

Beyond Algorithms: Employee Perceptions of Fairness, Trust and Engagement in AI-Driven Feedback Systems

A Qualitative Case Study on the Swedish Newspaper Organization Bonnier News



Liv Sterner and Gustav Nadrowski Olsson
Bachelor Thesis, December 2024

Abstract

This thesis investigates the impact of AI-driven feedback surveys on employee perceptions of fairness, trust, and engagement within Swedish newspaper organizations, with a case study on Bonnier News. Utilizing a qualitative approach, semi-structured interviews were conducted with employees to explore their experiences with the AI-based feedback system, Winningtemp. The findings reveal a complex interplay between trust in AI and human oversight, where trust in managers significantly influenced employees' acceptance of the system. Transparency in decision-making and the balance between frequent feedback opportunities and survey fatigue emerged as critical factors shaping engagement and fairness perceptions. While AI facilitated real-time insights and efficiency, the human role in contextualizing and acting on feedback remained indispensable. This study highlights the need for careful integration of AI in workplace feedback systems to ensure alignment with organizational justice principles, ultimately fostering trust and engagement.

Keywords: Artificial Intelligence, Employee Feedback Survey, Fairness, Trust, Engagement, Employee Perceptions, Newspaper Organizations

Authors

Liv Sterner (25646),
Gustav Nadrowski Olsson (23724)

Supervisor

Ebba Henrekson, Affiliated Researcher, SSE Institute for Research and Center for Educational Leadership and Excellence

Examinator

Laurence Romani, Professor, Department of Management and Organization

Acknowledgements

We would like to express our gratitude to our supervisor, Ebba Henrekson, for her guidance and support throughout this project. We are also deeply grateful to our supervision group for their constructive feedback along the way. A special thank you goes to all the interviewees who generously shared their time and insights, making this research possible. Your contributions have been instrumental in shaping our understanding and enriching our study. Thank you to each and every one of you.

Table of Contents

1 Introduction.....	1
1.1 Background.....	1
1.2 Problem Statement and Research Question.....	2
1.3 Scope and Limitations.....	2
1.4 Bonnier News.....	3
1.5 Winningtemp.....	3
2 Literature Review.....	4
2.1 Introduction to Key Concepts.....	4
2.2 Review on Previous Research.....	4
2.2.1. Traditional feedback and bottom-up communication.....	4
2.2.2 Traditional Surveys.....	5
2.2.3 AI in feedback and bottom-up communication.....	6
2.2.4 Taking on a Process-Perspective.....	7
2.3 Research Gaps.....	7
2.4 Limitations in Literature Review.....	8
3 Theoretical Framework.....	8
3.1 Fairness Heuristic Theory: How perceptions of fairness/trust can arise.....	8
3.2 Human Relations Theory: The role of interpersonal justice in the process.....	10
3.2.1 Relevance of HRT.....	10
4 Research Methodology.....	12
4.1 Research Philosophy and Approach.....	12
4.1.1 Research Paradigm.....	12
4.1.3 Research Approach.....	12
4.2 Research Design.....	12
4.2.1 Qualitative Research Design.....	12

4.2.2 Semi-Structured Interviews.....	13
4.2.3 Design of Interview Questions.....	13
4.3 Data Collection.....	14
4.3.1 Sampling Strategy.....	14
4.3.2 Interview Process.....	15
4.4 Thematic Analysis.....	16
4.5 Research Quality.....	16
4.5.1 Rigor.....	16
4.5.2 Reflexivity.....	16
4.5.2 Ethical Considerations.....	17
4.6 Usage of Artificial Intelligence.....	17
5 Empirics.....	17
5.1 Trust.....	18
5.1.1 Awareness of AI.....	18
5.1.2 Trusting the People Behind the System.....	19
5.1.3 Transparency and Its Impact on Trust.....	20
5.2 Engagement.....	21
5.2.1 Balance Frequency with Survey Fatigue.....	21
5.2.2 Anonymity Concerns.....	22
5.2.3 Communication and Follow-Up from Managers.....	23
5.3 Perception of Fairness.....	25
5.3.1 Fairness in Feedback Interpretation.....	25
5.3.2 Perceived Fairness in Representation.....	26
5.3.3 Transparency as a Component of Fairness.....	26
5.4 Conclusion.....	27
6 Analysis.....	27
6.1 The human role central in overall fairness.....	28

6.2 Transparency.....	30
6.3 Frequency & Representation.....	30
6.4 Anonymity and invasion of privacy.....	31
7 Discussion.....	31
7.1 Revisiting the Research Question.....	31
7.2 Comparison with Previous Research.....	32
7.3 Practical Implications.....	33
7.4 Directions for Future Research.....	33
8 Conclusion.....	33
9 References.....	34
10 Appendix.....	38
10.1 Appendix 1: Interview guide.....	38
10.2 Appendix 2: Email to Respondents.....	40
10.3 Appendix 3: Interview Sample.....	41

1 Introduction

1.1 Background

Artificial Intelligence (AI) has become a defining force of the 21st century, transforming industries, reshaping work processes, and driving innovation in ways previously unimaginable (Gong, 2023). Among its most profound impacts is its integration into the internal workings of organizations, revolutionizing how decisions are made and employees are managed. In recent years, AI has become a cornerstone of organizational transformation, with applications ranging from enhancing operational efficiency to reshaping employee management practices. Across various sectors, AI is streamlining processes, offering real-time insights, and improving decision-making. (McKinsey & Company, 2024).

This wave of AI adoption has not bypassed the Swedish newspaper industry sector, which is itself undergoing rapid transformation driven by digitalization and technological innovation (Statista, 2024). Newspaper organizations are now leveraging AI in multiple ways. Externally, for instance, AI enhances audience experiences by generating AI-written article summaries (Dagens Media, 2023). Internally, it is finding its way into human resource management (HRM) practices, fundamentally altering how newspaper companies engage with and support their employees (Bonnier Group AB, 2023).

One prominent example is Bonnier News, Sweden's largest media group and a key player in the Nordic media landscape (Nordicom, 2022). Bonnier News owns several influential publications, including Dagens Nyheter and Expressen, and reaches millions of readers daily through digital and traditional channels (Bonnier News, n.d.). Recently, the company implemented AI-driven employee feedback surveys to replace traditional annual surveys (Bonnier Group AB, 2023).

AI-based employee feedback surveys hold promise for addressing the limitations of traditional surveys, such as infrequent data collection and slow response times (Garrad & Hyland, 2020), but they also introduce new challenges. For instance, reliance on algorithms may lead employees to perceive AI-based decision-making as less fair or transparent, raising critical questions about

trust, engagement, and the overall impact on workplace culture (Miller, 2023). These concerns are particularly pertinent in creative and decentralized environments, such as media organizations, where interpersonal communication and trust play a pivotal role (Wang, 2024).

Although AI offers significant potential for enhancing HRM practices, it also requires careful implementation. While these tools streamline operations, their success hinges on employee trust, engagement, and perceptions of fairness. Without these critical components, the very systems meant to empower employees and improve decision-making risk creating skepticism or resistance. Understanding how employees perceive AI-based feedback systems in newspaper organizations, like Bonnier News, is therefore vital. As these tools continue to shape organizational dynamics, examining their impact on employee trust, engagement, and perceptions of fairness provides valuable insights for companies navigating this new transition.

1.2 Problem Statement and Research Question

As newspaper organizations are adopting AI-driven tools to gather employee feedback, understanding how these systems influence employee perceptions becomes essential. AI-based employee feedback surveys offer opportunities to enhance efficiency and provide data-driven insights (McKinsey & Company, 2024), but may also raise concerns about employee trust, engagement and perceptions of fairness (Bankins et al., 2022). This thesis aims to examine how AI-based employee feedback surveys shape employee perceptions of trust, engagement and fairness within large Swedish newspaper organizations. By exploring these perceptions, we seek to provide valuable insights into the broader implications of adopting AI-driven employee feedback systems in creative industries. This study centers around our research question:

How do AI-based employee surveys influence employee perceptions of trust, engagement and fairness within Swedish Newspaper Organizations?

1.3 Scope and Limitations

This study focuses on the media sector in Sweden, with a specific emphasis on Bonnier News. The scope of this research is limited to AI-based employee feedback surveys within Bonnier News. While the findings may offer insights into broader trends in AI adoption across other

industries, the study is limited by its focus on a single organization within the media sector. Additionally, the study will focus primarily on employee perceptions, which may vary depending on individual roles, levels of interaction with the AI system, and organizational culture.

1.4 Bonnier News

Bonnier News has been chosen as a case study for exploring the impact of AI-based employee feedback surveys on employee perceptions. As one of the largest newspaper organizations in the Nordic Region (Nordicom, 2022), Bonnier News plays a central role in the Swedish media landscape, making it an ideal setting for this research. It provides a diverse range of content, including daily news, investigative journalism, opinion pieces and lifestyle features, and business reporting, through its major publications such as *Dagens Nyheter*, *Expressen*, and *Dagens Industri*. With over 2 000 employees across various roles, Bonnier News represents a multidisciplinary organization with diverse operational environments. While the company includes functions like marketing IT, and administrative departments, this study focuses specifically on the editorial environment, where journalists and other creative professionals collaborate to produce content (Bonnier News, n.d.). The editorial environment is particularly dynamic and creative, making communication, collaboration, and trust essential components of its functioning (Bonner AB, n.d.).

