
Article:
Tafvelin, S., Nielsen, K. orcid.org/0000-0001-9685-9570, Simonsen Abildgaard, J. et al. (3 more authors) (2019) Leader-team perceptual distance affects outcomes of leadership training: Examining safety leadership and follower safety self-efficacy. Safety Science, 120. pp. 25-31. ISSN 0925-7535

https://doi.org/10.1016/j.ssci.2019.06.019

Leader-Team Perceptual Distance Affects Outcomes of Leadership Training:

Examining Safety Leadership and Follower Safety Self-Efficacy

Susanne Tafvelin[1,2], Karina Nielsen[3], Johan Simonsen Abildgaard [4], Anne Richter[1,5], Ulrica von Thiele Schwarz[1,6] Henna Hasson[1,5],

[1] Medical Management Centre, Department of Learning, Informatics, Management and Ethics Karolinska Institutet, Sweden

[2] Department of Psychology, Umeå University, Sweden

[3] University of Sheffield, UK

[4] National Research Centre for the Working Environment, Copenhagen, Denmark

[5]Centre for Epidemiology and Community Medicine, Stockholm County Council, Sweden

[6] Mälardalen University

Author Note

*Susanne Tafvelin, Department of Psychology, Umeå University, 90187 Umeå, Sweden. E-mail: susanne.tafvelin@umu.se, phone: +46-705906668.

Karina Nielsen, Institute of Work Psychology, Sheffield University Management School, Sheffield S10 1FL, Sheffield, UK. E-mail: k.m.nielsen@sheffield.ac.uk

Johan Simonsen Abildgaard, National Research Centre for the Working Environment, Copenhagen, Denmark. E-mail:jss@arbejdsmiljoforskning.dk

Ulrica von Thiele Schwarz, Medical Management Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, 171 77 Stockholm, Sweden. E-mail: ulrica.schwarz@ki.se

Anne Richter, Medical Management Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, 171 77 Stockholm, Sweden. Email: anne.richter@ki.se

Henna Hasson, Medical Management Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, 171 77 Stockholm, Sweden. Email: henna.hasson@ki.se

*Corresponding author

This work was supported by AFA Försäkringar [Grant 160070]

Leader-Team Perceptual Distance Affects Outcomes of Leadership Training: Examining Safety Leadership and Follower Safety Self-Efficacy

Abstract

Whether leaders and their teams agree or not on perceptions of leadership has been found to impact follower well-being and performance. Less is known about how agreements or disagreements play a role in relation to safety and leadership training. The present study examined the effects of leaders' and followers' perceptual distance on safety leadership prior to a leadership safety training. Forty-eight leaders and a total of 211 followers from the paper industry completed surveys before and after training. Polynomial regression with response surface analyses revealed that the agreement between leaders and their followers regarding safety leadership before training was positively related to training outcomes including safety leadership and followers' safety self-efficacy. Line managers who overrated themselves on safety leadership before training had less favorable training outcomes. Our findings suggest that 360-degree feedback may not be sufficient for motivating leaders to change their behaviors during leadership training.

Keywords: safety leadership, leader-follower perceptual distance, leadership training, safety self-efficacy, polynomial regression, response surface analysis

1. Introduction

In an effort to achieve a proactive approach in the prevention of workplace injuries, organizations have turned towards key predictors of safety, such as leadership (e.g., Zohar, 2002). Given that studies consistently show that leadership is an important antecedent to employee safety perception, attitudes, and behaviours (Clarke, 2013; Kapp, 2012; Mullen, Hoffmeister mf från safety science), leadership training has been put forward as a promising strategy to improve workplace safety (Kelloway & Barling, 2010). However, a meta-analysis revealed only modest results in terms of the return on investment, and 34% of leadership training programs do not achieve their intended outcomes (Avolio, Reichard, Hannah, Walumbwa, & Chan, 2009).

One reason for these disappointing findings may be that leadership training programs have failed to consider essential psychological mechanisms (Avolio et al., 2009). First, effective leadership training requires not only passive learning but also changes in leaders' behaviors once leaders return to the workplace. These changes in leaders' behaviors should not only be observed by followers but also should lead to changes in followers (Avolio et al., 2009; Kirkpatrick, 1994). Second, the organizational context is critical to the successful transfer of learning from leadership training (Nielsen, Randall, & Christensen, 2010). On the one hand, followers may not welcome the changes that leaders attempt to implement post training. On the other hand, leaders may not perceive that any changes in their behaviors are required. Importantly, leaders and followers may not be in agreement about whether changes are needed. Leaders and followers need to have a shared mental model of the leader's behaviors before any training takes place; any disagreements prior to training may have profound effects on the training outcomes (Hasson, von Thiele Schwarz, Nielsen, & Tafvelin, 2016).

The aim of the present study was to understand the impact of agreement and disagreement on leaders' safety behaviors among leaders and followers pre-training and how these impact leaders' changes in safety behaviors and the changes in followers' safety self-efficacy post-training. The leadership training program aimed at improving both leaders' safety leadership and followers' safety self-efficacy, i.e., the extent to which followers feel empowered to proactively deal with safety issues. We contribute to the literature of safety leadership in two main ways. First, we use models of leader-team perceptual distance (i.e., the impact of agreement or disagreement between leader and their teams) to examine their relevance in relation to safety leadership and employee safety self-efficacy. Second, we examine if perceptual distance may serve as a psychological mechanism predicting outcomes of a safety leadership training, which is important to gain a fuller understanding of when training in safety leadership may succeed in achieving the intended outcomes.

