Recursive Impacts of Algorithmic Management on Trust and Employee Productivity in Professional Work Settings

Alexander Korogodsky
University of Miami
alex.korogodsky@miami.edu

Abstract

This research proposes a framework that describes the influences of algorithmic management on trust and worker productivity. The framework illuminates the presence of a powerful, recursive force that exhibits both enabling and inhibiting effects of algorithms on trust and explains the consequences for employee productivity. Our findings result from an in-depth single case study investigating the dynamics of adopting algorithms to manage employees in professional work settings. Drawing on advances in Actor-Network Theory, we observed that algorithms mature from being enabling tools to emerge as equal actors that both influence and are influenced by the trust relationships. This research advances algorithmic management and trust theories by describing this phenomenon in terms of unexpected effects on employee-to-employee and employee-to-algorithm trust and identifying its impact on worker productivity.

Keywords: Algorithmic Management, Trust, Employee Productivity, Actor-Network Theory

1. Introduction

Recent advances in data science and artificial intelligence facilitate the emergence of Algorithmic Management (AM) as an area of increased scholarly interest. AM is defined as the automation of management and administration of employees by delegating managerial functions to algorithms (Jarrahi et al., 2023).

With technological innovation, algorithms grow in sophistication, gain complexity and precision, and become a part of our everyday lives, especially at work (Jain et al., 2024). This improved maturity of algorithms changes their role from just *technology tools* that humans use to *participants–actors* equal to humans (Latour, 1987), introducing new recursive dynamics of human-algorithm relationships that ultimately affect workers' performance. Some of these influences are positive and enabling, while others demonstrate adverse effects on worker productivity.

While most AM literature focuses on the platform companies' space (Möhlmann et al., 2021; Waldkirch et al., 2021; Gregory et al., 2021), little has been studied in professional work settings. Kellogg et al. (2020) describe the use of AM in the context of human resources for the purposes of managing, controlling, and influencing human behavior. They offer a classification of organizations' use of algorithms to *direct* workers by restricting and recommending, *evaluate* workers by recording and rating, and *discipline* workers by rewarding or replacing them. The researchers also describe the employees' response to becoming vulnerable to algorithmic control.

Rousseau et al. (1998) define a psychological intention to accept vulnerability from those who mean well as "trust" – an excellent way to study the relationships between humans and algorithms for three reasons. First, trust is a ubiquitous phenomenon that exists in multiple contexts with individual actors (for example, between one employee and another employee or a supervisor) and with a group of actors (for example, between an employee and a team of employees). Second, it enables us to trace the effects of AM on specific trust dimensions such as ability, benevolence, and integrity (Mayer et al., 1995), connecting a highly theoretical concept of trust with a concrete organizational case. Third, multiple studies (Likert, 1967; McGregor, 1967) show that trust affects individual and organizational performance, and, therefore, two research questions can be asked and can be answered: *How does algorithmic management affect trust in professional organizations, and what are the consequential impacts of these influences on employee productivity?*

To answer these questions, we adopt the theoretical foundations of the Actor-Network Theory to understand and describe the new dynamics between

URI: https://hdl.handle.net/10125/109499
978-0-9981331-8-8

algorithms and trust with the intent of contributing to relevant literature on algorithmic management and trust theories.

We explore the case of FinPro, a prominent financial services firm with a rich history of outstanding market performance that is recognized for its unique organizational ethos and algorithmic management of employees. We demonstrate that at FinPro, sophisticated algorithms claim their own agency by polymathically aggregating massive amounts of information about the employees, serving as a decision nexus, and establishing influence ascendancy by directly affecting workers' behaviors.

Along with planned enabling influences, these dynamics are responsible for the unanticipated recursive effects of algorithms on trust, leading to credibility imbalance, ratings inaccuracy, and negativity dominance. Counterintuitively, such recursiveness adversely affects worker productivity, directly and indirectly, further eroding trust.

This paper is structured as follows. We begin by reviewing relevant academic literature and creating a strong theoretical foundation for our study. We then discuss our research design, provide a brief case description, explore the methodological foundations of the study, and review the data-gathering process and analysis techniques. We present and discuss our research findings in the context of extending existing knowledge, develop a framework and, finally, describe our contribution to empirical science and practice.

