

The Age of Datafication: Balancing Employee-Leader Vulnerability, Trust, and Responsibility in Datafied Pay for Performance Systems

Alice Rickert*
alice.rickert@unisg.ch
University of St.Gallen

Daniela Frau*
frai@zhaw.ch
Zurich University of Applied
Sciences

Yeschi Deuss*
yeschi.deuss@unisg.ch
University of St.Gallen

*All authors contributed equally

The Age of Datafication: Balancing Employee-Leader Vulnerability, Trust, and Responsibility in Datafied Pay for Performance Systems

Abstract

This study investigates how vulnerability emerges in trust relationships between employees and leaders after implementing a datafied pay-for-performance technology. Using a dual perspective approach, we analysed 21 interviews with employees and their leaders. Our findings suggest that datafication technology introduces new forms of vulnerability for both employees and leaders. Based on Weibel et al.'s (2023) vulnerability framework, we develop a four-stage model of employee-leader vulnerability and trust. Initially, employees experience discontinuity vulnerability, prompting a re-evaluation of trust. Next, they perceive socio-emotional vulnerability, in response to which they expect their leaders to assume responsibility. However, leaders behave in a responsibility-averse manner, using HR technology to deflect responsibility back onto employees. As a result, employees experience a trust breach leading them to lower their overall trust towards their leaders. We contribute to the literature on trust, HR-technology, and responsible leadership by offering insights on managing employee-leader vulnerability in datafied workplaces.

Keywords: *Datafication; Vulnerability, Employee Trust, Responsible Leadership, Pay for Performance*

Word Count excluding tables and references: 7083

The Age of Datafication

Despite the increasing use of new datafication- technologies for core human resource (HR) processes and leadership tasks such as hiring and firing or performance evaluation, we still know little about how such datafication technologies influence trust in the workplace.

Datafication technologies, that is, technologies designed to quantify and measure social interactions for optimizing organizational processes, are significantly impacting and changing workplace dynamics at previously unprecedented speeds and in novel ways (Schafheitle *et al.*, 2020). For instance, Microsoft's Viva Insights provides predictive analytics on employee engagement and long-term performance based on automatically gathered data on organizational networks. This includes data such as employee contact frequency, interaction intensity (e.g., distribution of speaking times), and network diversity (Kasper and Chan, 2023).

Previous research investigating the effects of such datafication technologies has established that they amplify opacity, complexity, and uncertainties in decision-making processes. They challenge leaders because the underlying functions of these technologies are often undisclosed, and their self-learning capabilities result in constant changes to evaluation parameters (Leicht-Deobald *et al.*, 2019). Moreover, these technologies restrict employee autonomy and increase the possibilities for exerting organizational control while escalating the potential for unfair or unethical leadership (De Cremer *et al.*, 2022). Therewith, datafication technologies create new avenues for exploiting employees, intensifying overall employee vulnerability and threatening employee-leader trust (cf., Tsamados *et al.*, 2022; Weibel, Schafheitle and van der Werff, 2023). To mitigate this vulnerability, there is a growing consensus that leaders require new skills such as digital literacy and ethical awareness, enabling them to assume moral responsibility (De Cremer and Kasparov, 2022; Rickert, Schafheitle and Weibel, 2024; Weibel, Schafheitle and van der Werff, 2023). Consequently, equipping leaders with the skills to mitigate employee vulnerability, actively manage the trust relationship with employees, and become responsible leaders is one of the major challenges of HR management of the 21st century (Tsui, 2021; van Quaquebeke and Gerpott, 2023).

However, empirical studies investigating the perceptions of vulnerability and unfolding trust dynamics are still scarce. Even less is known about employee vulnerability and trust in the context of datafied HR technologies and responsible leadership behavior (Rickert, Schafheitle and Weibel, 2024; Weibel, Schafheitle and van der Werff, 2023). In sum, while current research offers theoretical explanations and highlights potential threats to employee-leader trust that HR must address, it still warrants empirical evidence to precisely determine the postulated underlying dynamics (e.g., vulnerability as an explanatory mechanism). Such empirical studies must delineate specific behaviors, practices, and levers for fostering employee-leader trust in datafied workplaces on the individual but also organizational level.

In this paper, we take a first step and explore the intricacies of the employee-leader trust relationship in the context of HR technology and datafication. We posit that the adverse effects of datafication technologies on employee-leader trust are exacerbated in pay-for-performance (PfP) contexts, where employees' remuneration is contingent on performance evaluation. This is because datafication within PfP significantly influences employees' financial status and, at times, their broader socioeconomic standing.

Therefore, we aim to answer the question: *How do perceptions of vulnerability in the trust relationship between employees and leaders emerge in a datafication-based HR technology context, such as Pay-for-Performance?* We adopt a dual-perspective approach to develop a model of employee vulnerability and trust. Our methodology involves semi-structured interviews with both employees and their respective team leaders. The interviews were conducted after implementing a datafication technology, namely a PfP system, in a quasi-experimental field setting.

We contribute to the literature on employee-leader trust, responsible leadership, and HR technologies by connecting morally responsible leadership with employee trust and vulnerability in a datafied PfP system context. Our data indicates that responsible leadership is necessary but challenging to achieve within vulnerable contexts. We demonstrate that employee experiences of vulnerability occur sequentially and discuss how leader awareness of employee vulnerability, perceptions of their own decision-making weaknesses (e.g., biases), and corresponding moral emotions can motivate leaders to experience responsibility aversion in the short term. This, in turn, leads to a loss of employee trust (long-term consequences) and is likely to escalate into suspicion or distrust.

In the following section, we provide a brief outlook on the theoretical background, offer a concise overview of the methodology, and quote both employees and leaders – the basis for delineating our model. Subsequently, we briefly discuss our findings and this study's implications and contributions.

Defining Trust and Vulnerability

The dynamic interplay between employees and leaders is intricately woven into workplace dynamics in contemporary organizational landscapes. Trust is central to these relationships, fostering effective collaborations and influencing a variety of workplace outcomes like heightened employee satisfaction, motivation, and performance (Dirks and Ferrin, 2002; Kannan-Narasimhan and Lawrence, 2012). Moreover, trust has been shown to increase employee commitment (Colquitt, Scott and LePine, 2007) and engagement (Ugwu, Onyishi and Rodríguez-Sánchez, 2014; Vanhala and Dietz, 2019).

Trust, conceptualized as a willingness to be vulnerable (Mayer, Davis and Schoorman, 1995), is not merely a static attribute but a dynamic force forming the bedrock of employee-leader relationships (Chughtai, Byrne and Flood, 2015; Dirks and Ferrin, 2001). In this context, the trustor is the social actor “rendering trust judgments” and thus, becomes vulnerable, while the trustee “is the party being trusted” (Jones and Shah, 2016, p. 392). Accordingly, a trustor's vulnerability arises from uncertainty regarding the prospective trustee's motives, intentions, and actions (Kramer, 2006).