2. Theoretical Background

Leaders and followers often differ in their perceptions of the leaders' behaviors (e.g., Atwater & Yammarino, 1992; Bass & Yammarino, 1991; Ostroff, Atwater, & Feinberg, 2004; Harris & Schaubroeck, 1988; Mabe & West, 1982; Van Velsor, Taylor, & Leslie, 1993) and other phenomena (e.g., goals, performance) in a workplace (Bashshur, Hernández, & González-Romá, 2011). Gibson, Cooper, and Conger (2009) proposed a model for studying perceptual distance between leaders and their followers and argued that differences in leaders' and followers' perceptions are detrimental to team performance because they hinder the team from maximizing collective cognition and reaching its full potential (Gibson et al., 2009). These differences are particularly problematic for leadership training programs because disagreement may mean that leaders and followers do not have a shared mental model regarding the current situation and the extent to which leaders need to change their behaviors (Hasson et al., 2016). Similarly, leaders who agree with their followers regarding their current

behaviors may be well suited to make the changes to their behaviors that their followers believe are needed.

2.1. Agreement as Sensemaking

A theoretical framework underpinning perceptual distance between leaders and followers is sensemaking (Weick, 1995). Sensemaking is crucially linked to the continuous negotiation of mental models and to achieving a collective mindset from which action can be taken (Weick, Sutcliffe, & Obstfeld, 2005). In the case of leadership training, a key prerequisite may be that leaders and followers agree in terms of their perceptions of the leaders' behaviors, i.e., having a shared mental model of the leaders' behaviors prior to training. When leaders' and followers' perceptions of the leadership behavior prior to a training are in disagreement, it is difficult for leaders to identify which behaviors they need to change due to the variations in leadership rating. This poses a challenge for reaching positive training outcomes. In particular, shared mental models have been shown to be relevant for making accurate and appropriate safety leadership judgements (Weick, 1989, 1993) and for sustaining a vigilant awareness of and collective focus on making sense of potential risks (Weick, 2010), i.e., leaders' continuously working toward taking actions to ensure safe environments. In the present study, we bridged the gap between the differential or shared mental models of the leaders' behaviors to ensure good safety leadership and to promote followers' safety self-efficacy, which answers recent calls by Zohar (2010) and Zohare & Luria (2005) suggesting to link sensemaking research to collective safety phenomena. We measured two types of safety self-efficacy: followers' perceived ability to give safety-specific feedback and their ability to proactively intervene when they observe unsafe work practices. We expected these three outcomes to be outcomes of this leadership training program due to the content of the training: Leaders were given feedback in their safety leadership behaviors,

and these behaviors were found to be related to followers' safety-specific self-efficacy (Katz-Navon, Naveh, & Stern, 2007).

2.2 Hypothesis Development

Previous research revealed that agreement between leaders and their followers is related to better work outcomes for both leaders and followers (Fleenor, Smither, Atwater, Braddy, & Sturm, 2010; Hasson, Tafvelin, & von Thiele Schwarz, 2013; Ostroff, Shin, & Kinicki, 2005; Tafvelin, von Thiele Schwarz, & Hasson, 2017). More specifically, in the leadership literature, leaders' and their teams' agreeing on the leaders' behaviors has been related to higher follower-rated leader performance as compared with those who disagreed (Atwater & Brett, 2005; Atwater & Yammarino 1992; Bass & Yammarino, 1991; Furnham & Stringfield 1994; Ostroff et al. 2004) and better follower outcomes in terms of job satisfaction and organizational commitment (Szell & Henderson, 1997). In a similar manner, studies in the area of perceptual distance showed that agreement between a leader and a team regarding, for instance, goal accomplishment and organizational support was associated with increases in team performance (Bashshur et al., 2011; Gibson et al., 2009). A limitation of these studies was that although they explored the impact of perceptual distance over time, they failed to explore how perceptual distance may influence leadership training outcomes. To the best of our knowledge, only a few studies have explored the challenges of having followers' and leaders' perceptions not aligned in leadership training programs. Nielsen and Daniels (2012) found that leaders who were trained in how to develop and implement teamwork reported less challenging work and poorer well-being if they returned to a group of followers who did not welcome the idea of implementing teams, compared with the control group. Hasson et al. (2016) studied agreement in relation to a leadership training program and found that when leaders and followers agreed on the pre-intervention organizational learning climate, the higher the followers rated the climate post intervention. In the present study, we extended

previous research to suggest that in teams where leaders and followers agree on the leaders' safety behaviors pre-training, leaders are more likely to understand the need for change, and thus, we will see greater improvements in the outcomes of the leadership training. We therefore developed the following hypotheses:

Hypothesis 1a: The higher the agreement between leaders and their followers on leaders' safety leadership behaviors pre-training, the greater the improvements in leaders' safety behaviors post training.

Hypothesis 1b: The higher the agreement between leaders and their followers on leaders' safety leadership behaviors pre-training, the greater the improvements in followers' self-efficacy in giving feedback on safety issues post training.

Hypothesis 1c: The higher the agreement between leaders and their followers on leaders' safety leadership behaviors pre-training, the greater the improvements in followers' self-efficacy in intervening against unsafe practices post training.