2. Theoretical foundations

2.1. Algorithmic management

Algorithmic Management has recently emerged as an area of scholarly interest focusing on controlling functions of algorithms. Duggan et al. (2020) describe algorithmic management as a "system of control where [...] algorithms are given the responsibility for making and executing decisions affecting labor, thereby limiting human involvement and oversight of the labor process". Lee et al. (2015) describe algorithmic management as software algorithms that assume managerial functions and explore how Uber and Lyft drivers respond to algorithms that assign work and evaluate their performance. Möhlman et al. (2021) also focus on using algorithms for matching and control in the online labor platform settings.

In the context of organizational work, Kellogg et al. (2020) describe how algorithms are used to direct workers by restricting and recommending, to evaluate them by recording and rating, and to discipline them by replacing and rewarding. Although their work does not explicitly study the effect of algorithms on the three dimensions of trust, they find that algorithms could be used to replace skilled workers, thus affecting the ability dimension; algorithms may lead to the creation of a dehumanized workplace and impact benevolence; algorithms could be subject to inaccuracies or bias and impact integrity dimension.

Leicht-Deobald et al. (2022) exemplify the use of AM in the human resources domain with *descriptive* performance management, *predictive* recruitment, and *prescriptive* performance improvement algorithms – de facto leading to algorithmic management of workers. Their research investigates the effects of algorithmic management and discovered that while the use of these tools can manifest an illusion of control as "algorithm-based decision-making lacks moral imagination" and, consequently, can harm employees. Consequently, these technologies threaten to erode employee trust (Leavitt et al., 2024.)

2.2. Trust and performance

While there are many conceptualizations of trust (Burke et al., 2007), a classical definition by Mayer et al. (1995) suggests that trust is the willingness to be vulnerable to another party and the extent to which the trustor is willing to take risks at the hands of the trustee.

Trust is multi-dimensional: Mayer et al. (1995) discuss three key dimensions and posit that ability, benevolence, and integrity form the core concept of trust. *Ability* is defined as a set of skills and competencies that warrant influence on others within a specific domain (Mayer et al., 1995). Naturally, higher levels of competence build confidence in stakeholders who tend to trust experts. *Benevolence* is a "perception of a positive orientation of a trustee toward the trustor" (Mayer et al., 1995, p. 719) or an extent to which a trustee wants to do good for the trustor, thus facilitating the forming of a strong relationship between them. Palanski and Yammarino (2007) examine the role of *integrity* ("fairness and loyalty" according to Butler, 1991) and discuss the importance of a set of ethical behavior principles and the consistency of their application in the formation of trust. Such a holistic understanding of trust across all three dimensions is essential, as a deficit in one dimension becomes a "weak link," undermining trust altogether.

Trust is a foundational measure of performance, many scholars agree (Likert, 1967; McGregor, 1967; Mellinger, 1956; Read, 1962). The main idea of these studies is that trust enhances workplace behavior and improves performance. This can happen *directly* – the more people trust each other, the more likely they are to exchange information and collaborate, or *indirectly* – by creating favorable conditions that facilitate performance improvement (Mayer et al., 1995). Dirks and Ferrin (2002) discovered that higher levels of trust result in more positive employee attitudes, higher levels of cooperation between them, and, therefore, improved performance. Benton et al. (1969) argue that trust does not have a primary effect on team processes and performance but indirectly facilitates them.

Lumineau et al. (2023) expose the dependency between levels of trust and control mechanisms in an organization and performance. They further propose the integrative control-trust dynamics framework, which involves a broad range of actors such as individuals, organizations, and institutions. Our work argues that there is a new actor on stage: the algorithm, which not only *affects* the trustor-trustee relationships but *becomes an actor* itself.

2.3. Actor-Network Theory

We adopt ANT as an alternative to studying algorithms in organizations from "the tool and medium perspective" (Anthony et al., 2023, p.1673). We use it as a lens for the evolution of algorithms' maturity from passive enabling tools to an agency that is no longer waiting to be used but assuming a larger responsibility (Baird & Maruping, 2021). The agency approach examines interdependencies and relevant complexities of human/machine relationships when both are actors.

Bruno Latour challenges the excessive focus of researchers studying how people interact with technology on human factors and their intentions. He suggests that human and non-human participants, including algorithms, should be treated as equal *actors* (Latour, 1987), thus significantly expanding the range of influences that need to be studied and understood. We believe this approach is important for two reasons. First, such portrayal of employees as actors is more than a semantical distinction: actorhood is not and cannot be prescribed; it emerges through the experience of institutional belonging and their sense of selves (Patriota, 2020). Second, an algorithm acting as a replacement or augmenter of a human instead of an interactant (Pakarinen & Huising, 2023) creates a new dimension for studying the effects of AM on workers – a "third pole" in the interaction of humans and non-human actors (Gutiérrez, 2023).