1.5 Winningtemp

In 2023, Bonnier News adopted an AI-driven employee feedback system through the Winningtemp platform, replacing traditional annual surveys (Bonnier Group AB, 2023). Winningtemp is an AI-driven employee platform designed to enhance employee engagement and organizational decision-making. By leveraging machine learning algorithms, it provides real-time insights into employee sentiment, job satisfaction, stress levels, and turnover risk. Unlike traditional survey methods, Winningtemp's system is supposed to tailor questions based on previous responses, ensuring relevance and encouraging consistent participation (Winningtemp, 2023). The system utilizes short, frequent surveys, which, according to Winningtemp, takes approximately 15 seconds to complete. (Winningtemp, n.d.)

2 Literature Review

2.1 Introduction to Key Concepts

Fairness: The definition of fairness used in organizational justice theories and the most recognized definition of justice is "To render to each their due" (Feinberg, 1974). Examples include gratitude, respect, blame, and appreciation. In organizational justice literature, fairness has commonly been described as flowing downstream from justice, and so the two terms will be used interchangeably (Colquitt & Rodell, 2015).

Human dignity: Human dignity is a noncomparative principle of justice from which interactional and procedural justice theories hinge (Bies, 2015). It refers to the worth held by each individual whose innate rights determine their due. Examples include respect and recognition in workplace interactions (Bies, 2001). Interactional justice ensures that employees feel valued, which AI systems must uphold to foster fairness and trust (Angerschmid et al., 2022).

Feedback-Driven Decision-making: Our focus is on decision-making rooted in employee feedback, not traditional managerial decision-making processes. Feedback-driven decision-making involves analyzing and action on insights collected from employees (Bankins et al., 2022).

2.2 Review on Previous Research

2.2.1. Traditional feedback and bottom-up communication

Traditional feedback mechanisms provides valuable insights into the dynamics of bottom-up communication and its role in fostering perceptions of fairness. The concept of *interactional justice*, emphasizing that respectful treatment and truthful explanations during interpersonal exchanges, significantly enhance perceptions of fairness (Bies and Moag, 1986). This framework forms the basis for understanding how fairness is perceived in feedback processes. Hallowell (1999) explored the significance of interpersonal dynamics in his work *The Human Moment at*

Work. Drawing on qualitative case studies, Hallowell emphasized that face-to-face interactions, which he termed "human moments," are critical for building trust, collaboration, and innovation in the workplace. He argued that authentic interpersonal connections provide emotional resonance often missing in indirect communication, such as emails or surveys.

Team psychological safety provides a model for understanding how safe environments enable bottom-up communication. Research involving 51 manufacturing teams revealed that when employees feel secure expressing concerns without fear of retribution, team performance and engagement significantly improve (Edmonson, 1999). Open communication channels have also been shown to drive employee engagement and innovative outcomes, further reinforcing the importance of bottom-up communication (De Dreu & West, 2001). Feedback technologies, such as 360-degree feedback and coaching, enhance engagement and performance systems when feedback is relevant, timely, specific, and understandable, fostering trust and alignment with organizational goals (Baker, 2010). Effective feedback must meet these criteria to maintain engagement and improve performance.

This body of research underscores the importance of fairness and trust in traditional feedback mechanisms while providing a foundation for understanding the transition to modern, technology-driven systems, as explored in the upcoming sections.

2.2.2 Traditional Surveys

Traditional annual surveys have been a cornerstone of organizational feedback mechanisms, providing a structured method for collecting employee perspectives on workplace satisfaction, engagement, and climate. Surveys that align with principles of organizational transparency have been shown to positively influence perceptions of fairness, fostering trust and engagement, while inactions following survey results can lead to skepticism and disengagement (Nadler & Tushman, 1980, Burke & Litwin, 1992).

Surveys can serve as trust-building tools when administered effectively, but the time gap between annual surveys often leads to delayed action, which can diminish employee engagement and undermine trust in the process (Church & Wacławski, 1998). Pulse surveys, introduced to capture real-time insights into employee sentiment, address limitations of traditional surveys by

offering shorter, more frequent feedback cycles. However, there is a risk of survey fatigue, emphasizing the importance of effective follow-up and clear communication (Welbourne, 2016). The success of pulse surveys, while improving real-time feedback collection, depends on organizational transparency and actionable responses to maintain trust and engagement (Kim et al., 2021).

This shift from static, periodic surveys to more iterative and responsive feedback mechanisms provides a foundation for understanding the appeal of AI-driven systems.

2.2.3 AI in feedback and bottom-up communication

With the speed of the development of AI, it is difficult to stay up to date to the latest research on AI in HR. In some aspects, what was relevant just a year ago is no longer relevant. Nevertheless, there are transcendent themes that do not depend as much on incremental technological developments and it is precisely these that will be more relevant to our study.

The advent of artificial intelligence (AI) has transformed feedback mechanisms, introducing new efficiencies while raising questions about fairness, transparency, and human interaction. Algorithmic systems, if coupled by a reduction in human involvement, limit employees ability to appeal to decision-makers, thereby reducing accountability and addressing complex emotional factors (Kellogg et al., 2020; Lee et al., 2015). Such limitations highlight the challenge of balancing efficiency with fairness, as algorithms often fail to account for nuanced human emotions and motivations (Lee et al., 2015). Over-reliance on AI in sensitive HR contexts underscores the importance of human oversight to ensure fairness and accuracy (Lee et al., 2018; Binns et al., 2018).

Concerns about dehumanization in AI-driven systems stem from employees perceiving AI as less capable of understanding human complexities, which can lead to skepticism and distrust (Binns et al., 2018). Studies suggest that transparency and communication are critical for mitigating these issues. For example, employees are more likely to trust AI systems when their processes are well-explained, and efforts are made to address algorithmic bias (Bankins et al., 2022; Angerschmid et al., 2022). These findings emphasize that while AI enhances feedback processes,

its success depends on integrating transparency, explainability, and human judgment to address fairness concerns effectively.

2.2.4 Taking on a Process-Perspective

A process perspective emphasizes the need to evaluate the entire lifecycle of feedback systems - from data collection to decision-making - rather than isolating specific components.

Ambrose and Schminke (2009) in their study of 330 employees across 54 organizations demonstrated that overall justice perceptions mediate the effects of distributive, procedural, and interactional justice. Focusing on a unique justice facet in analyzing justice in organizations, may underrepresent the total understanding of justice perceptions (Hausenstein, McGonigle, & Flinder, 2001). Their findings underscored the importance of embedding feedback processes within broader organizational justice frameworks to enhance fairness and trust. Binns et al. (2018) advanced the process perspective by exploring algorithmic decision-making in feedback systems. Their research demonstrated that perceptions of fairness in AI systems arise from the interplay between human and algorithmic roles. They emphasized the importance of taking the full context into account, arguing that fairness perceptions are shaped not only by the algorithms themselves but also by how they are integrated into existing feedback processes.

2.3 Research Gaps

The research gap driving this study lies in the limited research specifically focused on AI-based employee feedback surveys and their impact on employee perceptions of fairness, trust, and engagement. While existing studies have extensively explored employee perceptions in traditional feedback mechanisms and surveys, highlighting the importance of interpersonal relationships, transparency, and psychological safety, these findings are predominantly rooted in conventional, human-driven contexts. Similarly, research into AI in feedback systems has primarily centered on algorithmic decision-making and fairness perceptions in broader HR applications, such as performance evaluations or recruitment processes, rather than focusing on surveys as a tool for bottom-up communication. What is notably scarce, however, is research that bridges these two domains: applying insights from traditional feedback and survey studies to the

emerging and unique context of AI-driven feedback surveys. The implications of these dynamics remain underexplored.

2.4 Limitations in Literature Review

The literature review faced challenges due to inconsistent terminology, with studies using varied terms like "employee feedback surveys," "feedback questionnaires," and "employee engagement surveys," complicating source identification. This fragmentation highlights the difficulty in mapping the field systematically. Additionally, many studies identified through Boolean search commands focused on surveys as a methodological tool rather than as the central concept under investigation, further complicating the search for relevant literature.

3 Theoretical Framework

Research on both traditional feedback mechanisms and bottom-up communication, as well as studies on the impact of AI in human resources (HR), highlight a common theme: human dignity. Traditional research on employee surveys and bottom-up communication emphasizes the importance of understanding and listening to employees, which is rooted in principles of justice and human dignity. This understanding ultimately influences employee engagement. Similarly, recent studies examining the effects of AI also focus on human dignity, particularly in relation to fairness, trust and engagement (Bankins et al., 2022; Angerschmid et al., 2022). Two key themes emerge from this exploration: the factors that contribute to perceptions of fairness and trust (the *how*), and the reasons behind their development (the *why*).