Vulnerability is characterized by an individual's susceptibility to having their interests jeopardized or fulfilled depending on the actions of others (cf., Nienaber, Hofeditz and Romeike, 2015). Importantly, vulnerable individuals lack the power or capability to shield themselves from harm or guarantee the fulfillment of their interests without assistance from these other parties (Weibel, Schafheitle and van der Werff, 2023). They are, in other words, unable to independently satisfy, safeguard, or sustain their own interests at any given time (CIOMS, 2002; Nienaber, Hofeditz and Romeike, 2015). Consequently, they rely on the actions or cooperation of other parties.

Vulnerability in the Context of Datafied HR Technologies

Within the employee-leader relationship, vulnerability is reciprocal. However, with the integration of datafication technologies, this once bilateral trust dynamic becomes more intricate. Some scholars even argue that a triangular relationship emerges between employees, leaders and datafication technologies (Gal, Jensen and Stein, 2020; Schafheitle, Weibel and Rickert, 2021; van Quaquebeke and Gerpott, 2023). In such triangular relationships, employees are particularly vulnerable as they face the risk of not being heard while they must navigate contradictions between HR technology and leadership. However, also leaders are under pressure as they must identify technological biases, navigate the blurring of decision-making responsibilities between them and HR-technology, and manage a multitude of expectations from various stakeholders such as employees and customers. In essence, HR technologies become a new source feeding into both employees' and leaders' perceptions of being vulnerable. Hence, employee-leader trust is easily strained in the context of datafication-based HR technologies. *In this study, we, therefore, focus on vulnerability perceptions in the context of datafied HR technologies such as pay for performance systems.*

To study vulnerability, Weibel, Schafheitle and van der Werff (2023) have delineated a framework distinguishing two specific types of employee vulnerability nurtured by technology: First, *discontinuity vulnerability* arises when employees experience uncertainty for example due to the opaque functioning of technology. Such opaqueness can disrupt the employee-leader relationship creating discontinuity. Specifically, this vulnerability is rooted in the unpredictability of new workplace dynamics, a lack of familiarity, and the heightened complexity introduced by technology. Consequently, employees' ability to navigate such situations is compromised, exacerbating their experience of being vulnerable. Second, *socio-emotional vulnerability* arises when employees lack the feeling of being treated as human, that is, social connectedness, respect, and benevolence in the workplace. Specifically, this vulnerability arises when employees feel dehumanized, reduced to just numbers or mere “cogs in the machine” (Weibel, Schafheitle and van der Werff, 2023). Datafication technologies have a high potential to elicit such perceptions as they automate a multitude of critical HR processes that significantly impact employees' professional lives and careers, including promotions, payment, or termination procedures. Therefore, the authors conclude that employers and HR practitioners should actively manage trust to navigate the evolving vulnerabilities introduced by data-driven HR technologies (Weibel, Schafheitle and Rickert, 2024; Weibel, Schafheitle and van der Werff, 2023). By taking a step back to dissect the mechanisms at play, we aim to comprehensively explore the underlying factors contributing to this scenario.

Active Trust Management Through Responsible Technology Deployment

From our review above, we have so far established two insights: For one, datafication technologies heighten the perceptions of employee and leader vulnerability, which is an important component of trust in the workplace. For another, datafication technologies affect trust relationships to an unprecedented extent, taking control and signalling a lack of trust from employers and leaders.

Based on these two insights, previous research focusing on datafication technologies has repeatedly emphasized the importance of mitigating negative consequences and employ these technologies responsibly. Responsible technology use encompasses moral and ethical

considerations, including the protection of employees' privacy, guarding against bias, and ensuring human centricity (De Cremer and Kasparov, 2022; Jarrahi, 2018; Kellogg, Valentine and Christin, 2020). Particularly, HR governance processes and leaders play a crucial role in determining how technology is ultimately deployed. Initial research even suggests that the interpretation of responsibility, whether instrumental or genuinely benevolent, is essential for fostering trust in the future (cf., Rickert, Schafheitle and Weibel, 2024).

However, responsible decision-making alongside HR technology is challenging for leaders, too. As outlined above, technologies may create two-leader situations, in which neither employees nor leaders understand the technology's function, reliability, limitations, or scope (Lanz, Briker and Gerpott, 2024; Leicht-Deobald *et al.*, 2019; Schafheitle, Weibel and Rickert, 2021). This can create especially for leaders a diffusion of (moral) responsibility; as decision-making competencies blur, the process becomes more complex and uncertain. Gal, Jensen and Stein (2020) report a case where a highly productive consultant was almost dismissed because a datafication-based performance evaluation suggested she was underperforming. In reality, she finished her tasks ahead of time and therefore, had extended periods of screen inactivity. This complexity, opaqueness, and uncertainty can infringe on leaders' willingness to assume responsibility. Accordingly, leaders might be tempted to leave uncomfortable or challenging decisions, such as navigating moral dilemmas or determining fair bonus distributions, to technology. Such actions may stem from the desire to protect themselves from the potential vulnerability of being judged by their employees or peers (Bonnefon, Rahwan and Shariff, 2024; Dong and Bocian, 2024; van Quaquebeke and Gerpott, 2023). Therefore, whether datafication technologies are deployed responsibly may also determine the extent to which employees experience and perceive the vulnerabilities described above. To gain a deeper understanding of how perceptions of vulnerability arise during pivotal HR processes among employees and leaders, we have adopted an inductive narrative approach.

Datafied Pay-for-Performance Systems as Highly Vulnerable HR-Context

When identifying essential datafication technologies for strategic HR processes, it's crucial to understand their profound impact across the entire employee life cycle. Drawing from Pfeffer's (2018) influential work, "Dying for a Paycheck," we argue that while all aspects of this cycle matter for employee experiences, some hold greater significance for both professional and personal lives. This, in turn, affects employee vulnerability and trust. For example, performance management and payment processes determining termination, promotions, or basic salary directly impact employees' personal lives and interests, including their access to healthcare and overall well-being (Jarrahi *et al.*, 2021; Kellogg, Valentine and Christin, 2020; Pfeffer, 2018). Organizations aim to leverage highly relevant employee processes to achieve specific objectives within strategic HR management programs, utilizing practices such as PfP (Paauwe and Boselie, 2005). PfP is tailored to reward employees for meeting performance targets, aligning employee competence and behaviour with the organization's strategic goals. In essence, it incentivizes positively evaluated behaviour while penalizing deviations from predefined standards (Sapegina and Weibel, 2017).

Although such strategic HR processes impact the fundamental interests of the employees, they have only very limited power, control, or ability to influence either the process or the outcome. With the increasing deployment of datafication technologies in HR processes,

previously often diffuse processes such as payment and bonus determination, become even more complex and opaque. As a result, such processes can be seen as particularly relevant but also highly vulnerability-inducing. Consequently, they are a perfect starting point for investigating vulnerability-related trust dynamics.