Studies on agreement between leaders and followers have also found that the level of ratings is important for outcomes (Fleenor et al., 2010). When leaders' and followers' ratings were high and in agreement (rather than low and in agreement), the performance outcomes were greater (Bashshur et al., 2011, McKay et al., 2009). In a similar manner, the study that Hasson et al. (2016) conducted showed that organizational learning improved more in teams where leaders and their followers agreed, and reported high levels of organizational learning climate pre-training. Transferring these results to the safety leadership training context implies that when both leaders and teams have high ratings of the leaders' safety behaviors pre-training, leaders are able to change their behaviors to those needed to meet followers' need for safety leadership. We therefore developed the following hypotheses exploring the impact of the level of safety leadership behaviors.

Hypothesis 2a: Followers' ratings of leaders' safety leadership behaviors post training

will be highest when the leaders' and the followers' perceptions of these behaviors are high and aligned pre-training rather than aligned and low.

Hypothesis 2b: Followers' ratings of their self-efficacy in giving safety feedback will be highest when the leaders' and the followers' perceptions of the leaders' safety leadership behaviors are high and aligned pre-training rather than aligned and low.

Hypothesis 2c: Followers' ratings of their self-efficacy in intervening against unsafe practices will be highest when the leaders' and the followers' perceptions of leaders' safety leadership behaviors are high and aligned pre-training rather than aligned and low.

If pre-intervention disagreement exists between the leader and the team, it becomes important to understand how these disagreements influence the outcomes of the leadership training. If leaders rate themselves higher than their followers do, they may make few attempts to improve their safety leadership behaviors, and without any changes in behaviors, they are unlikely to influence their followers' safety-specific self-efficacy. Previous studies revealed less favorable outcomes when leaders' ratings were higher than those of their followers, that is, when the leaders overestimated their leadership and other work-related factors. More specifically, leaders who rated their leadership (Van Velsor et al. 1993), organizational support climate (Bashshur et al. 2011), power distance (Cole, Carter, & Zhang 2013), implementation climate (Aarons, Ehrhart, Farahnak, Sklar, & Horowitz, 2017), and organizational learning (Hasson et al 2016) higher that their followers did receive lower performance ratings from followers compared with leaders who underestimated or were in agreement with their followers. Bashshur et al. (2011) suggested that when a leader's ratings are higher than his/her team's ratings, this may result in passive leadership because the leader fails to understand the needs of the team. This situation may be troublesome for followers not only because they are exposed to leaders who do not exert safety leadership behaviors but also because their leaders post training take no corrective action to improve their leadership

behaviors. Based on these findings, we hypothesized the following:

Hypothesis 3a: Followers' ratings of their leaders' safety leadership behaviors will improve less post-training when the leaders' ratings of their safety leadership behaviors are greater than their followers' ratings are pre-training, rather than when the leaders' ratings are lower than the followers' ratings are.

Hypothesis 3b: Followers' self-efficacy in giving safety feedback will improve less post-training when the leaders' ratings of their safety leadership behaviors are greater than their followers' ratings are pre-training, rather than when the leaders' ratings are lower than the followers' ratings are.

Hypothesis 3c: Followers' self-efficacy in intervening against unsafe practices will improve less post-training when the leaders' ratings of their safety leadership behaviors are greater than their followers' ratings are pre-training, rather than when the leaders' ratings are lower than the follower' ratings are.

3. Method

The present paper is based on data from a natural experiment; a leadership training program in the Swedish forest industry. Pre-training measures were collected in November 2011 with a follow-up in April 2013. The training program took place between December 2011 and March 2013 and targeted all leaders in the organization. The company initiated the training program, and organizational psychologists from the company's occupational health service designed and delivered it. A team of researchers was tasked to evaluate the effects of the training program on leadership, safety, and organizational learning. The effects of the intervention on organizational learning, leadership, and safety were previously reported (references withheld for the sake of anonymity).

3.1 Procedure

The company employed about 800 followers and 101 leaders. All leaders were expected to participate in the leadership training program, and prior to training, they were asked to answer a Web-based questionnaire. They were also asked to invite five followers (direct reports) to provide a feedback assessment of their leadership behaviors. The instruction was to include both individuals to whom they felt close and individuals who were more distant. Thus, the data from leaders covered the whole population, whereas the followers represented a sample of followers, whom each leader purposefully selected. All respondents received two emails: one from Human Resources giving a general introduction to the aim of the program and data collection, and one jointly from the researchers and consultants containing an introductory letter describing the study in more detail as well as a personal link to the questionnaire. It was emphasized that responding to the survey was voluntary, and all respondents provided informed consent for their data to be used in research. Two reminders were sent out during the three weeks the data collection was ongoing.

3.2 Participants

In total 101 leaders, including 56 line leaders and 45 technical engineers, participated in the training. In the current study we included the line leaders only. Line leaders were defined as the management level directly above non-managerial workers, assuming responsibility for employee and health and safety issues. As technical engineers had no role in working with safety, these were excluded from the current study. Out of the 56 line leaders, 54 answered the baseline questionnaire and consented for the data to be used in research. Of these, 48 line leaders answered the follow-up questionnaire and had a sufficient number of followers who responded to the surveys of relevance in this study. Among the line leaders, 76% were men, the mean age was 47 years (*SD* 8.2), and they had spent six years working in their current positions (*SD* 6.3) and 20 years (SD 11.1) in their organizations.

Of the 240 followers who were invited, 158 responded to the questionnaire and had sufficient data on the variables of interest in this study. Thus, the response rate in the employee sample was 67%. In the final follower sample, the majority (75%) were male, the mean age was 47 years (SD = 8.9), and they had a mean tenure in the organization of 23 years (SD = 10.9 years).