In fact, we were intrigued by Gutiérrez's (2023) belief that the intersection of ANT and the algorithm provides insights into the unacknowledged power relationships that were not yet uncovered and developed natural researchers' curiosity around what those power relationships might be. Hence, we aim to advance existing research by studying algorithmic management effects on trust dimensions. Following Jarrahi, Möhlmann, and Lee (2023), we focus on the dynamics between human agency and machine agency – algorithms – in professional work settings.

3. Methodology

The case of FinPro offers an insightful platform for our study. Following Gerring (2006), we selected this extreme case as it allowed us to delve into the prototypical and paradigmatic phenomenon of a firm that enabled its unique culture with algorithms at an unprecedented scale. The founder and CEO took FinPro from a 200-person small business to one of the largest investment management firms with $180b+ in assets and nearly 2,000 on board – within a decade. The unique corporate culture - so-called "Organizational Ethos" (OE) - is based on the premise of extremely open, honest communication and continuous feedback and is enabled by algorithmic solutions for candidate recruiting, employee performance evaluation, job mobility, and work assignment, as well as compensation and termination.

Some of the most notable tools used were the Feedback Collector, the Problem Log, and Worker Cards. The main idea behind leveraging algorithms in the human resources domain was that to achieve maximum performance, the firm needs to operate a "machine" built as a combination of a process and a human, both systematized and codified. This use of algorithmic management in day-to-day management is based on a distinct corporate culture rooted in the vision and mindset of the founders, who led the firm for multiple decades before stepping aside.

Relying on Gioia and Pitre (1990), we leveraged grounded theory and an inductive thematic analysis to inform theory development. Thus, we focused on engaging with FinPro's employees to learn about their specific experiences and reflections on their interactions with algorithms. From October 2023 through April 2024, we conducted 35 semi-structured interviews with former employees, among them 14 with employees in the core business and supporting functions, 11 with senior leaders, seven with "the culture carriers" (a particular group of people that oversaw the adoption of OE and supporting algorithms), and three with designers of algorithmic management solutions. Four additional interviews were conducted with senior management consultants.

Participant selection was continuously refined using theoretical sampling, whereas, after several interviews and preliminary coding, we finetuned interview questions and checked for gaps and saturation of categories (Charmaz, 2014) to ensure we did not cut off data sampling too early. We interviewed employees from several months to decades in tenure at different levels – from entry-level to executives and across several functional domains – from investments to administration and IT. The key for us was to have elaborate and insightful discussions yet not limit our participants with a pre-arranged canvas of questions.

We asked, for example, "What was algorithmic management to you at FinPro?" or "Please describe the meaning of the word 'trust' while you were at FinPro?". To enhance our study's credibility, validity, and reliability, we triangulated data collection by documenting over 275 social media posts by FinPro employees and studying 46 additional publicly available artifacts – cases, books, videos, presentations, and articles. We wanted to ensure that the discussion themes we pursued and the interview format we designed did not introduce bias to our findings.

We operated a team of three researchers – the principal investigator, the validation investigator, and the field investigator – to collect and analyze data and interpret results. After the interviews, researchers individually studied the transcripts to capture both the textual data ("*what* was said") as well as the emotional background ("*how* was it said"). They coded transcripts with the labeling of data closely following the participant's own language. Individually, in duos, and as a team, we revisited and reflected on data repeatedly to have the themes emerge directly from the participants' experience and with their lens in mind.

The first-order codes were abstracted into the second-order themes to begin forming more general theoretical constructs. For example, certain codes mapped into a common theme of "lack of data quality and reliability," others – into the "algorithmic rules definition inconsistency," and so on. The field and validation investigators iterated by comparing and refining these themes to capture the underlying insights into the interview data and held bi-weekly discussions with the principal investigator.

We then organized themes into aggregated dimensions that provided insights into the algorithm's assuming actorship and consequential multidimensional effects on trust. This abstraction was informed both by the data itself and existing theoretical frameworks to ensure we joined relevant scholarly conversations and aligned with established research. Aggregated dimensions also served as key tenets for the theoretical framework we constructed and the conceptualization of the overall consequences for employee productivity and efficiency.

4. Research findings

4.1 Algorithm as an actor

FinPro used three main algorithmic solutions to improve performance evaluation efficiency, transparency, and objectivity and automate and facilitate it: the Feedback Collector, the Worker Card, and the Problem Log. Together, these tools assumed agency that manifested itself through polymathy, decision nexus, and influence ascendancy, which we describe in Figure 1.