Building on Weibel, Schafheitle and van der Werff (2023), we analyse how employees experience heightened vulnerability when confronted with uncertainties arising from new PfP-evaluation technology (a new PfP system). Rickert, Schafheitle and Weibel (2024) find that only employee-centric understanding by leaders leads to high-quality, employee-leader trust relationships in datafied workplaces. Accordingly, we analyse whether employees will respond to this potential heightened vulnerability particularly when the new system is perceived as flawed or unjust. Failure of leaders to protect employees from such a flawed or unjust PfP system may be interpreted as a lack of benevolence, breach of psychological contract and trust, or failure of integrity (Gal, Jensen and Stein, 2020; Leicht-Deobald *et al.*, 2019; Rousseau, 1989; Schafheitle *et al.*, 2020; Shleifer and Summers, 1987). This can lead to diminished trust, suspicion, or outright distrust.

Research Context and Methodology

Research Design, Sample, and Case Description

In exploring the impact of datafication technology on the employee-leader relationship, we conducted a qualitative study investigating the perspective of employees and leaders, focusing on employee trust and vulnerability. Our objective was to conduct a thorough inquiry into the evolving dynamics over time. Therefore, we adopted a qualitative approach to listen directly to the involved employees and leaders, enabling us to delve deeply into the experiences and interpretations they lived through. Importantly, this bilateral perspective is difficult to achieve as access to employee and respective leader data often poses a considerable challenge due to privacy regulations and organizational constraints. This is especially true in the context of new HR technology that affects a sensitive topic like monetary rewards allocation. Consequently, datasets containing such information are rare and interviews can offer unique insights into the dynamics at play.

We adopted a purposeful sampling strategy (Palinkas *et al.*, 2015) and interviewed leaders and employees from a large international insurance company undergoing a digital transformation. As part of this transformation, a new datafication technology was introduced. One notable change created by this technology was the daily measurement of employee performance, based on the number of claim cases handled, average response time, process rate, and customer satisfaction. Initially, this new technology was unconnected to any bonus awards. It was used to distribute workload equally among available team members, depending on the competencies within different work groups. Subsequently, the technology was used to measure how much workload employees in their group handled daily. A performance scorecard was used for monthly performance meetings. Leaders not only discussed the team performance, but they also acted as coaches to help the employees achieve their performance targets. However, a year after the performance scorecard implementation, employees requested the inclusion of the technological evaluations in the reward process, anticipating increased transparency and fairness in bonus allocation. Ultimately, this led to the creation of a data-based (variable) pay-

for-performance system (PfP). Using this unique situation, a mixed-method study (dissertation, Frau, 2023) was conducted namely, a field-quasi-experiment embedded by two qualitative studies to understand the impact on fairness and trust perceptions. One department consisting of four teams at different locations served as intervention group. The data-based PfP was introduced in two variants: 1) an individual-focus PfP (2 teams) or 2) a team-focused PfP (2 teams), referring to performance evaluation on the team level. These post-qualitative interviews were only used in the mixed-method study to ensure that the intervention was adequately implemented. However, unexpected quantitative findings in control variables triggered a thorough analysis of the post-qualitative interviews (this study).

The mixed-method study (Frau, 2023) examined employee-employer trust and, thus, trust on the organizational level. However, employee trust in the leader, measured as a control variable, revealed an unexpected quantitative finding. It was assumed that employees' trust in their leader would remain stable or improve in the experimental group due to the transparent communication of the new PfP scheme (e.g., using a transparent variable pay card). However, contrary to this assumption, employees in the experimental group exhibited significantly less trust in their team leader after the PfP decision ($M = 5.51$, $SD = 1.18$) compared to before its implementation ($M = 6.22$, $SD = 0.6$; asymptotic Wilcoxon test: $z = -2.39$, $p = .017$, $n = 26$). The effect size was $r = -.47$, indicating a strong effect according to Cohen (1992). In the control group, there was no significant decrease in trust. Employees' trust in their leaders remained virtually unchanged after the bonus decision ($M = 5.46$, $SD = 1.05$) compared to before ($M = 5.37$, $SD = 1.41$; asymptotic Wilcoxon test: $z = -0.498$, $p = .619$, $n = 25$). The effect size was $r = .10$, indicating a weak effect according to Cohen (1992). To understand what happened in the employee-leader trust relationship of intervention groups, we analysed the post-qualitative data resulting in this study.

Data Collection.

We analysed data from semi-structured interviews and observational notes on leadership personality and work climate in the teams (see Table 1). In total 21 semi-structured interviews were conducted in 2018 at a company meeting room with members of the department that piloted the datafied-PfP system. The interviews were digitally recorded and transcribed (average duration 45 min., 9 men, 12 women, 0 divers). This included interviews with the head of the department ($n=1$), all team leaders of this department ($n=4$), and two employees of each participating team ($n=16$; $N=55$ who fully experienced a one-year implementation of the datafied PfP system). Because we wanted to explore the impact the datafied PfP had on the employee-leaders trust relationship, we only analysed interviews with participants that had first-hand experience with both the discretionary bonus system and the new datafied PfP system. Therefore, all interview participants had been members of the organization for at least two years. All employees and leaders took part voluntarily in the interviews. To counteract biases of self-selection, leaders were asked to encourage participation from employees who were both supportive and critical of the new system, while also highlighting the voluntary nature of the interviews. All participants knew the interviewer for more than a year as she had interviewed them about their work context and daily tasks prior to the PfP system implementation. Thus, there was already an initial confidential relationship that allowed both leaders and employees to speak openly about their perceptions of the new PfP system.

Two distinct interview guidelines were devised. The *employee interview guideline* focused on how the implementation of the new datafication technology was perceived, including its impact on daily work, communication surrounding the PfP decision, and how the system influenced employee trust. The *leader guideline* addressed their experiences with the PfP system, its impact on their leadership behaviour, communication surrounding the PfP decision, and employee reactions to the PfP decision.

Table 1. Data Inventory

Data Type	Data Source	Quantity		
		Participants (n=)	Population (N=)	Matched Sample (pre/post) (n=)
Interviews	Department Head	1	1	
	Leaders	4	4	
	Employees (Individual PfP)	8	26	
	Employees (Team PfP)	8	29	
Observational Notes	Retrospective Memos of Leaders	5	5	
Employee Survey Data	Intervention Group	35	52	26
	Control Group	41	42	25

Note. Survey data was only used for triangulation. Participants (n=) refers to the number of interviewees and number of employees who filled out the survey at least once. The matched sample (n=) indicates the number of participants that completed the survey both pre- and post-experiment.

Analytical Approach.

All interviews were manually transcribed and analysed using MAXQDA 22, applying an abductive approach. We analysed employee and leader interviews to delineate specific trust dynamics. This unique dual perspective permits us a holistic understanding of the respective trust relationships and accounts for differences in employee-leader trust dynamics in various team contexts.