3.3 The intervention

The leadership training program included in total 20 days of training, including both didactic and experiential learning activities aimed at improving both theoretical knowledge and practical skills. The main goal of the training was to improve leadership, safety, and learning. The leaders were divided into cross-departmental groups consisting of 20 individuals. The first block was mainly theoretical and focused on teaching leadership, organizational change, and follower motivation. In addition to lectures and discussions, this block included feedback and action planning based on the 360-degree feedback assessment of the leader's safety leadership and transformational leadership. The second block was focused on skill training, focusing on leadership behaviors, facilitating behavior change among followers, giving feedback, and coordinating activities. Each leader identified an area in which he or she wanted to improve and used this as a case during the training program. Examples of areas for improvement include safety leadership for a specific work station, collective leadership, collaboration, and information sharing on the team.

3.4 Measures

Safety leadership behavior was measured with the *supervisory action to ensure safety at work* from the Group Safety Climate Scale (Zohar, 2000). Example items are "My supervisor seriously considers any worker's suggestion for improving safety" and "My supervisor gets annoyed with any worker who ignores the safety rules, even minor rules."

This is a five-item scale rated on a five-point scale from “completely agree” to “completely disagree.” Thus, for safety leadership behaviors, a high value indicates positive safety leadership. The Cronbach alpha (Cortina, 1993)of the supervisory action subscale was .77 at Time 1 and .79 at Time 2.

Safety self-efficacy was measured with two three-item scales, one for feedback and one for safety intervention (Pettinger, 2000). Safety feedback was measured with the overall question of “How comfortable are you with giving the following persons feedback regarding their personal safety?” Employees then rated on a 10-point Likert scale the extent to which they would give feedback to (1) a colleague, (2) someone else at work, and (3) a supervisor or manager. The coefficient alpha for this scale was .81 at Time 1 and .64 at Time 2. Safety intervention was measured in the same manner with the overall question of “How comfortable are you with stopping the following individuals if you think they are acting unsafely?” The Cronbach alpha for this scale was .85 at Time 1 and .77 at Time 2.

3.5 Analyses

We used polynomial regression with response surface analysis (Edwards, 1994) to test the impact of agreement and disagreement among leaders and their followers on leaders’ safety behaviors and followers’ safety-specific self-efficacy. This analysis allows for a combination of two variables to be related to an outcome while retaining information about the differences between the variables. The data analytic approach involved aggregating employee ratings of leaders (K= 48) to the team level to make team level inferences about relationships among variables. Leader ratings of leadership were already at the team level. To justify aggregation of the employee data to the team level, intraclass correlation coefficients (ICC: Bliese, 2000) and within group agreement (rWG(j): James, Demaree, & Wolf, 1984) statistics were calculated; these are presented in Table 1. ICC reflects the proportion of the

variance that is at the team level (McGraw & Wong, 1996), while within group agreement reflects the consensus in the scores that respondents working in the same team provide (James, Demaree & Wolf, 1993) Overall, the analyses support the aggregation of team ratings.

(Insert Table 1 about here)

As recommended when studying leader-follower perceptual distance (Gibson et al., 2009), we followed the three-step procedure by Shanock et al. (2010). First, agreement and disagreement between leaders and followers was investigated to ensure that the level of disagreement was sufficient for proceeding with further analysis. The minimum level was set to 10% disagreement, defined as at least 0.5 SD of the standardized mean score on the two predictors, as Fleenor et al. (1997) suggested. Second, polynomial regression analysis was conducted, one for each of the three outcome variables. This analysis enables us to examine the combined impact of two variables on a third, but at the same time retaining information about the differences between the variables. It is the recommended type of analysis to examine perceptual distance (Edwards, 2002; Shanock et al., 2010) as it keeps leader ratings and team ratings separate, at the same time as also incorporating higher order terms such as squared and interaction terms which enables tests of more elaborates effects (Humberg et al., 2018) The polynomial regressions were performed on scale-centered variables to facilitate the interpretation of the findings (Edwards, 1994). The outcomes were regressed on leaders' ratings, followers' ratings, the cross product of leaders' and followers' ratings, and the square of leaders' and followers' ratings of the leaders' safety leadership behaviors. If the predictors explain variance in the outcome variable, R^2 of the polynomial regression is significant, and further analysis is justified. This includes conducting four surface tests: a_1, a_2, a_3, and a_4, based on unstandardized regression coefficients (Atwater et al., 1998; Edwards, 2002). Third, the surface test values are plotted in graphs (see Figures 1-3), and the graphs are interpreted.

The four surface test values represent the slopes and curvature of two lines. The first line runs diagonally from the nearest to the farthest corners of the graph. This is called the line of perfect agreement. *a1* is the slope and represents how agreement between the predictors relates to the outcome. *a2* is the curvature and shows whether this relationship (between the agreement and outcome) is linear or non-linear, that is, if the outcomes differ depending on whether the ratings are high and in agreement or low and in agreement. The second line runs diagonally from the left to the right corner. It is called the "line of incongruence," where the slope is reflected by *a3* and the curvature by *a4*. Similarly, regarding the line of perfect agreement, the curvature shows how disagreement between predictors relates to the outcome and the slope if the direction matters.

4. Results

4.1. Preliminary Analysis

In Table 2, descriptive statistics and intercorrelations of the studied variables are presented. The correlations between followers' and leaders' rated safety leadership pre-training are non-significant, indicating that variation exists between the ratings of teams and leaders, which suggests that perceptual distance analyses are justified. At baseline, the two self-efficacy measures (providing safety related feedback and intervening against unsafe work practices) are correlated; however, this relation is not found at Time 2.