3.1 Fairness Heuristic Theory: How perceptions of fairness/trust can arise

The heuristics and biases approach, developed by Kahneman and Tversky in the 1960s and 1970s, explains how people often rely on mental shortcuts, known as heuristics, to make judgments (Kahneman, D., & Tversky, A., 1982). By concentrating on critical key aspects in a domain, people make inferences about other domains where direct information is lacking. Heuristics simplify the decision-making process of individuals, allowing them to conserve cognitive resources and boost confidence in their actions. Griffin et. al. (2012) argue that

heuristics have often been misinterpreted as irrational shortcuts and that Kahneman and Tversky, who developed the heuristics and biases approach, never claimed heuristics to be inherently irrational. Instead, they described heuristics as useful mental shortcuts that can often yield good results but must also be prudently cautioned against leading to systematic errors.

The general heuristics approach was developed into Fairness Heuristic Theory (FHT) that specifically explains how individuals rely on perceptions of fairness to make judgements about aspects such as trust and legitimacy (Lind, 2001; Van den Bos, Lind, & Wilke, 2001). Research on fairness in the workplace has commonly been categorized into two main frameworks: procedural and interactional justice.

Procedural justice relates to the structure, factors such as transparency, voice and accuracy (Folger & Greenberg, 1985; Greenberg & Folger, 1983). Schminke et al., (2000) review 30 years of empirical research linking justice with structure, illustrating how structure plays a direct, mediating, and moderating role in the relationship between justice perceptions and outcomes such as trust and engagement. For example, if a process is transparent (procedural justice) or if individuals feel they are treated with respect and dignity (interactional justice), this may signal that the manager can be trusted which in turn leads to improved job performance (Colquitt, LePine, Piccolo, Zapata, & Rich, 2012). Hakonen and Lipponen (2008) found that a decrease in face-to-face interaction leads to ambiguity in interactional justice, leading people to rely more strongly on procedural justice information.

From the procedural justice perspective Bies and Moag (1986) developed interactional justice as an extension, specifically focusing on interpersonal treatment, such as respect and being treated with human dignity. They emphasized interactional justice as a foundational aspect of procedural justice after observing that even if the process is fair but the interpersonal treatment is not, individuals may still perceive the process as unfair. Further relational models of procedural justice, such as the *relational model of authority* (Tyler & Lind, 1992) and the *group engagement model* (Tyler & Blader, 2000) similarly highlight the interpersonal role in procedural justice, showing that people care about and react to procedural justice because of the relational message

conveyed by a fair process. This corroborates the findings of Elton Mayo who in the famous Hawthorne studies concluded that organizations are social in their nature (Mayo, 1933).

3.2 Human Relations Theory: The role of interpersonal justice in the process

In the era of the post-industrial revolution, the Harvard Business School, under the leadership of Elton Mayo (Mayo, 1933) and his associates, conducted research on the Hawthorne plant of the Western Electric Company (also known as the Hawthorne studies). His research led up to the first systematic conception of organizations as social systems. It destroyed some of the basic assumptions of the machine model and laid the groundwork for Human Resources. His findings on the role of interpersonal relations in the process lead up to the Human Relations Theory (HRT), which emphasizes that employees are not merely "mechanical wheels" in a system but individuals who are due human dignity. This finding, Bies argues, should underpin all decision-making because it forms core properties of interactional and procedural justice; human dignity (e.g. respect, recognition) and truth (e.g. honesty, transparency) (Bies, 2001). After three decades of research on interactional justice, Bies (2015) concluded that truth and human dignity are at the core of all theories on interactional justice. These findings extend into the procedural justice framework, in how decision-making procedures communicate information about whether people are valued by the organization and its managers and whether these managers can be trusted (Tyler & Lind, 1992). They argue that decision-making procedures carry symbolic meaning regarding people's social identity, highlighting the role of interpersonal justice in the process. The HRT's emphasis on treating individuals not as mechanical wheels in a machine, but as individuals who are due human dignity, underpins the core findings of Bies's overview of organizational justice in the workplace (Bies, 2015).

3.2.1 Relevance of HRT

It should be noted that the HRT preceded procedural and interactional justice theory. However, its relevance remains as the findings explain the foundational aspects of overall justice, which both procedural and interactional justice build on. Ambrose and Schminke (2009) demonstrated that overall fairness perceptions fully mediate the effects of procedural and interactional justice. Employees do not isolate their experiences into rigid categories; instead, they combine aspects of

procedural and interactional justice to form a general sense of fairness that drives their attitudes and behaviors. It is this overall sense of fairness and its key aspects that we seek to study, and this is precisely what the HRT addresses. Bies in his study on interactional justice concluded that the core properties of interactional justice cumulate into two principles; truthfulness and human dignity (Bies, 2015). The HRT arrives at the exact same conclusion, providing a general theory that we argue is still relevant to this day. The HRT-principle adds depth by providing insights into the synthesis of why fairness perceptions arise, which adds guidance as to why a feedback system could either aid or harm perceptions of overall fairness in the feedback process. If we can understand not only how they arise, but also the core principles behind how and *why* they arise, we are better situated to understand perceptions of fairness/trust and their impact on engagement. The procedural and interactional justice frameworks themselves do not by themselves address this synthesis which is central in the overall justice perceptions. This is a major drawback because it limits our ability to distinguish key aspects leading up to overall fairness perceptions.

The HRT thus acts as a complement to the categorical frameworks, adding nuance and precision into our analysis. It can be used to examine whether AI's role in the feedback system upholds key aspects of human dignity and truthfulness, enabling a more robust assessment of its potential to sustain fairness, trust, and engagement. This goes in line with our strategy to apply traditional research on feedback on AI. The downside of using a theory that does not directly deal with AI, is substantially mitigated by the fact that the underlying questions faced in the post-industrial revolution era as a consequence of feelings of dehumanization, have resurfaced with a growing need to address the fundamentals of human dignity in the workplace (Banking et al., 2022; Lee et al., 2015). The fundamental questions the HRT answers are the same fundamental questions modern research highlights as foundational in justice perceptions, and these are the questions we ask in our analysis.

4 Research Methodology

4.1 Research Philosophy and Approach

4.1.1 Research Paradigm

This study is based on an interpretivist paradigm, which focuses on understanding subjective meanings individuals assign to their experiences (Saunders et al., 2020). The paradigm is particularly suited to this research since it explores how employees perceive fairness in AI-driven decision-making processes. Instead of looking for objective truths, the interpretive paradigm emphasizes that reality is not fixed but constantly shaped and reshaped through individual interpretations (Bryman, 2016). In terms of the research context, the interpretivist paradigm helps explain how employees' subjective views on fairness, trust and engagement emerge in response to the AI-based employee feedback system. By prioritizing participants' lived experience, the paradigm allows us to uncover insights into how the new system is perceived.

4.1.3 Research Approach

The research adopts an abductive approach, combining both deductive and inductive elements (Bell et al., 2019). While existing theories will provide the theoretical foundation, the study remains open to themes that may challenge these theories. This is important as unexpected issues may arise during the interviews that were not anticipated in the initial theoretical framework. Employees might introduce new topics outside preselected theories, ensuring that research remains flexible. This abductive approach ensures a deep and nuanced understanding of how AI-based employee surveys affect the perception of decision-making within the organization.

4.2 Research Design

4.2.1 Qualitative Research Design

A qualitative research design was chosen as it is best suited to explore the complex, socially constructed nature of employee perceptions. Since the study seeks to understand subjective experiences and interpretations of fairness, a qualitative approach allows for deeper insights that

are difficult to capture through numerical data. Using a quantitative method, such as a survey, would not be appropriate in this context, as it would reduce the nuanced perspectives to numbers, ironically making the research process similar to the very type of system it aims to examine. Qualitative methods offer the flexibility to capture the richness of participants' viewpoints, which is essential for understanding the complexity of their experiences (Saunders et al. 2020). However, the reliance on interviews means that findings may be influenced by the subjective interpretations of both the participants and the researchers. Additionally, qualitative methods are typically more time-consuming and involve smaller sample sizes, which may limit the generalizability of the findings (Bryman, 2016). To address this, efforts have been made to ensure rigorous data analysis through thematic coding and reflexivity to minimize researcher bias.

4.2.2 Semi-Structured Interviews

Semi-structured interviews are selected as the data collection method, as they allow for flexibility while exploring key themes related to employee perceptions. This method enables asking both prepared and follow-up questions, adapting to the participants' responses to understand their experiences and perceptions. However, the flexibility of the semi-structured interviews can lead to inconsistencies across interviews, as different participants may be asked during follow-up questions depending on their answers. This can make the comparison of the results more difficult. Additionally, the reliance on verbal accounts may result in limited insights into non-verbal cues, especially as interviews are conducted digitally (Bell et. al 2019). These limitations will be mitigated by using a well structured interview guide and taking notes of non-verbal cues during the interviews.

4.2.3 Design of Interview Questions

The interview questions were designed using an “inside-out” approach, starting with themes that are more immediately accessible to participants - trust and engagement - before addressing the more abstract concept of fairness (see appendix 1). This strategy was informed by the understanding from previous theory and research that fairness is the foundational concept influencing trust and engagement (Colquitt et al., 2001; Ambrose et al., 2009). This allowed employees to articulate their experiences and reflections more naturally, which then provided a

basis for exploring fairness as the underlying theme. The organization of the empirical results follows the same progression - trust, engagement and perceptions of fairness - reflecting both the natural flow of discussions and the prominence of each theme as it emerged in the data.