We started following the inductive coding procedures of the Gioia Method (Gioia, Corley and Hamilton, 2013). This involved developing first-order codes closely aligned with the respondents' perspective. To this end, one author familiar with all interviews inductively coded four interviews (leader (n=2), employee (n=2)) and identified passages that were highly informative about the trust dynamics at play. Relevant interview passages were distilled in an iterative process and served as a basis for a discussion between all authors concerning the topics and concepts present. In this discussion (moral) responsibility of leaders emerged as a potentially relevant concept. Subsequently, we aggregated the inductively derived first-order codes into more abstract and comprehensive higher-order concepts (inductive).

To illustrate the trust dynamics in place, we developed an initial grounded theory model (Gioia *et al.*, 2010; Gioia, Corley and Hamilton, 2013) based on our second-order higher concepts. We identified the two forms of vulnerability proposed by Weibel, Schafheitle and

van der Werff (2023) in the employee statements. Therefore, we proceeded to re-contextualize our inductively derived higher-order concepts with existing theory and literature (deductive). To verify the reliability of the final abductive coding guideline, two authors re-iterated the coding of one leader interview. Deviations in coding were discussed with all authors until we had a unanimous coding. Afterward, one of the authors re-iterated the coding of all interviews. As part of this process, our iterative discussions led us towards a model of vulnerability experiences and respective trust dynamics over time, that draws on the existing theoretical framework of Weibel, Schafheitle and van der Werff (2023). In line with existing research, our data contains relationships where technology infringed on employee-leader trust. Importantly, vulnerability perceptions over-time played a crucial role as depicted in our final model (Figure 1).

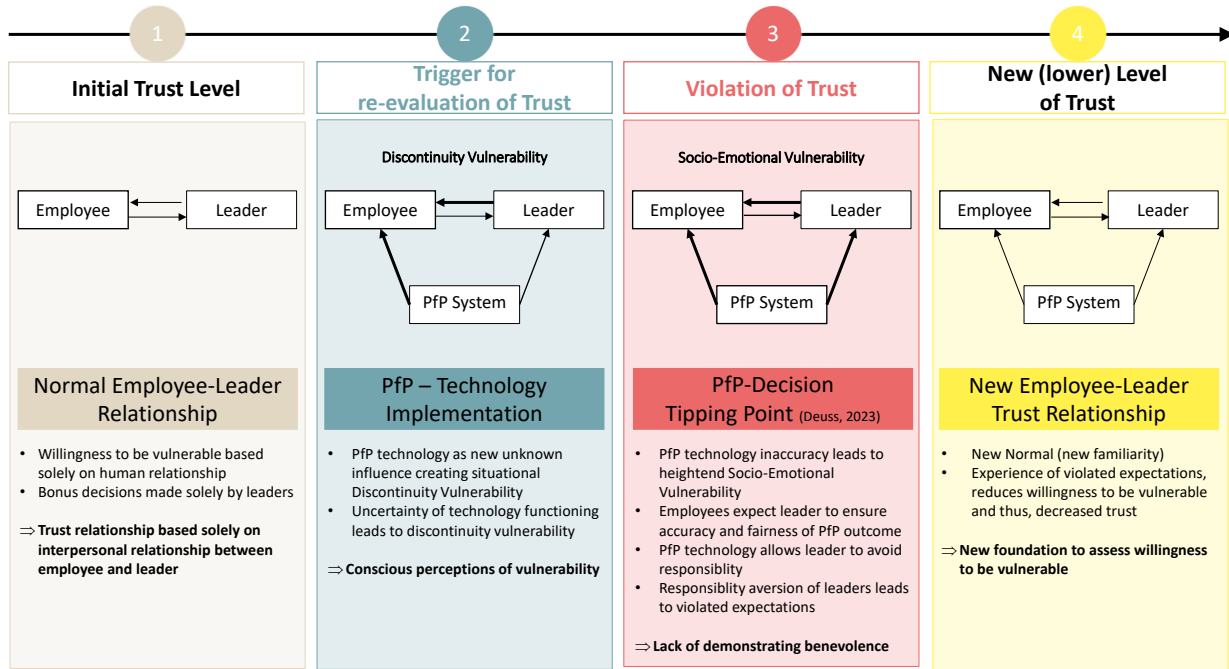
Results and Model

Interview data from employees and leaders' sheds light on the vulnerability related to trust dynamics in the context of datafied HR contexts. Our results focus on the employee perspective and the negative impact of datafication technology on employee-leader trust. We trace the dynamics that lower employee trust. Our analysis reveals that within the context of a datafied PfP-system, employees experience both discontinuity and socio-emotional vulnerability. These vulnerability experiences occur at different points in time and unfold sequentially (see Figure 1). The interaction between HR technology and leadership's responsibility assumption during the usage of the datafied PfP system emerges as a key factor shaping employees' vulnerability experiences. At the same time, we find that also leaders experience novel forms of vulnerability due to the datafied PfP system. The datafied PfP system allows leaders to avoid vulnerability by deflecting responsibility onto the employee and system, a behavior, we term responsibility aversion. We synthesize the employee-leader-technology dynamics in a four-stage model that illustrates the evolution of vulnerability experiences over time (see Figure 1). As this model is our main contribution to the trust and HRM literature, we present our data and insights corresponding to its four stages. For reader comprehensibility, employee and leader quotes will be presented separately.

Employee-Leadership Vulnerability and Trust Dynamics in Datafied HR-Contexts

Our model of employee vulnerability and trust over time illustrates four sequential stages in the employee-leader trust relationship following the introduction of datafication technology (see Figure 1). Stages One and Four represent episodes where a certain level of trust remains relatively stable over short duration, whereas stages Two and Three are episodes characterized by novel vulnerability perceptions. This is when the implementation of datafied HR-technology triggers new dynamics between employees and leaders, leading to novel perceptions and experiences of vulnerability on both sides. These vulnerability perceptions and experiences prompt a re-calibration of trust levels, leading to notable and long-lasting shifts in trust dynamics between employees and leaders.

Figure 1. Model of Employee Vulnerability and Trust over Time



Source: Author's own diagram

Stage One. The first stage depicts a “traditional” employee-leader relationship without any influence of datafied HR technologies on bonus payments. In our case, leaders determined discretionary employee bonuses on their own, mainly based on their subjective judgments of the employees’ individual performance. Therefore, employee willingness to be vulnerable was solely grounded in the interpersonal human relationship with their leader. The following quote shows the role of subjective evaluation in discretionary bonus settings.

Leader 4: “In the past, he [the employee] received XY amount and he didn't know whether it was a lot or a little and how it was made up. I always say that it was simply the supervisor's good feeling. You may have achieved the same targets as your desk neighbour, he had 4000, you had 3000, you didn't know why. And the line manager then simply explained to you somehow why you now have three and four and the targets were somehow not taken into account from his point of view.”