(Insert Table 2 about here)

4.2 Polynomial Regression with Response Surface Analysis

First, the analyses of agreement between the leaders' and the followers' perceptions of safety leadership showed that 23% were in agreement with their followers in their perceptions of safety leadership, while 38% of the leaders underestimated their safety leadership and approximately 40% over-estimated their safety leadership. Overall, the

discrepancy in leader and follower ratings on safety leadership behavior was larger than 10%, thus indicating that polynomial regressions are warranted for analyzing the perceptual distance represented in the data.

In the second step, a number of polynomial regression analyses were performed, see Table 3. Perceptions of safety leadership before training explained significant variance in followers' perceptions of safety leadership and self-efficacy to intervene against unsafe practices after training, whereas the explained variance in follower self-efficacy in giving safety feedback after training approached significance ($p = 0.079$). The range of explained variance varied between 13 and 43%. Based on these findings, we decided to calculate the surface test values, *a1-a4*, for all three outcomes (see Table 3).

In the third and final step, the surface test values were used to graph and thereby interpret the results. The surface test values, *a1-a4*, represented the slope and curvature of the two lines comprising the response surface pattern in the graph: the line of perfect agreement and the line incongruence. For the development of safety leadership, *a1* was significant supporting Hypotheses 1a and 2a. This suggests that when leaders and their followers' perceptions on safety leadership behaviors are in agreement before training, followers' ratings of safety leadership post training increase. As seen in Figure 1, the ratings of safety leadership after training increases along the line of perfect agreement, from the front right corner to the back left corner of the graph, where followers and leader–rated safety leadership are aligned. In addition, the lowest values are at the front of the graph, where both follower- and leader-rated safety leadership are low. In addition, a significant and negative *a3* value was found, suggesting that when leaders' ratings of safety leadership are higher than their followers' are before training, the followers perceive less improvement in safety leadership after training. These findings are in line with Hypothesis 3a, and the graph in Figure 1 show how followers' perceptions of safety leadership after training are low when leaders' ratings of their own

safety leadership are high and their followers' ratings are low pre-training (i.e., in the right back corner of the graphs).

Regarding followers' increased self-efficacy in giving safety feedback after training, the findings suggest that agreement on safety leadership pre-training is important, as *a1* and *a2* were significant supporting Hypothesis 1b. However, Hypothesis 2b was not supported given the negatively significant *a2,* which indicates that the relationship between agreement and followers' increased safety feedback self-efficacy post training is not linear but curvilinear, i.e., the positive slope of agreement decreases with higher values for safety leadership; see Figure 2. Hypothesis 3b was not supported given that the direction of disagreement between leaders and their followers did not matter regarding self-efficacy safety feedback (i.e., a nonsignificant *a3*).

Finally, for followers' changes in self-efficacy in intervening against unsafe work practices, no support was found for Hypotheses 1c and 2c, that is, the importance of agreement. On the contrary, leaders' disagreement was related to increases in followers' self-efficacy in intervening against unsafe behavior. As can be seen in Figure 3, high values for intervening can be found when leaders' ratings of safety leadership are high and the followers' ratings are low and vice versa. However, *a3* was negative and significant supporting Hypothesis 3c.

(Insert Table 3 and Figures 1-3 about here)

5. Discussion

In the present study, we explored the effects of pre-intervention agreement or disagreement on leaders' safety leadership on the outcomes of a leadership training program. Building on sensemaking theory (Weick, 1995) we proposed that it is important for leaders and followers to have a shared mental model of the leaders' behaviors before training for the

leaders to improve their safety leadership and improve followers' safety self-efficacy. In cases where leaders and followers disagreed on the leaders' safety leadership behaviors pre-training, we hypothesized that when leaders rated themselves higher than their followers did, this would have the most negative impact, as leaders would ignore their followers' ratings and would feel little need to change.

Hypotheses 1a and b were supported in the expected direction; when leaders and followers agreed, leaders' safety leadership behaviors and followers' self-efficacy to give safety-related feedback improved. Interestingly, the results also indicated that disagreements regarding safety leadership behaviors were important for followers' self-efficacy in proactively intervening when observing unsafe work practices. The importance of high and aligned ratings were significant only for leadership safety behavior (Hypothesis 2a). For safety self-efficacy feedback (Hypothesis 2b), the relationship was curvilinear, meaning that the positive effect flattens out the higher the leadership ratings. Our findings corroborate previous studies of perceptual distance (Bashshur et al., 2011; Gibson et al., 2009) but extend them into demonstrating the importance of agreement also for safety outcomes of leadership training.

Hypothesis 3 concerned the impact of disagreements when leaders overrated themselves compared with their followers' ratings. More specifically, Hypothesis 3a, that leaders who overrate their safety leadership behaviors compared with their followers will have followers who report fewer improvements in their leaders' safety leadership behaviors post training, was supported. Overestimation was also problematic for improvement in follower outcomes in terms of safety self-efficacy in intervening when observing unsafe practices, supporting Hypothesis 3c. We thereby extend previous findings of the harm of overrating yourself as a leader (Van Velsor et al., 1993) by showing that it not only hurts employee performance but also leaders' development during leadership training.