4.1.1 Algorithmic polymathy. The employee profile *"was the observational data that was gathered through the Feedback Collector,"* one of our participants explained. The Feedback Collector facilitated workers rating each other's strengths and weaknesses on the attributes that are aligned with OE. Another participant explained: *"Feedback collector was a data collection tool, but there were analytic engines that were consuming that information and then sending out signals to managers to say based on these Feedbacks, you should intervene or you should promote someone or you should fire someone" [P3].* These ratings – "Feedbacks" – would be assigned during the meetings - "at will" or at a specially designated time, and after an employee-to-employee interaction. Usage statistics were monitored with the quotas (out of 18 weekly Feedbacks, six must be negative as a feedback giver cannot be too friendly or too soft in the assessments) and the supervisors strictly enforced compliance.

Afterward, it would be aggregated into a Worker Card, which was used as the radar: *"Before we meet, let me look at the Worker card so we can understand each other better. It gave you a good radar of what this person appears to be good at. This is what this person appears to be bad at. I can see what feedback that person gives. Is this someone who typically gives hard and harsh feedback, and I should have that expectation? Is this someone who doesn't give harsh feedback?" [P8]*

The algorithms were also aware of employees' daily transactional activities via the Problem Log, a solution where workers could record any issues requiring further diagnostics and resolution. An issue would be recorded and assigned to a problem solver for follow-up and resolution. Often, resolving a problem would require employees to go through "diagnosis" – a well-defined and prescribed way to run these sessions, which, in turn, would result in additional "Feedbacks." This tool would often be used to push people outside of their zone, as one of the participants explained: *"I would Problem Log more than I can count, and it was comfortable… stretching people and pushing them beyond their comfort zone and knowing where their weaknesses are. With FinPro being "a safe place to fail," there were opportunities where you would make the decision whether or not you put someone in a position to fail for the sake of stretching and growing and whether or not the outcomes of that would be very impactful for the firm." [P2].*

Our results clearly demonstrate that at FinPro, algorithms were polymathic: they were aware of worker personality and capability profiles, they knew about employees' strengths and weaknesses, received weekly updates on employee performance in core OE dimensions, and they accumulated a wealth of situational and transactional performance data.

4.1.2. Decision nexus. Algorithms at FinPro participated in key HR processes. Our participant explained that there were *"Three applications: hiring – how can you use a profile to look for and screen for people like [Person A] to feed a recruiting demand signal. Plus, the interview scoring by asking similar questions was to figure out if this candidate is like [Person A] or not. Idea of matching people to either jobs or teams - based on this brighter picture of what [Person A] is like, what is he likely to be good or bad at. Combination of that - who is he likely to work well with." [P11].* Aside from the recruiting and job assignment, algorithms were active participants in handling promotions and compensation decisions and were *"more used for management types of decisions: who should we put on this project? Is this person performing well over time? If not, what can we do about it? If they are - should we promote them and give them more responsibility? What should their compensation level be?" [P19]*. Algorithms even assisted in releasing people: *"When we let people go, they were never combative because we would be able to go through the Feedbacks. Here's a picture of you. Do you agree? Do you not agree? Most people's reaction was, "Wow, I should not be in this job; I should move on." [P8].* Therefore, the algorithms were at the crux of all employee-related decision-making across a broad spectrum of people management processes.

4.1.3. Influence ascendancy. Finally, our results demonstrated that the wealth of knowledge of the workers and engagement in all key decisions led to algorithms affecting the opinions of humans: *"My only worry about it was I sometimes felt like people gave feedback that reaffirmed what the baseball card said versus truly being independent" [P8].* Also, we observed that the algorithms even influenced employees' behavior: *"I can say that data from the tools probably had a subconscious influence on how I interacted with people and how I went about solving problems." [P23].*

In summary, we conclude that at FinPro, algorithms were polymathic - they accumulated a great deal of knowledge about the employees, knew who they were and how they performed, and stored daily transactional details of that performance. In essence, algorithms became a decision nexus by playing a central role in many employee management processes and participating in all critical decisions from hiring to firing. Algorithms also exhibited influence ascendancy in that they could influence workers' behaviors, decisions, and actions. Hence, we confirmed that algorithms became actors at FinPro.