Stage Two. The creation of a data-based (variable) PfP system commenced a new stage. As employees requested the inclusion of the datafication-based performance evaluations in the bonus award process, they were aware that the PfP system would have a substantial impact on their payments. Although they hoped for increased transparency and fairness of the monetary reward process, they remained uncertain about how and to what extent the data-driven recommendations would be beneficial or detrimental to their interests.

Employee 1.1: “So that was still difficult to assess: Yes, what impact will it [the datafied PfP system] have in the end? What consequences does it have at all? Upwards or downwards [the final variable pay].”

Employee 2.4: *“Because of the transparency, people were perhaps a little uncertain. (...) how it [the datafied PfP system] is evaluated in relation to employees who have been with the company for a long time and employees who are newer or less experienced.”*

Thus, the awareness of the uncertainty surrounding the algorithm and the consequences of the PfP system has led employees to experience *“Discontinuity Vulnerability”*. This vulnerability appears to have been triggered by the newly introduced opaqueness in decision-making. Leadership communication emphasized that a new system was in place, with employee evaluations based on transparently communicated criteria. However, the technological balancing of these criteria, the behaviours counted towards each criterion, and the synthesis with leadership evaluations remained unknown to both employees and leaders. Due to this opaqueness of the process, neither leaders nor employees could predict the impact of technology on overall variable pay. This was illustrated by the surprise leaders reported of employees both receiving unexpectedly low and high bonuses.

Moreover, leadership behaviour in response to technological evaluations was unclear. Thus, the transparent communication regarding the new PfP system did not reduce unpredictability but instead emphasized what remained unknown concerning the PfP allocation and new workplace dynamics.

Importantly, in response to this conscious perception of vulnerability, employees closely monitor both the datafied PfP system (proactive engagement) and their leaders, thereby formulating moral expectations towards their leaders. These expectations entailed, proactive leadership engagement with technology, the rectification of performance evaluations (a novel expectation), ensuring measurement solely based on objective performance criteria, ultimately safeguarding the accuracy and fairness of PfP outcomes (see Stage 2). The following quote shows how some employees react with proactive engagement as they see their concerns concerning the PfP system being validated.

Employee 3.2: *“Yes, the problem is that some figures don’t add up, which doesn’t exactly inspire confidence. We have people keeping lists themselves, who write down everything they do, every recourse because the figures don’t add up. That can’t be right, so I don’t do it because that’s not my job. But some people keep ten lists (target 1, 2, 3...) because the numbers are never right.”*

Thus, the conscious perception of discontinuity vulnerability (here exemplified by insecurity in the performance evaluation process) acts as the initial trigger for re-evaluating the employee-leader trust relationship.

Stage Three. When employees begin to perceive the PfP system as flawed, opaque and infringing on their interests, they become aware that they need to rely on leaders to rectify the technological flaws and safeguard their interests (see Stage 2). This forced reliance on the leader to intervene on their behalf leaves them with the experience of *“Socio-Emotional Vulnerability”* (see Stage 3). Employees specifically experience three interlinked aspects. Firstly, employees feel that not only their work but also their personalities are being reduced to

numbers. This creates a foundation for feelings of being let down, betrayed or attacked by their leaders on a personal level.

Employee 2.2: *“Well, not so much because I’m disappointed about the way I was evaluated. I have no trust because it doesn’t matter. I get the feeling that no matter what I do, it wouldn’t be suitable. Because I feel that it’s not my work that is being evaluated but really me. That’s it. (...) If you’re evaluated based on your work, it’s clear, it’s mathematical. If you get an average of 4.4, you receive a set amount, and when you get an average of 2.3, you get less. I have a feeling that this is not the only thing that is being taken into account [by the PfP system and leader].”*

Secondly, if leaders fail to adequately meet the expectations employees formed in Stage 2, this can create a “tipping point” (cf., Deuss, Y. T., & Wiemann-Hügler, M, 2023) in their trust relationship. The perception of being harmed or deceived can mark a pivotal moment in the relationship, leading it to take a sharp turn toward the negative.

Employee 3.3: *That’s right, I didn’t get any support from the team leader. He didn’t notice anything, although it [the mistake] was clear. (...) [Referring to the new system] that was the beginning when I fought back. At other times [in the former system], I was not indifferent, my God, but what could I do. That is [referring to the fighting back] also where I mentioned the word bullying for the first time.*

Thirdly, employees describe a lack of leadership’s decisions being benevolent. When describing their experiences of socio-emotional vulnerability, employees often refer to their leaders as leaving PfP decisions solely to new technology, using the technological recommendation as an excuse and failing to verify the data to ensure a certain quality level of the bonus award process.

Employee 3.3: *“No, because my manager doesn’t clarify it. He shows us the figures as they are received without considering they might be wrong. If a person (handles) 200 cases too few, then as a supervisor, I would look more closely into it and not just say, ‘you’ve underperformed.’ That’s my opinion.”*

Interestingly, employees seem to have a quite accurate perception of their leaders’ behaviours and motives. Our data from the leadership interviews indicate, that some leaders indeed show, what we term “*Responsibility Aversion*” (see Figure 1, Stage 3). These leaders are often aware of the biases and shortcomings human decision-making processes contain. They feel that such monetary rewards processes sometimes are overly prone to reflect recent interactions with an employee instead of a holistic year picture. Moreover, they are aware of the subjectivity of performance evaluations noticing the differences in the evaluation of employees that emerge due to personal tendencies of being more lenient or strict in evaluations as a leader.

Leader 4: *“Because in the past it was like this: Yes, you do the categorization and then you have amounts, then you start to (...) differentiate, to round, and then you have to coordinate it among each other (between the leaders). And then he thinks: ‘Yes, but no, if yours has 3.4, then mine has 3.7’. [This time] it was much easier because it was thematic and there was no discussion with the other leaders: ‘Yes, I want to give him 100 more now’. Yes, it's just the system that tells you the amount.”*

Leader 2: *“(...) with the new process I have the whole year in mind. [With the bonus process] it's certainly often driven by the moment: if he [the employee] has perhaps just annoyed me, I simply have the moment in mind and am then perhaps no longer completely objective and have forgotten that he did a great job in January to March, because now perhaps for some reason the behaviour is a bit more tense. And with the new system: the figures speak for themselves. And I can also see from the figures when I start to think about it retrospectively and we have to assess the behaviour: Now, okay, we've always discussed the figures and I can see an increase in performance - I've reflected on the year differently with the new process.”*

This awareness made it previously difficult for them to justify their bonus decision towards employees, making them vulnerable towards criticism and feelings of guilt and stress. They experience the new system as a possibility to protect their own vulnerability toward employees by deflecting responsibility back on the employee and system. From their perspective, the datafied PfP system contributes to responsibility diffusion (Schafheitle et al., 2021), allowing them to bear less blame and guilt in monetary rewards decisions.