5.1 Theoretical and Practical Implications

The findings in our study have a number of implications for both research and practice. From a research perspective, our findings answer the calls for leadership research to identify psychological mechanisms that can explain when and why leadership training is successful or not (Avolio et al., 2009). Based on theories of sensemaking (Weick, 1995) our findings suggest that one such psychological mechanism is the extent to which leaders and their teams agree on the initial level of safety leadership before training. The level of agreement will then influence both the development of safety leadership during training, and employees' post training safety-self efficacy. In addition, when leaders overrated themselves in comparison with their followers' ratings, followers reported fewer improvements in their leaders' behaviors post training. In sum, our study suggest that perceptual distance between leaders and teams influence not only important employee outcomes (which has been the focus in previous research), but may also serve as an important mechanism that explains when and why leadership training is effective. We found that the proximal outcome of followers' perceived ratings of their safety leadership behaviors were more often supported compared with the more distal outcome of employee safety self-efficacy. This is perhaps not surprising considering the leadership training program targeted this outcome specifically. The results suggest the importance of conducting this type of analysis as a sense check: What is the impact of leader-follower agreement or disagreement on the immediate training outcomes?

For practice, our findings suggest that, although many leadership training programs include 360-degree feedback in an attempt to provide leaders with feedback on what they need to improve (Fleenor et al., 2010) such feedback may not be efficient. Leaders who rate themselves highly may not take followers' feedback and instead rely more on their own ratings. These results raise the question of the content of leadership training and what feedback and input should be given to leaders to motivate them to change their behaviors. We

also found that leaders and followers who had high ratings that were aligned had a curvilinear effect on followers' perceived ability to give safety-related feedback. A possible explanation for this might be that these followers do not see much room for giving feedback due to leaders' high attention to safety compared with when leaders and followers agree that leaders enact more moderate safety leadership behaviors. It thus becomes important also to understand the baseline levels to estimate the realistic outcomes of training.

For self-efficacy in intervening against unsafe work practices, we found the counterintuitive result that disagreement led to more willingness to intervene. A tentative explanation may be that when leaders and followers are unaligned, the possibility of followers' assuming informal responsibility and intervening to ensure their colleagues' safe working practices is higher. Interestingly, we found different results for the two safety self-efficacy outcomes. Agreement was important for followers when it came to giving safety feedback, whereas disagreement was found to be important for followers in terms of feeling able to intervene. Future research should explore these differences.

5.2 Methodological Considerations

Some limitations must be considered. First, we did not have a control group, meaning we cannot clearly attribute changes in outcomes to the intervention. Another limitation was the selection of followers, where leaders were asked to invite five of their followers to participate in the survey, which is a common way of recruiting survey respondents in leadership training studies involving 360-degree evaluations (Fleenor et al., 2010). We cannot be sure how they selected these staff members, and hence, the findings might be biased. However, the leaders did receive instructions to include both staff members to which they felt close and those from which they felt more distant, and it is clear that substantial disagreement occurred between leaders and teams on the variables of interest, indicating that not just individuals with whom the leaders were in agreement were recruited to participate in the

study. This study also focused specifically on safety outcomes, and other health-related outcomes need to be evaluated to illuminate the impact of perceptual distance across a broader range of outcomes. Finally, the study comprised only one company; more studies from several organizations are needed to further examine the generalizability of the findings. Concurrently, the study had several strengths. Among them was the use of a longitudinal intervention design, which is rare both in leadership training evaluations (Avolio et al., 2009) and in the perceptual distance literature (Bashshur et al., 2011). Another strength was the use of more sophisticated statistical analysis than that used in most previous studies on leader-follower agreement, which allowed the three-dimensional relationship rather than only the two-dimensional relationship to be studied (Edwards, 2002). These analyses made it possible to detect curvilinear relationships, and it mattered if ratings were in agreement and high or in agreement and low.

6. Conclusions

Leadership training is a top priority in many companies and may be a way of ensuring that leaders are equipped to understand and enforce safety priorities. Using feedback from 360-degree surveys is common in leadership training, building on the assumption that having a shared understanding of how the leader behaves is important both in itself and as a motivator for change. This study showed that the perceptual distance between leaders and followers before training has a significant impact on the outcomes of the training in that agreement is generally beneficial and particularly when both leaders and followers rate the leadership favorably. The present study also showed that when leaders over-rate themselves, feedback on discrepancy does not change the fact that they see little improvement in outcomes over time. This raises the question of whether 360-degree feedback is helpful for the group that would need it the most.

References

Aarons, G. A., Ehrhart, M. G., Farahnak, L. R., Sklar, M., & Horowitz, J. (2017). Discrepancies in leader and follower ratings of transformational leadership: Relationship with organizational culture in mental health. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(4), 480-491.

Alimo-Metcalfe, B., Alban-Metcalfe, J., Bradley, M., Mariathasan, J., & Samele, C. (2008). The impact of engaging leadership on performance, attitudes to work and wellbeing at work: A longitudinal study. *Journal of Health Organization and Management*, *22*(6), 586-598.

Atwater, L. E., & Brett, J. F. (2005). Antecedents and consequences of reactions to developmental 360 feedback. *Journal of Vocational Behavior*, *66*(3), 532-548.

Atwater, L.E., & Yammarino, F. J. (1992). Does self- other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology, 45*(1), 141-164.

Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly*, *20*(5), 764-784.

Bashshur, M.R., Hernández, A., & González-Romá, V. (2011). When managers and their followers disagree: A longitudinal look at the consequences of differences in perceptions of organizational support. *Journal of Applied Psychology, 96*(3), 558-573.

Bass, B.M., & Yammarino, F. J. (1991). Congruence of self and others' leadership ratings of naval officers for understanding successful performance. *Applied Psychology, 40*(4), 437-454.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein, & S. W. J. Kozlowski

(Eds.), Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions (pp. 349–381). San Francisco, CA: Jossey-Bass.