4.2 The effects of actor-algorithm on trust

When the algorithm emerges as an actor, it continues to deliver expected enabling effects - the same way it did when it was just an enabling tool. However, what is nuanced at FinPro is that the word "trust" was not part of the organizational vocabulary. A participant shared with us *"an OE principle regarding trust: 'trust but verify.' You're not supposed to trust anyone or anything just because there's a pattern of being able to trust that person." [P21].*

However, FinPro created a proxy for trust – so-called "credibility." Credibility was initially defined as *"score/weighting calculated more intuitively based on observations, experience" [P4]* and later became a function of Feedbacks collected from people. The same participant further explains: *"It's through not only providing Feedback both on content and on OE that one can earn a higher credibility ranking." [P4].* We described these effects in Figure 2.

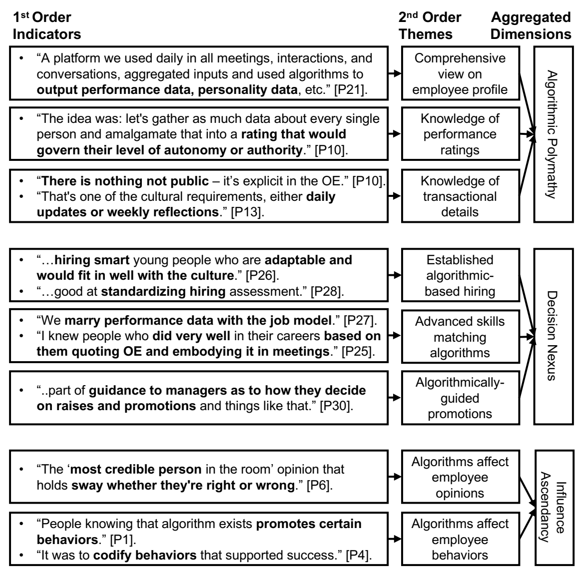

Figure 1. Algorithm as an actor

4.2.1. Credibility imbalance. However, the notion of credibility was an imbalanced metric where new employees, tenured employees, and the founders would have different levels of credibility. For new employees, even experienced ones, *"any experience or expertise external to FinPro basically didn't count for anything. You could have been the world expert in topic X when you joined the firm, and you were not going to be credible in X." [P13].* Another participant adds to those insights, highlighting that the credibility score was entirely different for the tenured employees: *"So you're either super tenured or you're not, that affected the culture - if you're there for 12 years, you're credible in 20 things. If less than two years - you're not." [P18].* With that, the founders' credibility was a golden standard, and everybody else's score was measured relative to that. This was not just an anecdotal point or an assumption; one of our participants explained that *"the founders were 10 out of 10 in every category – it's not an assumption; it's mathematically in the system. It's one of the drivers that has a hard-coded axiom." [P11].*

This credibility imbalance between groups of employees was not coincidental and could not be "de-biased" – it was part of the algorithm.

4.2.2. Ratings inaccuracy. At the same time, the ratings were not perceived by employees as very accurate. *"The data was messy." [P17]* First, the algorithms were prone to systemic data quality issues: errors, data incompleteness and inconsistency, and even initially ambiguous definitions. For example, *"Over time, the definition of that Feedback could change. So even in a time series of the data, the definition of it changed, so the data aren't as reliable: it means different things over time. [P12].* Second, the inconsistency of rules and definitions also affected the logic of the algorithms, which consisted of *"a bunch of subjective rules, expert rules, heuristics that were codified by people who are "credible" [P3].* On top of that, the ability to assign a metric to a poorly quantifiable definition would also contribute to the inaccuracy of evaluations: *"It's like you are measuring these fuzzy things, and trying to correlate them to other fuzzy things, and then trying to figure out how to drive action right with them." [P31].* Third, an algorithm's output was coarse and further subjected to interpretation. Participant 3 explained: *"The signals were two coarse: the engine would say promote someone or fire someone. The engine was not nuanced enough; it was too binary." [P3].* Another participant added: *"Behaviors are quantified. And I don't know what you do with that information: exposing it without explanation created some very strange incentives and behaviors." [P11]*

Evidently, a combined effect of data quality issues, rules inconsistency, and subjective output interpretation caused rating distortion and, consequently, resulted in employees questioning the reliability of an algorithm: *"We generally did not trust the data and, as a result, the outcomes, and the engines." [P2].*

4.2.3. Feedback manipulation. At FinPro, all algorithms and their usage were focused on negativity. For example, over 1/3 of weekly quotas in Feedback Collector were mandated to be negative because coming across as overly positive was against the OE. The Problem Log, by definition, was designed to collect and report the issues, though no records of positive actions were kept. One of the participants recalls: *"Friday afternoon comes up, and you are scrambling to think, 'Oh, I forgot to Feedback someone negatively.' And you're trying to think of anything that anyone did wrong, so that way you can leave negative Feedback before you bounce on for the weekend." [P7].*

Focus on negative feedback engraved in the culture and solidified by the quotas is lethal to the employees: *"I started as someone who was a highly capable individual that trained really fast and was above expectations within the first six months to a total failure." [P7].* The reason for such a quick transition "from hero to zero" was because of the enforced negative feedback.