4.3 Data Collection

4.3.1 Sampling Strategy

This study uses a purposive sampling strategy (Emmel, 2013), targeting participants with direct experience using AI-based employee feedback surveys within Bonnier News Local, a business area under Bonnier News. Bonnier News Local, established in 2020, comprises around 40 local newspapers (Bonnier News, n.d.). To maintain anonymity, the specific organization within Bonnier News Local included in this study will not be named.

The sampling strategy aims to capture diverse perspectives while focusing on employees who have experienced the transition from traditional annual surveys to the AI-based system. The final sample consists of 10 participants, evenly divided between men and women, to ensure gender diversity. Participants were also selected across various age ranges, from 35 to 65, to provide insights from different career stages. We contacted a manager within the chosen organization, who helped identify and reach out to employees interested in participating. This process allowed us to specify preferences for diverse demographics, including age, gender and managerial oversight. Potential participants were connected via email (see Appendix 2), where we broadly described the study aim. Out of 11 potential interviewees, 10 responded, and interviews were scheduled (see Appendix 3 for detailed interview sample).

While Bonnier News Local provides a rich context, there are limitations to generalizing findings across the entire Bonnier News organization or other subsidiary. Employees in smaller organizations under Bonnier News Local may experience the AI-based system differently than those in other subsidiaries, influenced by factors such as organizational size, culture and available resources. Additionally, participants are relatively new to the AI-based system and have not used it for an extended period, which may limit their ability to provide comprehensive reflections. However, efforts were made to mitigate this limitation by ensuring participants were selected from different departments and worked under various managers. This diversity

strengthens the representativeness of the findings. Additionally, while differences in perceptions across subsidiaries may exist, the shared technological framework of the AI-based system provides a foundation for comparable experiences. Further research could explore potential variations across subsidiaries and regions to address this limitation, as well as perceptions of employees with longer-term exposure to the system, to address these limitations.

Despite the benefits of the purposive sampling strategy, qualitative research involves smaller sample sizes, which may limit the study's broader applicability (Bryman, 2016). The sample size of 10 participants is, however, aligned with qualitative research norms (Creswell & Poth, 2018).

4.3.2 Interview Process

All interviews have been conducted digitally via Google Meet, as the researchers and participants were located in different locations. While in-person interviews can provide a more enriching exchange and capture non-verbal cues, remote interviews offer greater accessibility and convenience for both parties (Opdenakker, 2006). Each interview lasted approximately 20 to 40 minutes and was audio-recorded using the Voice Memo feature on an iPhone, with the participants' consent, to ensure accuracy.

An interview guide structured the conversation to maintain consistency across all interviews, while allowing flexibility for follow-up questions based on participants responses. The guide covered five key areas: background information, awareness of Ai in employee surveys, anonymity and frequency, and perceptions of the transition to AI-based surveys (see Appendix 1). This structured, but adaptive, approach ensured that interviewees captured comprehensive insights while exploring themes relevant to the research question.

All quotes used in the analysis have been translated from Swedish to English. Although we carefully worked to keep the original meaning and context, some nuances in language and tone may not have been fully captured. To mitigate this risk, translations were cross-checked with the original Swedish transcripts to ensure they accurately reflected what participants intended to say.

4.4 Thematic Analysis

Thematic analysis was conducted to identify key themes within the interview data, aligning with the study's abductive approach. First, interviews were transcribed using the software KlangAI. To ensure accuracy and familiarity with the data, we listened to and read through the interviews, and identified initial codes. The transcripts were then imported into Maxqda24, a qualitative analysis software, where both pre-identified and emergent codes were applied. Quotes and specific codes were subsequently organized into three broader thematic areas: trust, engagement and perceptions of fairness. While Maxqda was used for coding, the thematic analysis itself was conducted manually, as this approach allowed for deeper interpretive engagement with the data.

The process of open coding, focused coding, and axial coding, aligns with the abductive approach by integrating deductive and inductive elements (Williams & Moser, 2019). Pre-identified codes reflect deductive reasoning, while emergent codes capture unexpected themes raised by participants, demonstrating the inductive flexibility of the analysis.

4.5 Research Quality

4.5.1 Rigor

To ensure the quality and credibility of the research, multiple measures were taken to uphold rigor. Thematic analysis was conducted, allowing for a thorough examination of patterns and themes in the data. The use of Maxqda24 ensured the organization and management of codes, but manual interpretation was maintained to foster a deeper understanding of the data. Triangulation was employed through cross-referencing participant responses and aligning them with the theoretical framework, ensuring consistency and reliability. The structured interview guide (see Appendix 1) further enhanced rigor by maintaining a uniform approach across interviews while allowing for flexibility in exploring unique participant perspectives.

4.5.2 Reflexivity

We actively reflected on our potential biases and their influence on the interpretation of the data. Maintaining a reflexive journal during data collection and analysis allowed us to document our

thought processes, decisions, and any assumptions that could shape the findings. This practice aimed to reduce subjectivity and enhance transparency in the research process.

4.5.2 Ethical Considerations

Ethical considerations were considered to protect participant rights and ensure the integrity of the research. Participants were provided information about the study's objectives and methods and were asked to sign informed consent forms. Confidentiality and anonymity were upheld by not disclosing participants' names, age, roles, or identifying the specific organization within Bonnier News Local. All data was securely stored, and access was restricted to the research team.

4.6 Usage of Artificial Intelligence

Artificial Intelligence (AI) tools were used during the research process to enhance efficiency and accuracy in specific tasks, while ensuring that the integrity and rigor of the study were maintained. ChatGPT was used to refine the language and academic tone of the written text. This tool helped improve the presentation, without adding to or altering the content of the research, ensuring it remained solely the work of us. Additionally, KlangAI was used to transcribe the interview recordings, providing an initial draft of the verbal data. To ensure transcription accuracy, we reviewed the transcriptions while listening to the audio recordings, correcting any errors to maintain alignment with the original material. These AI tools streamlined specific aspects of the research process.

5 Empirics

This section explores the key themes identified from employee interviews regarding their perceptions of the AI-driven feedback system, focusing on trust, engagement, and perceptions of fairness in the order they emerged most prominently during the interviews. This organization reflects the natural flow of the discussions and allows for a detailed exploration of each theme as employees shared their experiences and reflections.

5.1 Trust

Trust plays a significant role in the effectiveness of any feedback system. This section examines how employees' perception of the system is influenced by their awareness of AI's involvement, their confidence in managerial oversight, and the transparency regarding the processing and utilization of feedback.

5.1.1 Awareness of AI

Awareness of AI's role in analyzing and processing feedback emerges as a critical aspect of employees' trust. To assess this, we asked participants if they were aware about AI's involvement in the system. For those who were unaware, we provided a brief explanation to ensure they could share informed reflections during the interview.

Employees who were initially unaware of AI's involvement felt that managers could have communicated this more clearly. They assumed the system functioned like earlier feedback tools, reflecting a lack of transparency rather than a criticism of the technology itself. Despite this, they reacted positively or indifferently upon learning about AI's involvement. They did not express significant skepticism toward AI, attributing this to their strong trust in management. This trust encouraged employees to focus less on AI's technical aspects and more on confidence in organizational processes.

“I wasn't aware that the system used AI and I don't really know what it means. I'm not negative toward it, but I don't know how much AI and how much humans are involved. [...] Since I trust that the managers still evaluate the suggestions themselves, I don't see an issue with it.” - Clara

“When talking with colleagues about the survey, it's often about practical things, like how frequently we receive the emails. [...] I've never heard anyone reflect on the fact that it's actually AI behind it.” - Lara

In contrast, employees aware of AI's involvement displayed more critical views. While they did not dismiss the potential benefits of AI, they questioned the system's transparency and the necessity of AI in processing employee feedback. Many informed employees were skeptical

about whether AI could adequately interpret the complexity of qualitative feedback. They also noted a lack of clarity on how AI personalized questions or generated insights, which added to their reservations.

“The results are apparently gathered by an algorithm that asks different questions each time, but I don’t know if everyone gets the same questions simultaneously. I also haven’t noticed that the questions reflect my previous answers.” - Erik

“The exact details about how AI is programmed in this system aren’t entirely clear” - Karl

Ultimately, the findings suggest that while awareness of AI is essential for informed engagement, it can also challenge trust in the system by encouraging critical reflection. Nonetheless, regardless of their awareness level, employees’ trust in the overall system remained tied to their trust in management - a theme explored further in the next section.

5.1.2 Trusting the People Behind the System

While employees’ trust in the AI-driven feedback system varied based on their awareness and understanding of AI, a critical insight emerged: trust in the system ultimately hinges on trust in human managers. Regardless of employees’ perceptions of AI - whether skeptical, positive, or indifferent - their confidence in the feedback process was rooted in the belief that managers play a critical role in interpreting and acting on the results.