Leader 1: *“Yes. I have a certain protective mechanism for myself because, in the past, I sometimes had a guilty conscience if I gave someone less than the previous year. And it probably felt more personal when I told the employee face-to-face. Now, I can hide behind these printouts and hand back full responsibility to the employee.”*

Interviewer: *“Or blame the system?”*

Leader 1: *“Yes, a combination of the two. It was also less of a burden for me because I was only interested in it to a limited extent. So, I saw myself as an agent of the system, but I didn't – or only felt partially responsible – for the final outcome. In the past, I really felt responsible when I had to allocate and then discuss variable pay.”*

Thus, at times leaders even themselves recognizes their responsibility toward the employee but refuse to embrace it (see Stage 3). They consciously hide behind the system to avoid responsibility. However, it seems that while leaders are aware of their reluctance or even aversion to assume responsibility, they seem to be unaware of the

socio-emotional vulnerability experiences and corresponding long-term consequences for the trust relationships with their employees.

Stage four. The employee-leader dynamic has reached a new (lower) level of trust. Due to violated trust expectations, a perceived lack of leader empathy and benevolence in light of increased vulnerability employee's willingness to be vulnerable is reduced. They feel "used" and dehumanized. Consequently, their trust in the leader decreases and reaches a new lower equilibrium. Based on the described experiences with the datafied PfP system a "new normal" and familiarity has emerged. These experiences build a new foundation for employees to determine their willingness to be vulnerable and hence, trust in the employee-leader relationship.

Employee 3.1: *"In the team, so I trust them anyway. (...) And with the team leader is actually still difficult (laughs)."*

Employee 1.1.: *"(...) it is not weighted correctly, it is not applied correctly. It is illogical in the way it is applied, not fair. It's often a matter of luck. (...)"*

Interviewer: *"Has it generally strengthened your trust in the team leader?"*

Employee 1.1: *"No."*

Interviewer: *"Can you tell me again the main points, the three main points, why not?"*

Employee 1.1.: *"Because he didn't ... Because he (...) lost the overview, I think. (...) Because ... after every one-to-one, for example, the process satisfaction (customer satisfaction) was just good, good, good. Yes, I had a bit of bad luck because of a claim rejection, (...) And at the end it says: [Goal] not achieved!"*

Discussion

We analysed the perceptions of employee and leader vulnerability in the context of datafied PfP systems. Drawing on interviews of employees and leaders we examined when and what form of vulnerability emerges and how it contributes to the employee-leader trust dynamics that unfold. We find that employees experience discontinuity and socio-emotional vulnerability in a sequential fashion. These diminish employee trust towards leaders. Moreover, we find responsibility aversion behaviour of leaders, who use the new datafied PfP system to deflect responsibility onto employees or the system to avoid own vulnerability towards their employees or peers. We want to highlight two contributions of our study and discuss them in more detail below. First, our study contributes theoretically to trust theory, as we are the first to illustrate that the forms of employee vulnerability proposed by Weibel, Schafheitle and van der Werff (2023) occur in a sequential fashion. Moreover, we also contribute to the literature on algorithmic management and datafication technology as we are, to our knowledge among the first, to provide empirical evidence for responsibility aversion of leaders using datafied HR technology. In the following section, we will discuss our insights and their significance for the respective fields in more detail.

First, our insights contribute to trust and datafication technology research. Our data indicates that employee vulnerabilities in a datafied HR technology context like PfP systems occur sequentially. The insecurities surrounding the new PfP system led to discontinuity vulnerability which in turn triggers a conscious re-evaluation of employee trust. Employees expect leaders to show their benevolence by actively engage with the system to ensure fairness and transparency. This dependency of leader benevolence results in socio-emotional vulnerability. When employees realize that their expectations are not fulfilled or violated, the relationship deteriorates – a lower form of trust relationship, or even a suspicion or distrust-based relationship is established.

With regards to the datafication and algorithmic management literature, these insights align with previous research that finds only leaders assuming the role of moral advocates for employees and aiming to protect them, can maintain employee trust in the context of datafication technologies. In contrast, leaders approaching this topic instrumentally can never be truly perceived as trustworthy by their employees (Rickert, Schafheitle and Weibel, 2024). Moreover, it also mirrors the predictions of Kellogg, Valentine and Christin (2020), who argue in their review of existing literature that datafication tools impact employee trust by redefining recognition and reward systems, which in combination with the heightened technological influence and control creates new adverse experiences of vulnerability for employees in the workplace. Similarly, several authors have argued that the growing use of datafication technology requires a stronger focus on human-centric approaches, taking into account employees' experience, emotions, and the potential individual hardships they may face, in order to mitigate the negative consequences associated with datafied HR technologies (Floridi *et al.*, 2018; Heukamp, 2020; Mittelstadt, 2019; Schafheitle *et al.*, 2020; Schafheitle, Weibel and Rickert, 2022).

For trust theory, our insights have three implications: Firstly, our results highlight, that trust and distrust are dynamically evolving concepts that change over time. Such dynamic processes have also been theoretically conceptualized by recent work of Gaito (2023), but also older works like Lewicki, Tomlinson and Gillespie (2006). However, the scholarly literature on trust measurement does not yet sufficiently account for these dynamic processes and fluctuations. Measuring vulnerability perceptions, and triggers for discontinuity or socio-emotional vulnerability might be an avenue, that allows to account for and predict such fluctuations and hence, trust development over time. Secondly, our analysis provides empirical evidence that the willingness to be vulnerable is indeed a core component of trust, crucial for explaining and understanding its evolution over time. It is important to recognize that it is not necessarily actual vulnerability that employees experience (even if present in our data), but rather that perceptions of vulnerability might be sufficient to elicit changes in the trust dynamic. Therefore, managing trust through vulnerability will always entail managing human emotions, fears, and needs, making trust an inherently human-centric concept. Thirdly, one of the most relevant insights from our research on employee-leader trust is that this trust also necessitates vulnerability from leaders. Leadership concepts like authentic or humble leadership have long emphasized the importance of leaders being and showing vulnerability (Bunker, 1997). Interventions to foster such vulnerability include well-known formats like “fuck-up (leadership) nights” or reflective learning journeys. Essentially these might be interventions that enable us to teach leaders how to embrace vulnerability by actively recognizing and dealing with their

own unpleasant emotions, instead of keeping up appearances. Such abilities will become crucial to safeguard employee-leader trust in an area of technology. Second, our insights also contribute to responsible leadership theory (Waldman, Siegel and Stahl, 2020) since our data indicates that employees form conscious expectations concerning the (moral) responsibility of leaders and monitor their leaders' actions accordingly. When leaders refuse to assume the responsibility employees expect, socio-emotional vulnerability is heightened, leading to a decline in their willingness to exercise trust. Interestingly, leaders are also aware of their moral responsibilities and role. Our data confirms the existence of responsibility aversion, it also highlights that contrary to the common notion in academic literatures (i.e., that leaders who are aware of their responsibility and care about their employees will accept responsibility; cf., Bonnefon, Rahwan and Shariff, 2024; Frangieh and Yaacoub, 2017; Gond *et al.*, 2017) some consciously chose responsibility aversion to reduce their own vulnerability. Therefore, our findings provide evidence for the recent predictions of several authors, that datafication technologies might be used as excuses for leadership behaviour (van Quaquebeke and Gerpott, 2023) and therefore, require even more responsibility from human leadership (De Cremer and Kasparov, 2022). Consequently, HRs job is not done by simply recruiting caring leaders with a high responsibility disposition. Instead, they must actively train and equip leaders to assume responsibility in the face of datafication technologies. This will safeguard the employee-leader trust relationship from spiralling into distrust.