Cole, M. S., Carter, M. Z., & Zhang, Z. (2013). Leader–team congruence in power distance values and team effectiveness: The mediating role of procedural justice climate. *Journal of Applied Psychology*, *98*(6), 962.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104. http://dx.doi.org/10.1037/0021-9010.78.1.98

Edwards, J. R. (1994). Regression analysis as an alternative to difference scores. *Journal of Management*, *20*(3), 683-689.

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression and response surface methodology. *Advances in Measurement and Data Analysis*, 350-400.

Fleenor, J. W., McCauley, C. D., & Brutus, S. (1996). Self-other rating agreement and leader effectiveness. *The Leadership Quarterly*, *7*(4), 487-506.

Fleenor, J.W., Smither, J.W., Atwater, L.E., Braddy, P.W., & Sturm, R.E. (2010). Self-other rating agreement in leadership: A review. *The Leadership Quarterly, 21*(6), 1005–1034.

Furnham, A., & Stringfield, P. (1994). Congruence of self and subordinate ratings of managerial practices as a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*, *67*(1), 57-67.

Gibson, C., Cooper, C. D., & Conger, J. A. (2009). Do you see what we see? The complex effects of perceptual distance between leaders and teams. *Journal of Applied Psychology, 94*(1), 62-76.

Harris, M. M, & Schaubroeck, J. (1988). A meta- analysis of self- supervisor, self- peer, and peer- supervisor ratings. *Personnel Psychology, 41*(1), 43-62.

Field Code Changed

Formatted: German (Germany)

Hasson, H., Tafvelin, S., & von Thiele Schwarz, U. (2013). Comparing Employees and Managers' Perceptions of Organizational Learning, Health, and Work Performance. *Advances in Developing Human Resources, 15*(2), 163-176.

Hasson, H., Thiele Schwarz, U., Nielsen, K., & Tafvelin, S. (2016). Are we all in the same boat? The role of perceptual distance in organizational health interventions. *Stress and Health*, *32*(4), 294-303.

Humberg, S., Nestler, S., & Back, M. D. (2018). Response surface analysis in personality and social psychology: Checklist and clarifications for the case of congruence hypotheses. *Social Psychological and Personality Science*, Advance online publication, doi: 10.1177/1948550618757600.

James, L. R., Demaree, R. G.,&Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. Journal of Applied Psychology, 69(1), 85–98.

L.R. James, R.G. Demaree, G. Wolf (1991). r_{WG}: an assessment of within-group interrater agreement The Journal of Applied Psychology, 78, 306-309.

Katz-Navon, T., Naveh, E., & Stern, Z. (2007). Safety self-efficacy and safety performance: Potential antecedents and the moderation effect of standardization. *International Journal of Health Care Quality Assurance*, *20*(7), 572-584.

Kirkpatrick, D. L.(1994): Evaluating Training Programs. *The Four Levels. San Francisco*.

McKinsey (2012). The State of Human Capital 2012—False Summit: Why the Human Capital Function Still Has Far to Go, a joint report from The Conference Board and McKinsey, October 2012, mckinsey.com.

Mabe, P. A., & West, S.G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30-46. doi:10.1037/1082-989X.1.1.30

Formatted: German (Germany)

McKay, P.F., Avery, D.R., & Morris, M.A. (2009). A tale of two climates: diversity climate from subordinates' and managers' perspectives and their role in store unit sales performance. *Personnel Psychology, 62*(4), 767-791. doi: 10.1111/j.1744-6570.2009.01157.x

Nielsen, K., & Daniels, K. (2012). Does shared and differentiated transformational leadership predict followers' working conditions and well-being?. *The Leadership Quarterly*, *23*(3), 383-397.

Nielsen, K., Randall, R., & Christensen, K. B. (2010). Does training managers enhance the effects of implementing team-working? A longitudinal, mixed methods field study. *Human Relations*, *63*(11), 1719-1741.

Ostroff, C., Shin, Y., & Kinicki, A.J. (2005). Multiple perspectives of congruence: Relationships between value congruence and employee attitudes. *Journal of Organizational Behavior, 26*(6), 591-623.

Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding self-other agreement: A look at rater and ratee characteristics, context, and outcomes. *Personnel Psychology, 57*(2), 333-375.

Pettinger, C.B., (2000). Improving occupational safety & health interventions: A comparison of safety self-efficacy and safety stagees of change. Virginia Polytechnic Institute & State University, Blacksburg, Virginia

Szell, S, & Henderson, R. (1997). The impact of self-supervisor/subordinate performance rating agreement on subordinates' job satisfaction and organisational commitment. *Journal of Applied Social Behaviour, 3*(2), 25-37.

Tafvelin, S., von Thiele Schwarz, U., & Hasson, H. (2017). In agreement? Leader-team perceptual distance in organizational learning affects work performance. *Journal of Business Research*, *75*, 1-7.

Van Velsor, E., Taylor, S., & Leslie, J.B. (1993). An examination of the relationships among self- perception accuracy, self- awareness, gender, and leader effectiveness. *Human Resource Management, 32*(2- 3), 249-263.

Weick, K. E. (1989). Mental models of high reliability systems. *Organization & Environment*, *3*(2), 127–142.

Weick, K. E. (1993). The collapse of sensemaking in organizations: The mann gulch disaster. *Administrative Science Quarterly*, *38*(4), 628. https://doi.org/10.2307/2393339

Weick, K. E. (1995). *Sensemaking in Organizations*. Thousand Oaks: Sage.