Employees widely agreed that while AI could collect, organize, and highlight trends in feedback, it could not replace the human ability to interpret and act upon the results. Many employees felt that the system’s credibility required a distinctly human touch. Employees were uncertain whether the feedback process could deliver meaningful outcomes without this human intervention.

“I think it’s good to have both anonymous feedback and open conversations. [...] But personal conversations are often more honest and clearer. [...] I’m a bit worried that AI-based surveys might reduce the personal dimension.” - Anna

This highlights a fundamental limitation in how employees perceive AI. While they see it as an efficient tool, its outputs hold value only when humans review and act on them. Employees need reassurance that managers carefully evaluate the feedback and make informed decisions based on it.

For employees, trust in the system stemmed not from the technical aspects of AI but from their confidence in managers to contextualize and act on the data. Even those skeptical of AI expressed trust in their managers to bridge any perceived gaps in the system's capabilities. This reliance on human oversight reflects a broader trust in leadership rather than in the AI system itself.

“It feels a bit strange that a computer would initiate a decision, but I trust that the managers still assess whether the suggestion is relevant and put their human touch on it.” - Clara

Employees trust managers, but this trust comes with a condition: they want to see managers actively involved in evaluating and using the feedback. Transparency about *how* managers play their role in this process is essential. We will explore this further in the next section.

5.1.3 Transparency and Its Impact on Trust

While awareness of AI influenced employees' perceptions to some degree, the desire for clarity about how feedback was processed, interpreted, and used proved to be a far more significant determinant of trust. Regardless of employees' awareness or skepticism about AI, the need for visible human involvement and transparency was consistent across all groups.

Employees consistently emphasized that managers play a critical role in meaningfully evaluating and acting upon feedback. While they generally trusted managers to handle results responsibly, they strongly desired more transparent communication about this process. Many employees felt that the lack of clarity undermined their confidence in the feedback system, which left them uncertain about the actual impact of their input. Their trust in the system stemmed not just from

AI's capabilities but from the visible actions of human managers. Employees sought reassurance that their managers thoroughly reviewed, contextualized, and ethically acted on their feedback.

“I think trust in the algorithm is crucial, but we don’t have any insights into how it’s calibrated or when it’s adjusted. [...] It’s important to know how the feedback is actually used.” - Anna

5.2 Engagement

Engagement also emerged as a critical factor in the success of the AI-driven feedback system. This section explores the dynamics of engagement, focusing on how the balance between system design, organizational culture, and communication shapes employees' willingness to participate meaningfully.

5.2.1 Balance Frequency with Survey Fatigue

One of the most debated aspects of the feedback system was the frequent distribution of surveys. Some employees appreciated the opportunity for regular input, feeling that the system's real-time nature allowed for timely and relevant feedback. By maintaining a high survey rhythm, employees could address issues while they were still fresh in their minds, leading to a sense of immediacy and alignment with ongoing workplace dynamics.

“Receiving the survey every three weeks makes it feel current and keeps events fresh in my memory. The frequent setup suits me well; it provides a better opportunity to influence my situation compared to the old system” - Karl

Others emphasized the shortness of the surveys, noting that the streamlined design enabled them to respond without disrupting their workflow. The simplicity of the questions helped maintain focus and minimized the cognitive load associated with providing feedback.

“Now, you don’t get a long series of questions where you can lose interest after a while. [...] You can actually stay focused from start to finish here.” - Albin

However, not all employees viewed the high frequency of surveys positively. Many expressed frustration with what they perceived as survey fatigue, citing the repetitive nature of the process. For these employees, the frequency diluted the significance of each survey, leading them to view it as "just another task" rather than a meaningful opportunity to share insights. In some cases, employees admitted to completing surveys only to dismiss the reminders, which frequently filled their email inboxes, raising concerns about the authenticity of their responses.

"I complete the surveys just to get rid of the reminders on email." - Patrik

While the connection between AI and survey frequency may not have been immediately apparent to all employees, the system's design and its ability to process feedback efficiently were enabled by AI technology. One employee mentioned that the frequent distribution of surveys was a product of AI's capacity to analyze data quickly, allowing the system to generate actionable insights in near real time. However, the relationship between AI and engagement remains nuanced.

"Now the surveys are sent out frequently, which I understand is possible thanks to AI, and the results are analyzed quickly." - Anna

The findings highlight the need for a balance between frequent feedback opportunities and survey fatigue. While some employees appreciate real-time feedback, others feel that high frequency may reduce overall engagement. To enhance effectiveness, the system must clarify the purpose and outcomes of these surveys, ensuring employees see their input as meaningful. Effective communication about the value of frequent surveys is crucial to sustain long-term engagement while mitigating fatigue.

5.2.2 Anonymity Concerns

Anonymity emerged as a factor influencing employees' engagement with the AI-driven feedback system, particularly the open-text comments. The system first asks employees to select one of four emojis representing different levels of sentiment - ranging from very positive to very negative. After selecting an emoji, employees can elaborate further with an open-text comment.

Managers can respond directly to these comments within the system but do not know the commenter's identity.

Many employees appreciated the opportunity to provide anonymous feedback but expressed concerns about being identified through their writing style, specific content, or how they expressed themselves. The ability of managers to respond directly to comments within the system amplified this worry. While the managers could not see who had written the comment, employees felt "seen" in a way that undermined their perception of anonymity.

“I noticed the first time I used it, my manager went in and started responding with their name. [...] It felt really uncomfortable when that happened in real-time. I felt immediately seen and identified.” - Tina

This led some employees to avoid leaving detailed or sensitive comments, especially in smaller teams where identifiability was more likely. For these employees, the risk of being recognized - real or perceived - was enough to diminish their willingness to fully engage with the survey.

“If I write that I’m the youngest or oldest in the editorial team, they’ll probably figure out who it is. Writing style also plays a role, as they might recognize the way I express myself.” - Lara

However, for some employees, anonymity was not a significant factor influencing their engagement with the feedback system. These individuals highlighted the open and communicative culture within their workplace, which fostered an environment where they felt comfortable expressing their opinions directly to their managers.

“I wouldn’t say anonymity is particularly important to me. The things I bring up in the survey are things I’d just as easily say openly in a meeting” - Johan

5.2.3 Communication and Follow-Up from Managers

A significant issue highlighted by employees was the need for follow-up discussions after survey submissions in the AI-driven system. In the previous annual survey system, managers frequently

organized meetings to discuss the results, asked clarifying questions, and outlined specific actions based on the feedback. This structured approach gave employees confidence that their input was taken seriously and used to inform meaningful changes.

“Previously, we had discussions around the results, and the discussions were often more important than the actual numbers, helping identify what needed to be done. With Winningtemp, because the surveys are so frequent, the type of in-depth conversation is missing” - Anna

In contrast, the current system seemed to lack visible acknowledgment of feedback, which left many employees feeling disconnected from the process. Some questioned whether their input led to concrete actions, reducing their motivation to engage with the system. Without this communication loop, employees struggled to see the value in providing feedback

“It feels like our opinions don’t have any value, and that we’re not being listened to. We want to know that what we’re saying actually makes a difference.” - Lara

Employees emphasized that they need to see a clear connection between their feedback and management's actions. When managers fail to communicate how they use this input, employees feel their contributions are overlooked. This lack of acknowledgment leads to disengagement and even resistance to the feedback process.

“If adjustments are presented without showing how they relate to employee feedback, it can create resistance. It’s important for us to see a clear connection between what we raise in our responses and the actions that are taken.” - Lara

Some employees also described the current system as overly focused on collecting data rather than fostering meaningful dialogue. The absence of in-depth discussions made the process feel impersonal and transactional, reducing their engagement with the surveys.

“You can easily become a slave to the system instead of having natural conversations” - Erik

5.3 Perception of Fairness

Following the discussions on trust and engagement, fairness emerged as another central theme in employees' perceptions of the AI-driven feedback system. This section presents employees' views on how fairly the system interprets and represents their input, the impact of survey frequency on representation, and the role of transparency in addressing concerns about equitable treatment.

5.3.1 Fairness in Feedback Interpretation

A significant advantage of the AI survey system was its efficiency, which many employees valued for its quick and straightforward approach to gathering input. This efficiency was a notable improvement over previous, more time-consuming systems. However, employees also expressed reservations about whether the AI could adequately interpret complex and emotional nuances in their feedback.

"The questions are clear and simple, with only five or six at a time, making it easy to focus without losing interest. It feels like a smooth and efficient setup." - Albin

"AI can signal that something isn't going well, but reality is more complex. It's about understanding the reasons behind dissatisfaction. That's something the system can't determine."
- Erik

The limited response options, particularly the use of four sentiment-based emojis, amplified these concerns. Employees felt that the lack of a neutral or middle-ground option forced them to choose polarized responses that might not fully reflect their perspectives.

"Now we have fewer response options compared to before, which forces a choice between two negative and positive answers. I know many are frustrated by this because it doesn't always reflect how they really feel. A middle-ground option is missing." - Anna

Additionally, the predefined sentiment emojis influenced how employees formulated their open-text comments. One employee noted that selecting a particular sentiment shaped the tone of their responses, potentially limiting their ability to fully express themselves.