In response, some employees "voted with their feet" – in our interview sequence, 15% of the participants had a tenure of 1 year or less (three and six months). Others reflected that *"I always thought it was a shame that we only logged bad things and not good things because I think they're equally important." [P13]*

Those who stayed, to avoid negative dominance, needed to play along, which, in many cases, rewarded not the real skills or actual performance but perception of performance – what our participant called *"somewhat arbitrary, politically influenced perception of me." [P13].*

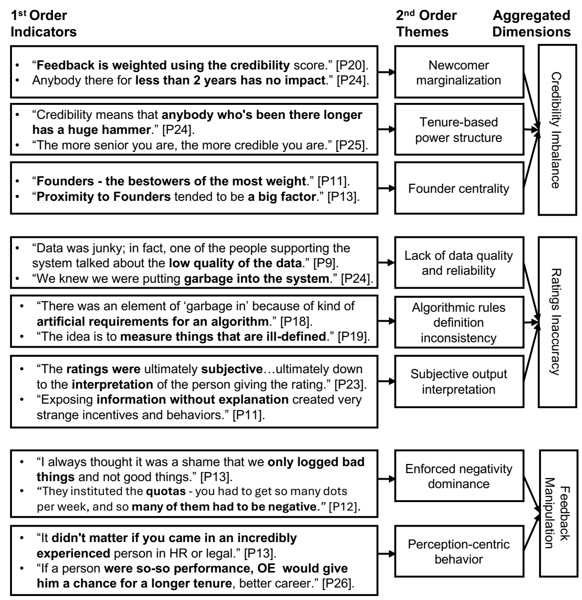

Figure 2. Effects on trust

This happened in group settings: *"People spent weeks to pregame meetings. By the time the team shows up [to the meeting], there's nothing organic about the conversation they rehearsed" [P16].* This also happened at the individual level: *"I am going to do this meeting because I know that there is a person there who has been assessed to be strong in this attribute because they are in proximity to a very important person who also is good at that attribute. So, if I can say something smart in front of that person and I get him or her to give me that Feedback, I just boost my score in that dimension radically up." [P3]*

For many, though, *"after a while - usually 3-4 months - you felt like most people felt like their actions had no impact on the result. I can do a terrible job and be totally distracted and get glowing feedback. Or I can be doing what I think is a really good job, or at least trying really hard, to get negative feedback." [P24].* Consequently, employees were less interested in changing their actual behavior or improving real performance. Top of mind was creating the right perception of themselves in the system. Hence, the enforced negativity dominance coupled with skillful perception-centric behavior created fertile grounds for feedback manipulation.

In essence, three mostly unexpected influences on trust are present: credibility imbalance, ratings inaccuracy, and feedback manipulation.

4.3. Employee productivity

Now that we have seen in the data that algorithms become full actors at FinPro and that "actorship" causes some unanticipated effects on trust, we were intrigued to trace the direct and indirect consequences on employee productivity.

4.3.1. Direct effect on productivity. Direct effect results from the organization explicitly requiring employees to invest time using the algorithms. *"We called it a social contract within FinPro that we all had with one another to adhere to that, incorporating the OE into our daily life. And on average, we spend about 20% of our week Feedbacking one another." [P21].* However, another participant explained that not only did workers spend time using the systems, but they also needed to do extensive pre-planning on how to use them to do the right thing – properly reflecting on others' perceptions of them. Another participant reports: *"Full 20% of the time is by definition set up for you to do that kind of stuff: not work, but Feedback Collector, Problem Log, and so on... And then you will be sitting there writing and rewriting and rethinking how you should phrase this to tr other things. It's like everything becomes a double, at least right from the remaining 80%. And then you're down from your 100% productivity down to maybe 30-40% if you're lucky." [P32].* This math is further confirmed: *"50% of my time is now in conversations discussing the problems that I'm having in my role that's causing more problems because now I don't have enough time to do the job, creating the compounding effect." [P7].*