Limitations and Future Research

We acknowledge several limitations in this study. Firstly, the interviews were initially planned to verify the accuracy of an intervention. Therefore, they were not tailored to address topics such as vulnerability and responsibility. While this presents a limitation, it underscores the pronounced themes within the data and guarantees that the interviewers' assumptions and own theories have not biased the data collection. Moreover, this design aligns with the exploratory nature of this study.

Secondly, we collected a sample of only two employees per leader, drawing participants from four different teams. Ideally, we would have included all employees but resource constraints and passed time between data collection and analysis (roughly one and a half years) prevented a sample extension. Despite the limited number of interviews, we, fortunately, followed a purposeful sampling strategy, interviewing both supporters and critics of the new datafied PfP system from different teams. The purposeful cross-team sampling and one researcher's extensive on-site presence, provide us with confidence that the findings reflect general trends rather than an extraordinary phenomenon. Moreover, the context of an organizational experiment encouraged the honesty of participants concerning the new PfP system's benefits and drawbacks. Importantly, they also shared information that could have potentially negative effects on their careers. Consequently, we believe that we have captured the full range of emotions and honest experiences prevalent.

Nevertheless, future research could benefit from comparisons across multiple cases to verify the mechanisms at play in different organizational contexts. Further testing of the dynamics would enhance the robustness of the findings. Additionally, we were unable to delve deeply into the issue of leaders' responsibility aversion, particularly regarding the contextual and organizational factors influencing its emergence. Future studies could address this

limitation and examine the concept of responsibility aversion over time. They could explore specific strategies and interventions to promote responsible leadership practices in datafied organizational contexts. Longitudinal studies tracking the long-term effects of datafied PfP systems on employee trust and leader behaviour would also provide valuable insights into trust relationships in evolving organizational landscapes. Such insights are crucial to manage trust in the era of datafication.

REFERENCES

- Bonnefon, J.-F., Rahwan, I., & Shariff, A. (2024). The Moral Psychology of Artificial Intelligence. *Annual Review of Psychology*, 75(1), 653–675.
<https://doi.org/10.1146/annurev-psych-030123-113559>
- Bunker, K. A. (1997). The power of vulnerability in contemporary leadership. *Consulting Psychology Journal: Practice and Research*, 49(2), 122–136.
<https://doi.org/10.1037/1061-4087.49.2.122>
- Chughtai, A., Byrne, M., & Flood, B. (2015). Linking Ethical Leadership to Employee Well-Being: The Role of Trust in Supervisor. *Journal of Business Ethics*. Advance online publication. <https://doi.org/10.1007/s10551-014-2126-7>
- CIOMS. (2002). *International Ethical Guidelines for Biomedical Research Involving Human Council for International Organizations of Medical Sciences*.
- Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *The Journal of Applied Psychology*, 92(4), 909–927.
<https://doi.org/10.1037/0021-9010.92.4.909>
- De Cremer, D., & Kasparov, G. (2022). The ethical AI—paradox: why better technology needs more and not less human responsibility. *AI and Ethics*, 2(1), 1–4.
<https://doi.org/10.1007/s43681-021-00075-y>
- De Cremer, D., Narayanan, D., Deppeler, A., Nagpal, M., & McGuire, J. (2022). The road to a human-centred digital society: Opportunities, challenges and responsibilities for humans in the age of machines. *AI and Ethics*, 2(4), 579–583.
<https://doi.org/10.1007/s43681-021-00116-6>
- Deuss, Y. T., & Wiemann-Hügler, M. (2023.). Sensemaking and Tipping Points: Investigating the Process of Distrust Formation. In FINT (Chair), *12th FINT Workshop on Trust Within and Between Organizations*, Helsinki, Finland.
- Dirks, K. T., & Ferrin, D. L. (2001). The Role of Trust in Organizational Settings. *Organization Science*, 12(4), 450–467. <https://doi.org/10.1287/orsc.12.4.450.10640>
- Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: Meta-analytic findings and implications for research and practice. *The Journal of Applied Psychology*, 87(4), 611–628. <https://doi.org/10.1037/0021-9010.87.4.611>
- Dong, M., & Bocian, K. (2024). Responsibility gaps and self-interest bias: People attribute moral responsibility to AI for their own but not others' transgressions. *Journal of Experimental Social Psychology*, 111, 104584.
<https://doi.org/10.1016/j.jesp.2023.104584>

- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). Ai4people-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Frangieh, G. C., & Yaacoub, K. H. (2017). A systematic literature review of responsible leadership. *Journal of Global Responsibility*, 8(2), 281–299. <https://doi.org/10.1108/jgr-01-2017-0004>
- Frau Daniela. (2023). *The Impact of Variable Pay for Mixed Performance on Trust in the Employer Through the Lens of Fairness Heuristics. A Longitudinal Mixed-Method, Field Quasi-Experiment.: Doctoral Dissertation (no. 5328)*. University of St. Gallen
- Gaito, T. (2023). Rising from the Darkness: Transforming Intractable Conflicts through Stakeholder Engagement. *Academy of Management Proceedings*, 2023(1), Article 13008. <https://doi.org/10.5465/AMPROC.2023.104bp>
- Gal, U., Jensen, T. B., & Stein, M.-K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, 30(2), 100301. <https://doi.org/10.1016/j.infoandorg.2020.100301>
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, 16(1), 15–31. <https://doi.org/10.1177/1094428112452151>
- Gioia, D. A., Price, K. N., Hamilton, A. L., & Thomas, J. B. (2010). Forging an Identity: An Insider-outsider Study of Processes Involved in the Formation of Organizational Identity. *Administrative Science Quarterly*, 55(1), 1–46. <https://doi.org/10.2189/asqu.2010.55.1.1> (Zeitschriftenaufsatz).
- Gond, J.-P., El Akremi, A., Swaen, V., & Babu, N. (2017). The psychological microfoundations of corporate social responsibility: A person-centric systematic review. *Journal of Organizational Behavior*, 38(2), 225–246. <https://doi.org/10.1002/job.2170>
- Heukamp, F. (2020). Ai and the Leadership Development of the Future. In J. Canals (Ed.), *IESE Business Collection. The Future of Management in an AI World: Redefining Purpose and Strategy in the Fourth Industrial Revolution* (pp. 137–148). Springer International Publishing AG. https://doi.org/10.1007/978-3-030-20680-2_7
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>

