, & Tetlock, 1999; Schroeder, Steel, Woodell, & Bembenek, 2003). Our research then contributes to the understanding of how social identity salience affects the experience of negative emotions. Thus, social identity can be seen as an important component for the perception of negative emotions as a reaction to toxic behavior.

As with any research, there are some limitations. First, our manipulation of social identity could not be strongly confirmed in the manipulation check. This is surprising given that the

manipulation was based upon previous research. However, our results show clear interaction effects of social identity on emotions, cooperation and retribution, which provides confidence in our results.

Second, the participants in the present research primarily were high school students. Such a population has certain limitations when it comes to extrapolating the obtained results. Nevertheless, there is no reason to assume that high school students might differ from other populations such as college students or workers. Moreover, a strength of the current approach is the mundane realism in which the study was established. That is, participants in vivo experienced the tasks and procedure of this study (in contrast to for example computer-based experimental designs).

Third, participant groups always consisted out of 12 people. Therefore these results are not to be extrapolated to other group sizes. Indeed, Thomas & Fink (1963) argue that group size has an effect on many group dynamics and thus outcomes, such as the nature of interaction and member satisfaction. Therefore we encourage future research to vary in group size.

Fourth, the toxic behaving confederate was a woman. As such, our findings cannot be generalized to the situation were a male leader behaves in a toxic manner. Nevertheless, research investigating the role of the leader's gender on the evaluation by followers has only found small effects (Eagly, Makhijani & Klonsky, 1992).

Finally, in our study the leader was appointed in contrast to the random selection or election of the group leader. It may well be that group members' reactions may differ as a function of the appointment of the group leader. Research should further explore whether the results can be consistent depending on leader selection.

The major strength of this research is that our approach provided innovatory insight into the boundary conditions on people's reactions to a toxic leader. Toxic behaving leaders appear to

be more penalized by followers if social identity is high. As such, this research makes an important contribution to the endeavors of fencing off the perpetuation of toxic leadership.

References

- Ashforth, B., & Mael, F. (1989). Social identity and the organization. *Academy of Management Review*, *14* (1), 20-39.
- Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychology research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.
- Bies, R., & Tripp, T. (1996). Beyond distrust: "Getting even" and the need for revenge. In Kramer & T. R. Tyler (Eds.), *Trust in Organizations*, 246–260.
- Bies, R., & Tripp, T. (2002). Hot flashes, open wounds: Injustice and the tyranny of its emotions. In S.W. Gilliland, D.D. Steiner, & D. P. Skarlicki (Eds.), *Emerging perspectives on managing organizational injustice* (pp. 302-221). Greenwich, CT: Information Age.
- Brown, M., & Trevino, L. (2006). Ethical leadership: A review and future directions. *The Leadership Quarterly*, *6*, 595-616.
- Crocker, J., & Luhtanen, R. (1992). A collective self-esteem scale: Self evaluation of one's social identity. *Personality and Social Psychology Bulletin*, *18*, 302-318.
- De Cremer, D., & Van Dijk, E. (2005). When and why leaders put themselves first: Leader behavior in resource allocations as a function of feeling entitled. *European Journal of Social Psychology*, *34*, 553-563.
- Eagly, A., Makhijani, M., & Klonsky, B. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*, 3-22.
- Einarsen, S., Aasland, M., & Skogstad, A. (2007). Destructive leadership behavior: A definition and conceptual model. *The Leadership Quarterly*, *18*, 207-216.