4.3.2. Indirect effect on productivity. The indirect effects are much more complex and more difficult to account for. For example, to maintain such a rigorous focus on OE and the algorithms that supported and enabled it, the organization deployed so-called "management workers" often perceived as *"incredibly heavy weight of overhead" [P13]*. A participant explains: *"There are people who do the work, there are people who sit in the Management Committee, and then there are people, the managers, in the middle." [P3].* Another participant estimated how many of those workers FinPro employed: *"There were 1,000 people, basically management jobs, that weren't directly attributable to the performance of FinPro." [P11].*

Another example of an indirect effect would be employee productivity in the meetings – starting with the Management Committee down. *"Even the management committee spends most of its time talking about whether they themselves are operating according to the OE and what bad outcomes they need to deal with. They are not talking about the business of the company." [P10]*

Evidently, employee productivity was directly and indirectly impacted by algorithmic management's recursive impact on trust. Based on the participant feedback, we estimate that employees spent at least one full business day managing the algorithms every week and at least another business day - planning and pregaming meetings. This effectively leaves three days out of five to attend work duty, which justifies a description of FinPro given by an interviewee: "*FinPro was not at any point a well-run company; it was extremely inefficient." [P10].*

5. Discussion

When algorithms "grow out" of being enabling tools and become actors, along with human and non-human objects, a new trust dynamic is introduced: employees (individually, as humans, and collectively, as teams, sub-cultures, and organizations) need to trust the algorithm. This new dynamic results from a change in trustees and perception of control. Previously, when the algorithm was used as an enabling tool, workers perceived that they were in control. They trusted their intentions rather than the algorithm that facilitated benevolence, ensuring tasks were accomplished more efficiently. However, when an algorithm becomes an equal actor, it is now making decisions that affect human relationships, reduce the autonomy of human actions, and force evaluations and task assignments. If the algorithm demonstrates maturity in producing reliable and consistent outcomes – it will enhance productivity. Otherwise, if its ability is questioned due to bias, data quality concerns, rules inconsistency, or other problems, the use of the algorithms gradually subsides until it finally vanishes.

5.1 Multi-dimensional recursive effects

When an algorithm becomes an actor, it introduces a complex effect on trust. However, instead of just describing it, we add depth by demonstrating influences *across all three dimensions of trust* – ability, benevolence, and integrity. We also suggest that the recursive nature of algorithms' effect on trust can be understood through *primary effects* and *counter-effects,* both displayed in Figure 3 as solid arrows and dotted arrows, respectively.

In the ability dimension, algorithmic tools enhance employee skills and allow them to make more informed decisions quickly, automate routine tasks, free up employees to focus on higher-value activities, and assure consistency in which these skills enable employee competencies. As a result, this can significantly improve both employee-to-employee trust and employee productivity and consequently strengthen employee-to-algorithm trust, creating a primary effect with a positive feedback cycle. The counter-effect here is due to ratings inaccuracy fueled by questionable data quality and ambiguously defined algorithmic rules coupled with subjective interpretation of algorithmic output that creates a distorted picture of skills and competencies one can trust. The distortion causes a misleading picture of one's true abilities, triggering misinformed decisions about performance, development needs, and career mobility. The result is that employee-to-employee trust is eroded; employees perceive algorithms as unreliable, trust them less, and resist their adoption, and the negative feedback cycle continues, impeding trust.

In the benevolence dimension, the openness and loyalty of the trustee signal their "goodwill" and reinforce the expectation that the trustee will act in favor of the trustor's interests. Hence, the algorithms' primary effects are to promote transparency and facilitate open communication channels through which precise data-driven performance assessments are made available. This positively affects both employee-to-employee and employee-to-algorithm trust because transparency reinforces employees' belief that the organization – supervisors and colleagues – is committed to developing them, further enhancing the trust in the system and algorithms supporting it. However, we see how algorithms true up performance assessments with credibility, which marginalizes new employees regardless of their experience level. Tenured workers are empowered with higher credibility ratings, making the founders, as well as senior leadership in their direct proximity, the most credible. Credibility imbalance becomes a counter-effect that impedes trust and results in workers' skepticism, adding to their reluctance to rely on algorithmic systems. This perpetuates a cycle of mistrust and non-usage or complacent usage at best.