“The choice of the four faces likely influence how one formulates their comment. [...] Even though the faces give a quick overview, there’s a risk that they constrain thoughts and limit the ability to fully develop the response.” - Johan

This highlights that there is a tension between the system’s streamlined design and its capacity to represent employees’ experiences fairly.

5.3.2 Perceived Fairness in Representation

The frequency of surveys generated mixed perceptions about fairness. Some employees valued the frequent surveys for providing a more consistent and balanced view of workplace sentiment over time. They believed this regularity helped reduce the influence of temporary mood swings, leading to a fairer overall representation.

“It’s good that questions are asked often because it can balance out those temporary responses.”
- Anna

However, others expressed concerns that survey fatigue led to declining participation rates which undermined the system’s fairness. Employees noted that when fewer individuals responded, the results became less representative of the broader workplace sentiment, raising doubts about the accuracy and equity of insights derived from the data.

“I wonder how much you can trust the results if, for example, only half of the employees respond.” - Erik

5.3.3 Transparency as a Component of Fairness

Employees emphasized that fairness perceptions could be strengthened by increasing transparency around how the AI system processes and uses their feedback. Many noted that the

lack of clarity on how individual responses contributed to broader trends and actions created skepticism about whether the system treated all input equitably.

“It would be helpful if management communicated more clearly how they use the results from the employee survey. Right now, we don’t know if the system handles the responses fairly or if it oversimplifies or misses important nuances.” - Lara

“I’m unsure if the system captures nuances and my state of mind in a fair way.” - Anna

Employees wanted assurance that the AI system not only processed data efficiently but also avoided oversimplifications or biased patterns in its follow-up questions and actions.

5.4 Conclusion

While the empirical findings provide detailed insights into how trust, engagement and perceptions of fairness shape employees’ perceptions of the AI-driven employee feedback system, the interconnectedness of these themes requires further exploration. In the following analysis chapter, we draw on theoretical frameworks to examine how these themes influence and reinforce one another.

6 Analysis

According to FHT, individuals form trust based on perceptions of fairness, and this trust subsequently influences engagement. Central to the analysis will be identifying key explainable variables that underpin this fairness-trust-engagement link and examining how AI affects these dynamics. It is essential to address the observation that the findings related to fairness in the empirical section are less developed than those on trust and engagement. This is partly because these themes are deeply interconnected, and much of what emerged during the interviews - particularly regarding trust and engagement - can be rooted in or connected to perceptions of fairness, even if fairness itself was not explicitly discussed. The interviews did not frequently focus on “fairness” in isolation; instead, the interviewees often linked their experiences and reflections to trust, implicitly touching on fairness without framing it as such. As a result, while

the fairness theme may seem less pronounced in the empirical section, it remains a critical underlying element.

Ambrose and Schminke (2009) demonstrated that overall fairness perceptions fully mediate the effects of procedural and interactional justice. Employees do not isolate their experiences into rigid categories; instead, they combine aspects of procedural and interactional justice to form a general sense of fairness that drives their attitudes and behaviors. A categorical analysis strictly divided into procedural and interactional justice would oversimplify perceptions of fairness by not taking into account their relation and interconnectedness. The HRT principle of integrating interactional justice into procedural justice helps us interpret the findings through an overall sense of justice - and when needed - this is complemented with separate analysis of the individual justice frameworks. What follows is a list of key justice variables and the subsequent analysis leading up to them.

6.1 The human role central in overall fairness

Without human intervention in interpreting the output of the AI, employees expressed distrust in the outcome. The human involvement acted as a fairness heuristic that mitigated distrust in AI. The trust in human oversight reflects the relational nature of trust. Employees saw human managers as more capable of understanding complex, nuanced feedback due to their ability to empathize and contextualize human experiences – qualities algorithms lack (Lee et al., 2015).

Employees who trusted their managers also trusted that the feedback system would be used ethically, even when aware of the black-box nature of the AI (how the operations of the AI are not visible). The uncertainty of procedural justice in the process could have prompted employees to emphasize interactional justice, in line with FHT (Hakonen & Lipponen, 2008). More specifically, individuals in the absence of information from one justice framework, may draw inferences based on justice information from another framework. The focus on trust in managers became especially prevalent among employees aware of AI involvement, possibly prompted by their awareness of the process's uncertainty.

It is here worth mentioning that it is not always the case that human-decision making is assumed to result in greater trust, such as in the case of the *machine heuristic*, where algorithmic systems are assumed to be less biased (Araujo et al., 2020), leading to greater perceptions of trust compared to a human decision-maker. While we largely did not find the *machine heuristic* to hold true in our context, the reason for it may be grounded on the same fundamental aspect (or rather the absence of it) as in the machine heuristic - that of truthfulness and accuracy.

One way of analyzing the role of human involvement through the lens of interactional justice is through the need for recognition. Recognition could manifest itself in the procedural justice framework through voice; the ability to express their views and provide input. However, recognition must permeate the formal structure of the employee survey not to lead to a system that feels disconnected and ineffective. The interpersonal justice framework informs understanding of the procedural justice framework, once again relating to the HRT principle. Recognition, in turn, comes down to the same foundational justice variables found to be central in all interactional justice; human dignity and truthfulness (Bies, 2015). Recognizing employees and giving them a listening ear is essential, but if this is done insincerely, they are not truly being listened to. More than providing a tool of feedback is needed, the execution of the tool is also important. The HRT closely relates the nature of feedback systems to the fair treatment of employees, emphasizing the importance of hearing and seeing employees, which is in line with viewing them not as mechanical wheels in a machine but as individuals. The question then arises whether the AI tool aids the feedback process or fails to account for the human abilities, emotions, and motivations appropriate to the situation (Lee et al., 2015).

Our findings reveal that employee feedback survey acts as a supplement rather than a replacement for the human aspect of feedback. Most employees express their nuanced feedback on a day-to-day basis with their managers. The employee-manager relationship was initially strong, and these dialogues largely remain. AI serves the role of complement with the added value of frequency/representation and anonymity. These will be further addressed later in the analysis.

However, the findings also demonstrate that engagement with the feedback system is closely tied to follow-up and communication from managers. Employees wanted more than frequent surveys; they sought reassurance that their feedback was acknowledged, understood, and acted upon. Without this, the system risked feeling like a formality rather than a meaningful tool for change. This reinforces a broader theme: while AI facilitates frequent surveys and efficient data processing, the human role in contextualizing and communicating results remains indispensable for fostering meaningful engagement.

6.2 Transparency

Transparency also emerged as a key aspect of forming trust. This aligns with FHT, which posits that employees are more likely to trust both the process and the decision-makers when they perceive transparency. The Black Box nature of AI presented an uncertainty in the process that was mitigated by trust in managers. However, not everyone knew about the involvement of AI. Those unaware were surprised upon finding out and expressed disappointment that management had not communicated this. With needing more transparency in the process, the foundation of procedural justice weakens, and employees may rely on interactional cues - such as trust in their manager's intent or interpersonal treatment - to fill the gap. Suppose managers are not transparent in communicating how feedback is processed and used. In that case, employees may interpret this as a lack of recognition or respect, undermining trust in the system and the managers themselves. This demonstrates how the procedural and interactional justice frameworks intersect, again highlighting the HRT principle.

6.3 Frequency & Representation

Regular reporting provides a more balanced picture of employee experiences, avoiding the bias from a once-a-year survey, which could capture an unusually good or bad week rather than a true reflection. By spreading feedback across time, the AI tool enables the company to more accurately and consistently listen to its employees, aspects central in the procedural justice aspect of accuracy and the interactional justice aspect of truthfulness.

Employee engagement in the AI-driven employee feedback system reflects their perceptions of fairness and trust. Employees who perceived the system as fair and trusted their managers were

more likely to engage meaningfully. Engagement was not simply about completing surveys; it was about feeling that participation was valued and impactful. As HRT suggests, engagement is driven by a sense of belonging and the assurance that one's voice matters. However, engagement was negatively affected by survey fatigue and a lack of follow-up.

6.4 Anonymity and invasion of privacy

Anonymity played a dual role in shaping engagement. The anonymity of the survey encouraged honest feedback, which may be explained by Schein (1977), who argues that if the procedure is too psychologically intrusive or invasive to the individual, the information gathering may be perceived as an invasion of privacy. However, concerns about identifiability undermined its benefits. For instance, Tina felt uncomfortable when her manager responded directly to comments in real-time, stating, “[...]I felt immediately seen and identified.” This tension highlights the need to balance anonymity with effective communication, ensuring employees feel heard and protected.

7 Discussion

7.1 Revisiting the Research Question

How do AI-based employee surveys influence employee perceptions of trust, engagement and fairness within Swedish Newspaper Organizations?

AI-driven feedback surveys at Bonnier News influence employee perceptions through the foundation of fairness, established by transparent and consistent processes. This foundational fairness fosters trust in the AI systems, which in turn enhances engagement by assuring employees that their feedback is valued and impactful. Maintaining effective human oversight ensures that AI enhances rather than replaces human decision-making, preserving a crucial personal touch. Additionally, clear communication about AI's role and its operational processes strengthens this trust, effectively integrating AI systems into the organizational culture and enhancing their acceptance among employees.