Weick, K. E. (2010). Reflections on enacted sensemaking in the Bhopal disaster. *Journal of Management Studies*, *47*(3), 537–550. https://doi.org/10.1111/j.1467-6486.2010.00900.x

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science*, *16*(4), 409–421.

Zohar, D. (2000). A group-level model of safety climate: testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology*, *85*(4), 587.

Zohar, D. (2010). Thirty years of safety climate research: Reflections and future directions. *Accident Analysis & Prevention*, *42*(5), 1517-1522.

Zohar, D., & Luria, G. (2005). A multilevel model of safety climate: cross-level relationships between organization and group-level climates. *Journal of Applied psychology*, *90*(4), 616.

Table 1

Subscale intraclass correlation coefficients (ICC) and within group agreement

Subscale	Number of items	ICC	Mean rWG (j)
Safety leadership	5	.20*	.80
Safety self-efficacy feedback	3	.10	.72
Safety self-efficacy intervention	3	.09*	.79

Note. *K*= 48 leaders, * p< .05

Table 2
Descriptive statistics and Pearson correlations among all study variables

	Mean (std)	1.	2.	3.	4.	5.	6.
1. SLB T1, leader	3.64 (.58)	-					
2. SLB T1, team	3.33 (.54)	.15	-				
3. Feedback T1, team	8.23 (1.42)	.16	.30*	-			
4. Intervention T1, team	8.44 (1.29)	.16	.24	.74**			
5. SLB T2, team	3.34 (.51)	.08	.72**	.31*	.35*	-	
6. Feedback T2, team	8.61 (.92)	.05	.25	.25	.40**	.33*	-
7. Intervention T2, team	8.52 (1.06)	-.06	.04	.32*	.45**	.02	.26

Note: SLB= Safety leadership behaviour. Sdt= standard deviation. T1 = Before training, T2= After training. * $p < .05$, * *$p < .01$

Table 3

Polynomial regression analyses and surface values for the safety leadership behaviors, safety feedback, and safety intervention

	Employee rated training outcomes at Time 2		
	Safety leadership	Safety feedback	Safety intervention
Constant	3.12*	7.00*	4.172*
Outcome, T1	-	.15	.53*
Leader-rated, T1 (b_2)	-.05	1.01*	-1.24
Team-rated, T1 (b_1)	.81*	1.03	1.07
Leader-rated squared, T1 (b_5)	.07	-.59	1.07*
Leader-rated * team-rated, T1 (b_4)	-.14	-.1.03	-1.88*
Team-rated squared, T1 (b_3)	-.07	.08	.22
R^2	.46*	.13	.31*
Surface tests			
$a_1 = (b_1 + b_2)$	.76*	2.04*	-.18
$a_2 = (b_3 + b_4 + b_5)$	-.14	-1.54*	-.56
$a_3 = (b_1 - b_2)$	-.85*	-0.02	-2.31*
$a_4 = (b_3 - b_4 + b_5)$	.13	.53	3.14*

Note. $N = 47$ leaders. R^2 = *Explained variance.* * $p < .05$

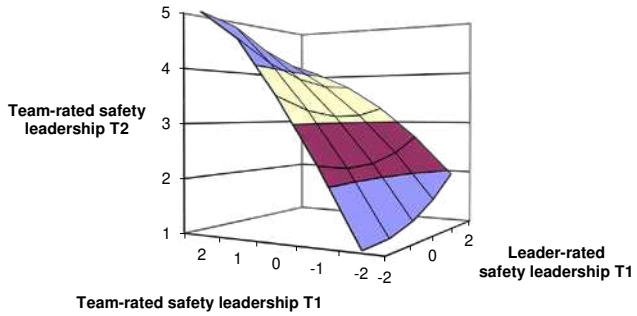

Figure 1. Leader-team perceptual distance regarding safety leadership before training (T1) and teams' ratings of safety leadership after training (T2).

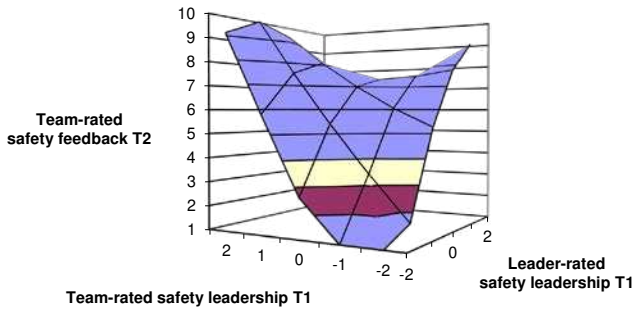

Figure 2. Leader-team perceptual distance regarding safety leadership before training (T1) and teams' ratings of safety self-efficacy to give feedback after training (T2).

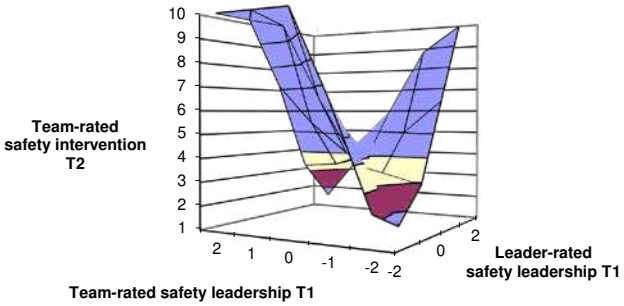

Figure 3. Leader-team perceptual distance regarding safety leadership before training (T1) and teams' ratings of safety self-efficacy to intervene after training (T2).