- Jarrahi, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211020332>
- Jones, S. L., & Shah, P. P. (2016). Diagnosing the locus of trust: A temporal perspective for trustor, trustee, and dyadic influences on perceived trustworthiness. *The Journal of Applied Psychology*, 101(3), 392–414. <https://doi.org/10.1037/apl0000041>
- Kannan-Narasimhan, R., & Lawrence, B. S. (2012). Behavioral Integrity: How Leader Referents and Trust Matter to Workplace Outcomes. *Journal of Business Ethics*. Advance online publication. <https://doi.org/10.1007/s10551-011-1199-9>
- Kasper, J., & Chan, M. (2023). *Introduction to Organizational Network Analysis with vivainsights in R*. <https://microsoft.github.io/vivainsights/articles/network-analysis.html>
- Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*, 14(1), 366–410. <https://doi.org/10.5465/annals.2018.0174>
- Kramer, R. M. (Ed.). (2006). *Oxford management readers. Organizational trust: A reader* (1. publ). Oxford University Press. http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014828087&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA
- Lanz, L., Briker, R., & Gerpott, F. H. (2024). Employees Adhere More to Unethical Instructions from Human Than AI Supervisors: Complementing Experimental Evidence with Machine Learning. *Journal of Business Ethics : JBE*, 189(3), Article 12392, 625–646. <https://doi.org/10.1007/s10551-023-05393-1>
- Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S. D., Wildhaber, I., & Kasper, G. (2019). The Challenges of Algorithm-Based HR Decision-Making for Personal Integrity. *Journal of Business Ethics : JBE*, 160(2), 377–392. <https://doi.org/10.1007/s10551-019-04204-w>
- Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of Interpersonal Trust Development: Theoretical Approaches, Empirical Evidence, and Future Directions. *Journal of Management*, 32(6), 991–1022. <https://doi.org/10.1177/0149206306294405>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, 20(3), 709. <https://doi.org/10.2307/258792>
- Mittelstadt, B. (2019). AI Ethics – Too Principled to Fail? *SSRN Electronic Journal*. Advance online publication. <https://doi.org/10.2139/ssrn.3391293>

- Nienaber, A.-M., Hofeditz, M., & Romeike, P. D. (2015). Vulnerability and trust in leader-follower relationships. *Personnel Review*, 44(4), 567–591. <https://doi.org/10.1108/PR-09-2013-0162>
- Paauwe, J., & Boselie, P. (2005). HRM and performance: what next? *Human Resource Management Journal*, 15(4), 68–83. <https://doi.org/10.1111/j.1748-8583.2005.tb00296.x>
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health*, 42(5), 533–544. <https://doi.org/10.1007/s10488-013-0528-y>
- Pfeffer, J. (2018). *Dying for a paycheck: How modern management harms employee health and company performance - and what we can do about it* (First edition). Harper-Collins Publishers.
- Rickert, A., Schafheitle, S. D., & Weibel, A. (2024). Unleashing the power of employee trust. The uncharted influence of responsible leadership in technology permeated workplaces. In T. Bui (Ed.), *Proceedings of the 57th Annual Hawaii International Conference on System Sciences*. <https://www.alexandria.unisg.ch/handle/20.500.14171/118138>
- Rousseau, D. M. (1989). Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal*, 2(2), 121–139. <https://doi.org/10.1007/BF01384942>
- Sapiegina, A., & Weibel, A. (2017). The Good, the Not So Bad, and the Ugly of Competitive Human Resource Practices: A Multidisciplinary Conceptual Framework. *Group & Organization Management*, 42(5), 707–747. <https://doi.org/10.1177/1059601117730238>
- Schafheitle, S. D., Weibel, A., Ebert, I. L., Kasper, G., Schank, C., & Leicht-Deobald, U. (2020). No stone left unturned? Towards a framework for the impact of datafication technologies on organizational control. *Academy of Management Discoveries*. Advance online publication. <https://doi.org/10.5465/amd.2019.0002>
- Schafheitle, S. D., Weibel, A., & Rickert, A. (2021). The Bermuda Triangle of Leadership in the AI Era? Emerging Trust Implications From “Two-Leader-Situations” in the Eyes of Employees. In T. Bui (Ed.), *Proceedings of the 54th Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2021.665>
- Schafheitle, S. D., Weibel, A., & Rickert, A. (2022). Exploring the Chemistry of Datafication Control – Pathways for a Trust-Enabling Use of Smart Workplace Technology. In T. Bui (Ed.), *Proceedings of the 55th Annual Hawaii International Conference on System Sciences*. <https://doi.org/10.24251/HICSS.2022.713>

- Shleifer, A., & Summers, L. (1987). *Breach of Trust in Hostile Takeovers*. National Bureau of Economic Research. <https://doi.org/10.3386/w2342>
- Tsamados, A., Aggarwal, N., Cows, J., Morley, J., Roberts, H., Taddeo, M., & Floridi, L. (2022). The ethics of algorithms: key problems and solutions. *AI & SOCIETY*, 37(1), 215–230. <https://doi.org/10.1007/s00146-021-01154-8>
- Tsui, A. S. (2021). Responsible Research and Responsible Leadership Studies. *Academy of Management Discoveries*, 7(2), 166–170. <https://doi.org/10.5465/amd.2019.0244>
- Ugwu, F. O., Onyishi, I. E., & Rodríguez-Sánchez, A. M. (2014). Linking organizational trust with employee engagement: the role of psychological empowerment. *Personnel Review*, 43(3), 377–400. <https://doi.org/10.1108/PR-11-2012-0198>
- van Quaquebeke, N., & Gerpott, F. H. (2023). The Now, New, and Next of Digital Leadership: How Artificial Intelligence (AI) Will Take Over and Change Leadership as We Know It. *Journal of Leadership & Organizational Studies*, 30(3), 265–275. <https://doi.org/10.1177/15480518231181731>
- Vanhala, M., & Dietz, G. (2019). How Trust in One's Employer Moderates the Relationship Between HRM and Engagement Related Performance. *International Studies of Management & Organization*, 49(1), 23–42. <https://doi.org/10.1080/00208825.2019.1565092>
- Waldman, D. A., Siegel, D. S., & Stahl, G. K. (2020). Defining the Socially Responsible Leader: Revisiting Issues in Responsible Leadership. *Journal of Leadership & Organizational Studies*, 27(1), 5–20. <https://doi.org/10.1177/1548051819872201>
- Weibel, A., Schafheitle, S. D., & Rickert, A. (2024). Vertauen Aktiv Management. *Personalmagazin*(02). <https://www.alexandria.unisg.ch/handle/20.500.14171/119482>
- Weibel, A., Schafheitle, S. D., & van der Werff, L. (2023). Smart Tech is all Around us – Bridging Employee Vulnerability with Organizational Active Trust-Building. *Journal of Management Studies*, Article joms.12940. Advance online publication. <https://doi.org/10.1111/joms.12940>