7.2 Comparison with Previous Research

Our findings align with existing theory and research, particularly regarding the foundational aspect of truthfulness and human dignity in shaping perceptions of overall fairness. Ambrose and Schminke (2009) emphasized that overall fairness perceptions mediate the effects of procedural and interactional justice, and our findings mirror this by showing that employees rely on both justice frameworks to form a general sense of fairness. For example, employees valued human involvement in interpreting feedback, aligning with the interactional justice principle of truthfulness (Bies & Moag, 1986) and the procedural justice principle of accuracy. However, it was the intersection of the two frameworks that provided the needed synthesis to understand the individual roles of the two frameworks.

Furthermore, we extend prior research by confirming that studies on traditional surveys and general AI feedback are also applicable to AI-driven feedback surveys. This addresses our identified research gap, suggesting that established theories of fairness perceptions in conventional feedback contexts hold relevance in the context of AI-enabled systems.

An additional contribution of our study is employee awareness of AI involvement, a dimension that has not been extensively studied in prior research. We found that greater awareness of AI's role in feedback processes often led to increased criticism of the system. This suggests that awareness can act as a critical moderating factor in shaping perceptions of fairness and trust, highlighting the importance of clear communication and education about AI systems in organizational contexts.

Our research confirmed the findings of (Hakonen & Lipponen, 2008) by showing that when procedural transparency was insufficient, employees relied more on interpersonal cues to share their perceptions of fairness. This highlights the role of human relations where procedural fairness is lacking. However, this does not imply that we can disregard deficiencies in the procedure, as doing so could transfer the lack of transparency in the process onto the managers, with employees interpreting this as a reflection of how managers treat them (Tyler & Lind, 1992).

7.3 Practical Implications

The primary challenge identified in the empirics was a lack of transparency, which stemmed not only from the black-box nature of the AI but also from employees' uncertainty about how their feedback was utilized. While improving transparency may be difficult, ignoring the issue could result in employees transferring their distrust onto managers, undermining trust in both the feedback process and leadership. This can be managed through clear and honest communication. If the black-box nature cannot be addressed on a process level, it is important to be transparent about this limitation and to explicitly communicate the intentions of the employee survey and the specific choice of the survey to employees. Acknowledging any system drawback is better than obscuring them, as the acknowledgement itself can foster trust in leadership, which in turn can transfer onto the process. Key aspects such as truthfulness and human dignity in simple day-to-day managerial interactions with employees, but also in the approach itself, will not only improve interpersonal relations but also have effects on the process itself. This returns to the key principle of the Human Relations Theory, that organizations are social in their nature and seeing employees not as mechanical wheels in a machine but as individuals who are due human dignity, is foundational.

7.4 Directions for Future Research

Future research should address limitations such as the small sample size, single-case focus on Bonnier News, and participants' limited exposure to the AI system. Broader studies across industries and more prominent, more diverse samples could enhance the generalizability of findings. Emerging challenges also warrant attention, including algorithmic biases, strategies to improve AI-human collaboration, and approaches to mitigate privacy concerns. Central to these challenges is integrating AI into the workplace in a way that upholds and works in harmony with human dignity.

8 Conclusion

This study examined how AI-driven feedback systems influence employee perceptions of fairness, trust, and engagement, using Bonnier News as a case study. Employees valued the

efficiency, anonymity, and frequency of AI-based surveys, which provided more timely and balanced insights than traditional methods. However, concerns about effectiveness, transparency, survey fatigue, and AI's limitations in interpreting nuanced feedback emphasized the continued importance of human oversight. Trust in the system was closely linked to confidence in managers' ability to contextualize and act on feedback. Employees viewed AI as a tool that complements human decision-making rather than replacing it. Transparency in how feedback is processed and used was critical for perceptions of fairness, while strong interpersonal connections with managers remained essential for engagement. Our findings highlight the need for organizations to integrate AI tools thoughtfully, ensuring they enhance, rather than replace, human elements of trust and fairness.

9 References

- Ambrose, M. L., & Schminke, M. (2009). Assessing roadblocks to justice: A model of fair behavior in organizations. *Research in personnel and human resources management*, 28, 219–263.
- Angerschmid, A., Zhou, J., Theuermann, K., Chen, F., Holzinger, A. (2022). Fairness and Explanation in AI-informed Decision Making. *Machine learning and knowledge extraction*, 2022-06, Vol.4 (2), p.556-579.
- Araujo, T., Helberger, N., Kruijkemeier, S., de Vresse, C. H., (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & society*, Vol.35 (3), p.611-623
- Baker, N. (2010). Employee feedback technologies in the human performance system. *Human Resource Development International*, 13(4), 477–485.
- Bankins, S., Formosa, P., Griep, Y., et al. (2022) AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human resource management context. *Information Systems Frontiers*, vol. 24, no. 4, pp. 857–875.
- Bell, E., Bryman, A., & Harley, B. (2019). Business Research Methods. 5th ed., *Oxford University Press*, pp. 45–48, 396–400.
- Bies, R. J. (2015). Rethinking Organizational Justice: Comparative and Noncomparative Principles. In R. S. Cropanzano & M. L. Ambrose (Eds.), *The Oxford Handbook of Justice in the Workplace*. *Oxford University Press*, pp. 93-96.
- Bies, R. J. (2015). Rethinking Organizational Justice: Interactional Justice: Looking Backward, Looking Forward. In R. S. Cropanzano & M. L. Ambrose (Eds.), *The Oxford Handbook of Justice in the Workplace*, *Oxford University Press*, pp. 89-101.
- Bies, R. J. (2001). Interactional (in)justice: The sacred and the profane. In J. Greenberg & R. Cropanzano (Eds.), *Advances in organizational justice*, Palo Alto: *Stanford University Press*, pp. 89–118.
- Bies, R. J., & Moag, J. F. (1986). Interactional justice: Communication criteria of fairness. In R. J. Lewicki, B. H. Sheppard, & M. H. Bazerman (Eds.), *Research on negotiations in organizations* (Vol. 1, pp. 43–55). Greenwich, CT: *JAI Press*.

- Binns, R., Kleek, M., Veale, M., Lyngs, U., Zhao, J., Shadbolt, N. (2018). "It's reducing a human being to a percentage": Perceptions of justice in algorithmic decisions. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Paper 377, pp. 1–14.
- Bobocel, D. R., Gosse, L. (2015,) Fairness Heuristic Theory/Uncertainty Management Model. In R. S. Cropanzano & M. L. Ambrose (Eds.), *The Oxford Handbook of Justice in the Workplace*, p 70-71.
- Bonnier Group AB. "Annual Report 2023" 2023, <https://www.bonnier.com/wp-content/uploads/2024/05/Bonnier-Group-Annual-Report-2023.pdf>. Accessed 15 October 2024.
- Bonnier News. "Journalism is our core." n.d, <https://www.bonniernews.se/home-en#:~:text=The%20Bonnier%20News%20Local%20business,of%20the%20Expressen%20business%20area>. Accessed 24 September 2024.
- Bonnier AB. "Sustainability at Bonnier: Five focus areas." n.d, <https://www.bonnier.com/en/five-focus-areas-within-sustainability/>. Accessed 27 September 2024.
- Bryman, A. (2016). Social Research Methods. 5th ed., *Oxford University Press*, pp. 26–28, 408–410, 420–422.
- Creswell, J. W., & Poth, C. N. (2018). Qualitative Inquiry and Research Design: Choosing Among Five Approaches. 4th ed., *SAGE Publications*, pp. 148–150.
- Colquitt, J. A., LePine, J. A., Piccolo, R. F., Zapata, C. P., & Rich, B. L. (2012). Explaining the justice-performance relationship: Trust as exchange deepener or trust as uncertainty reducer? *Journal of Applied Psychology*, 97, pp. 1–15.
- Colquitt, J. A., & Rodell, J. B. (2015). Measuring justice and fairness. In R. S. Cropanzano & M. L. Ambrose (Eds.), "The Oxford Handbook of Justice in the Workplace" (pp. 188). *Oxford University Press*.
- De Dreu, Carsten & West, Michael. (2001). Minority Dissent and Team Innovation: The Importance of Participation in Decision Making. *Journal of applied psychology*, Vol.86 (6), p.1191-1201
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350–383.
- Emmel, N. (2013). Sampling and Choosing Cases in Qualitative Research: A Realist Approach. *SAGE Publications*, pp. 33–38.
- Feinberg, J. (1974). "Noncomparative justice". *The Philosophical Review*, 83, 297–338
- Folger, R., & Greenberg, J. (1985). Procedural justice: An interpretive analysis of personnel systems. In K. Rowland & G. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 3, pp. 141–183). Greenwich, CT: JAI Press.
- Garrad, L. K., and Hyland, P. K. (2020). "Employee survey research". *Employee Surveys and Sensing: Challenges and Opportunities*, Oxford University Press, pp. 374–390.
- Gong, Z. (2023). Artificial intelligence in the 21st century. *Research on Intelligent Manufacturing and Assembly*, vol. 2, no. 1, pp. 52–59.
- Griffin, D. W., Gonzalez, R., Koehler, D. J., Gilovich, Thomas. (2012). Judgemental Heuristics: A historical overview. In K. J. Holyoak & R. G. Morrison (Eds.), "The Oxford Handbook of Thinking and Reasoning" (pp. 322-341). Oxford University Press.