- Ellemers, N., Kortekaas, P., & Ouwerkerk, J.W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology, 29*, 371-389.
- Frijda, N. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.
- Goldberg, J., Lerner, J., & Tetlock, P. (1999). Rage and reason: The psychology of the intuitive prosecutor. *European Journal of Social Psychology, 29*, 781-795.
- Gross, J. (1998). The emerging field of emotion regulation: An integrative view. *Review of General Psychology, 3*, 271-299.
- Haslam, A. (2004). *Psychology in organizations: The social identity approach*. Thousand Oaks: Sage Publications.
- Haslam, A., Ryan, M., Postmes, T., Spears, R., Jetten, J., & Webley, P. (2006). Sticking to our guns: Social identity as a basis for the maintenance of commitment to faltering organizational projects. *Journal of Organizational Behavior, 27*, 607-628.
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575-604.
- Hogg, M. (2001). A social identity theory of leadership. *Personality and Social Psychology Review, 3*, 184-200.
- Hogg, M., & Terry, J. (2000). Social identity and self-categorization processes in organizational contexts. *Academy of Management Review, 25* (1), 121-140.
- Hogg, M., & Turner, J. (1987). Intergroup behavior, self-stereotyping and the salience of social categories. *British Journal of Social Psychology, 26*, 325-340.
- Hogg, M., Martin, L., & Van Knippenberg, D. (2003). *Leadership and power: Identity processes in groups and organizations*. London: Sage Publications.
- Janis, I. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*. Boston: Houghton-Mifflin.

- Kidd, R., & Utne, M. (1978). Reactions to inequity: A perspective on the role of attributions. *Law and Human Behavior*, 2, 301-311.
- Leijonhufvud, A. (2009). Out of the corridor: Keynes and the crisis. *Cambridge Journal of Economics*, 33, 741-757.
- Lipman-Blumen, J. (2004). The allure of toxic leaders: Why we follow corrupt bosses and corrupt politicians – and how we can survive them. Oxford: Oxford University Press.
- Lount, R., & Philips, K. (2007). Working harder with the outgroup: The impact of social category diversity on motivation gain. *Organizational behavior and Human Decision Processes*, 103, 214-224.
- Mikula, G., Scherer, K., & Athenstaedt, U. (1998). The role of injustice in the elicitation of differential emotional reactions. *Personality and Social Psychology Bulletin*, 24, 769-783.
- Padilla, A., Hogan, R., & Kaiser, R. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly*, 18 (3), 176-194.
- Schroeder, D., Steel, J., Woodell, A., & Bembenek, A. (2003). Justice within social dilemmas. *Personality and Social Psychology Review*, 7, 374-387.
- Shaver, K. (1985). The attribution of blame: Causality, responsibility, and blameworthiness. New York: Springer-Verlag.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation model. *American Sociological Association*.
- Stouten, J., De Cremer, D., & Van Dijk, E. (2005). All is well that ends well, at least for proselves: Emotional reactions to equality violation as a function of social value orientation. *European Journal of Social Psychology*, 35, 767-783.

- Stouten, J., De Cremer, D., & Van Dijk, E. (2006). Violating equality in social dilemmas: Emotional and retributive reactions as a function of trust, attribution and honesty. *Personality and Social Psychology Bulletin*, 32, 894-906.
- Stouten, J., & Tripp, T. (2009). Claiming more than equality: Should leaders ask for forgiveness? *The Leadership Quarterly*, 20, 287-298.
- Tajfel, H. (1981). Human groups and social categories. *Studies of Social Science*. Cambridge University Press.
- Tajfel, H., & Turner, J. (1979). *An integrative theory of group conflict*. Oxford University Press.
- Thomas, E., & Fink, C. (1963). Effects of group size. *Psychological Bulletin*, 60, 371-384.
- Tyler, T., Blader, S. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and Social Psychology Review*, 4, 349-361.
- Weiner, B. (1985). An attribution theory of achievement motivation and emotion. *Psychological Review*, 92, 548-573.
- Weiss, H., Suckow, K., & Cropanzano, R. (1999). Effects of justice conditions on discrete emotions. *Journal of Applied Psychology*, 84, 786-794.

Appendix

Table 1

Means and standard deviations for negative emotions, cooperation, and retribution

	Toxic Leader				Toxic Follower			
	Low identity	social	High identity	social	Low identity	social	High identity	social
Dependent measure	M	SD	M	SD	M	SD	M	SD
Negative Emotions	3.67	0.76	4.56	1.29	4.54	0.98	3.64	1.23
Cooperation	2.65	1.37	1.75	1.32	1.33	0.49	1.80	1.51
Retribution	3.29	1.65	5.54	1.93	5.08	2.11	4.45	1.91

Note: Higher ratings for Negative emotions indicate a higher experience of negative emotions.

Higher ratings for cooperation indicate a higher desire to cooperate. Higher scores for retribution indicate a higher desire to get back at the transgressor.