- Gustavsson, P. "Benefits of using Winningtemp - for leaders." *Winningtemp*, 2023, <https://help.winningtemp.com/hc/en-gb/articles/5022763533330-Benefits-of-using-Winningtemp-for-leaders>. Accessed 15 October 2024.
- Guttman, A. (2024). "Digitalization in the media industry in Sweden - statistics & facts." *Statista*, Jan. 10, 2024, <https://www.statista.com/topics/6007/digitalization-in-the-media-industry-in-sweden/#topicOverview>. Accessed 29 October 2024.
- Hakonen, M., & Lipponen, J. (2008). Procedural justice and identification with virtual teams: The moderating role of face-to-face meetings and geographical dispersion. *Social Justice Research*, 21, 164–178
- Hallowell, E. M. (1999). The human moment at work. *Harvard business review*, 77(1), 58–66.
- Harter, J.K., Schmidt, F.L. and Hayes, T.L. (2002) Business-Unit-Level Relationship between Employee Satisfaction, Employee Engagement, and Business Outcomes: A Meta-Analysis. *Journal of Applied Psychology*, 87, 268-279.
- Hauenstein, N. M. A., McGonigle, T., & Flinder, S. W. (2001). A meta-analysis of the relationship between procedural justice and distributive justice: Implications for justice research. *Employee Responsibilities and Rights Journal*, 13, 39–56.
- Heed, R. "How Winningtemp's AI helps you find trends and analyse the temperature." *Winningtemp*, 2023. <https://help.winningtemp.com/hc/en-gb/articles/360018016580-How-Winningtemp-s-AI-helps-you-find-trends-and-analyse-the-temperature>. Accessed 15 October 2024.
- Kahneman, D., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.
- 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.
- Lee, Min Kyung. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*. 5.
- Lind, E. A. (2001). Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. In J. Greenberg & R. Cropanzano (Eds.), *Advances in organizational justice* (pp. 56–88). Stanford, CA: Stanford University Press.
- Mayer, R. C., Davis, J. H., Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, Vol. 20, No. 3, p. 709-734.
- Mayo, E. (1933). *The Human Problems of an Industrial Civilization*. Macmillan.
- McKinsey & Company. "The state of AI in early 2024: Gen AI adoption spikes and starts to generate value." *QuantumBlack AI by McKinsey*, May 30, 2024, <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai/>. Accessed 7 November 2024.
- Miller, K. "AI Overreliance Is a Problem. Are Explanations a Solution?" *Stanford HAI*, 13 Mar. 2023, <https://hai.stanford.edu/news/ai-overreliance-problem-are-explanations-solution>. Accessed 3 October 2024.
- Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 2015 CHI Conference on Human Factors in Computing Systems*. ACM, 1603–1612.

Næss, J. "Aftonbladet låter AI summera artiklar." *Dagens Media*, May 26, 2023, <https://www.dagensmedia.se/tidningar-och-tidskrifter/kvallspress/aftonbladet-later-ai-summera-artiklar/>. Accessed 3 October 2024.

Nordicom. "Two new factsheets: The 10 largest newspaper companies and the 25 largest media companies in the Nordics, 2022." 26 Jan. 2024, <https://www.nordicom.gu.se/en/latest/news/two-new-factsheets-10-largest-newspaper-companies-and-25-largest-media-companies>. Accessed 3 October 2024.

Opdenakker, R. (2006). Advantages and Disadvantages of Four Interview Techniques in Qualitative Research. *Qualitative Social Research*, vol. 7, no. 4.

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2020). Research Methods for Business Students. 8th ed., *Pearson*, pp. 149-187.

Schminke, M., Ambrose, M. L., & Cropanzano, R. S. (2000). The effect of organizational structure on perceptions of procedural fairness. *Journal of Applied Psychology*, 85(2), 294–304.

Schein, V. E. (1977). Individual privacy and personnel psychology: The need for a broader perspective. *Journal of Social Issues*, 33, 154–167.

Tyler, T. R., & Blader, S. (2000). Cooperation in groups: Procedural justice, social identity, and behavioral engagement. Philadelphia, PA.: Psychology Press.

Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 115–191). New York, NY: Academic Press.

Van den Bos, K., Lind, E. A., & Wilke, H. A. (2001). The psychology of procedural and distributive justice viewed from the perspective of fairness heuristic theory. *Justice in the Workplace: From Theory to Practice*, 2, 49–66.

Wang, D. (2024). Proceedings of the 2024 10th International Conference on Humanities and Social Science Research (ICHSSR 2024). *Atlantis Press*, pp. 615–620.

Welbourne, T. M. (2016) The Potential of Pulse Surveys: Transforming Surveys into Leadership Tools. *Employment Relations Today*, 43(1), 33-39.

Winningtemp. "Medarbetarundersökning som gör skillnad." n.d, <https://www.winningtemp.com/sv/medarbetarundersokningar#:~:text=V%C3%A5r%20medarbetarunders%C3%B6kning%20bygger%20p%C3%A5%20%C3%B6ver, strategiska%20beslut%20baserat%20p%C3%A5%20realtidsdat>. Accessed 15 October 2024.

10 Appendix

10.1 Appendix 1: Interview guide

Translated from Swedish to English

Background Information

- Could you describe your role within the company?
- How long have you been working at the company?
- What is your educational background and work experience related to your current role?
- Have you participated in both the previous and the current employee survey systems?
- How would you describe your relationship to your manager? How often do you communicate with them about work-related feedback?

Awareness of AI in Employee Surveys

- How familiar are you with the current employee survey system? How would you describe it?
- Are you aware that the employee surveys use AI? If so, how did you learn about it?
- What were your initial thoughts when you found out AI was involved in the survey process? Were you surprised, concerned or indifferent?

Fairness and Trust in Decision-making

- Do you feel that AI-based surveys allow you to express everything important to you? Is there any feedback you feel is missed or overlooked?
- Can you give an example of a decision that was made based on survey feedback? Did the decision reflect something you mentioned in your feedback, or did it seem disconnected from what you felt was important? How did that decision affect you or your team?
- How are decisions based on survey feedback communicated to you? Is the communication clear and helps you understand why certain decisions are made?
- Do you trust that the decisions made based on the AI system's feedback are fair? Why or why not?
- Are there more or fewer in-person meetings with managers about survey feedback now? How has this affected your understanding of how decisions are made? Has it impacted relationships between managers and employees?

- Do you know what happens after you submit your feedback?
- Have there been situations where you felt the decisions based on survey feedback were unfair? Can you share an example?
- In your opinion, does the AI system contribute to more or less fairness compared to human-driven decision-making processes?

Anonymity and Frequency

- Do you feel that your responses in the survey system (e.g., Winningtemp) are completely anonymous, or are there situations where you're concerned that your manager may know who submitted the feedback?
- How does the potential lack of anonymity affect your willingness to provide honest feedback?
- How do you feel about the frequency of the surveys? Do you think responding frequently is an advantage or a disadvantage?

Perceptions of the Transition to AI-based surveys

- Since the introduction of AI-based surveys, have you noticed any changes in the decision-making process? For example, do decisions now feel different compared to when feedback was managed by human managers?
- Have you noticed any differences in the types of questions asked in the current surveys compared to the previous ones? Do the surveys focus on different issues now?
- Have you noticed the surveys asking more questions based on your answers to previous ones? If so, did it feel like the survey was adapting to your input in a thoughtful way, or did it feel more automated? Was it helpful or did you experience them irrelevant?
- Do you think attitudes toward AI-based surveys differ between your profession and other professions? If so, why?
- Looking back, how do you think the decision-making process has changed, positively or negatively, since the transition from traditional surveys to AI driven surveys? Do

you feel the change has improved the system, or are there challenges that have emerged?

10.2 Appendix 2: Email to Respondents

Translated from Swedish to English

Dear [Recipient],

We are final-year students at the Stockholm School of Economics, currently writing our bachelor's thesis in management. Our research focuses on exploring modern feedback and survey systems.

We were informed that you might be interested in participating in an interview for our study. The interview will take approximately 30 minutes and will be conducted over Google Meet at a time that suits you best.

Please let us know your availability in the coming weeks, and we will do our best to accommodate your schedule.

Thank you in advance for considering our request. Your insights would be invaluable to our research, and we look forward to hearing from you.

Kind regards,
Liv and Gustav

10.3 Appendix 3: Interview Sample

Alias (pseudonym)	Date	Length
-------------------	------	--------

Tina	2024-10-05	24:53
Anna	2024-10-06	44:57
Patrik	2024-10-07	23:31
Johan	2024-10-09	32:18
Clara	2024-10-10	22:49
Erik	2024-10-10	34:15
Lara	2024-10-11	31:33
Karl	2024-10-24	23:20
Albin	2024-10-24	52:24
Carina	2024-11-14	20:04