Finally, in the integrity dimension that is concerned with perceived adherence to ethical values such as fairness and honesty, the primary effects enhance the trust between employees and between the employees and the algorithm. Naturally, fair feedback and straightforward performance assessments performed by an algorithm reduce uncertainty among employees and enable a culture of accountability, making sure decisions and follow-up actions are recorded, analyzed, and improved. It contributes to more frequent use of algorithms and generates more data and feedback, which then further improves the algorithm's accuracy and reliability. This positive feedback loop reinforces trust over time. However, excessive negative dominance coupled with perception-centric behavior causes an unexpected counter-effect where employees refrain from using algorithms in good faith, and feedback manipulation is introduced. Employees *a priori* are skeptical about the feedback they receive and prefer to invest time in manipulating the feedback instead of improving their actual skills to earn favorable ratings. In the culture of extreme transparency, these practices quickly become public, which snowballs into workers' complete mistrust of each other and the algorithms. Moreover, confronted by the usage quotas, employees complacently use algorithms too formally or, worse yet, engage in deception tactics, trying to "game the system." Both activities result in an avalanche of poor-quality data that is unfit for the purpose of algorithmically managing the employees. Algorithms become less accurate, leading to more errors and biases – thus eroding trust in the algorithm's "good intent" and creating a downward spiral.

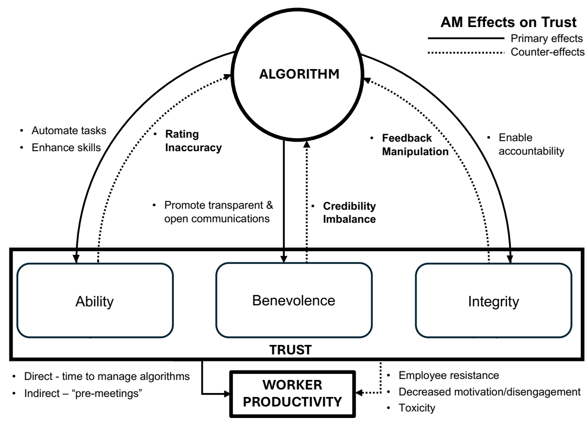

Figure 3. Multi-dimensional recursiveness

Therefore, primary effects and counter-effects in each of the trust dimensions demonstrate an aggregate impact of this multidimensional recursiveness. The strength of the combined effect on trust will depend on the strength of individual primary effects and counter-effects within each trust dimension and the nature of the feedback loops we described.

5.2 Productivity effect

Some of the described influences are direct: employees allocate time to manage algorithms and engage in feedback manipulation. Others, such as investing time in pre-gaming meetings or collaborating with "management workers" - are indirect.

Despite the anticipated improvement in employee productivity from algorithmic management, the recursive effects that impact all three dimensions of trust explain the consistent exponential degradation of employee productivity we observed in the case of FinPro. We attribute it to three primary reasons.

First, employees who distrust the quality, integrity, and relevance of algorithmic decisions resist. Such resistance manifests in non-compliance or attempts to game the system, both of which require time that would be typically taken away from handling core work activities. Our study suggests that this impact on productivity is significant and cannot be ignored.

Second, erosion of trust leads to decreased employee motivation and commitment, ultimately resulting in partial or complete disengagement. Naturally, if the algorithms prioritize metrics that are misaligned with employees' perceptions of valuable work – employees sign off. Disengaged employees are less likely to be productive and could put in less effort, take more time off, and even leave the organization, creating a high turnover effect.

Third, a distrust at work combined with algorithmic management and monitoring leads to a toxic environment - one of fear, increased stress, and anxiety, further degrading workers' productivity.

Therefore, if professional organizations are focused on improving worker productivity using algorithms, they must recognize the complexities of the multidimensional recursive effect of algorithmic management on trust.

5.3. Summary, limitations, research directions

Engaging in the algorithmic management of workers, professional organizations need to expect that algorithms grow out of being enabling tools and become full actors. This manifests through attributes of an agency: algorithmic polymathy, decision nexus, and influence ascendancy. When this occurs, along with anticipated effects on trust, unexpected influences are observed, such as credibility imbalance, ratings inaccuracy, and negativity dominance. Such complex dynamics of the algorithmic management effect on trust between actors exhibit recursiveness that strikes all three dimensions of trust: ability, benevolence, and integrity. The resulting effect on worker productivity is entirely unexpected: instead of anticipated productivity gains due to algorithmic enablement, employee productivity is adversely affected.

However, we observe boundary conditions when using a single, extreme case study for theory construction. First, an extreme case, by definition, is not representative of typical situations and, thus, has limited generalizability. Second, we are aware of causal ambiguity: it is difficult to establish whether the observed dynamics are due to the algorithmic management effects or the extreme nature of the case.

Nonetheless, this work's novelty is in discovering these new dynamics and proposing a framework that accounts for known and expected influences of algorithmic management on trust and identifies the unexpected but significant recursive effects. This is academically significant because, while relying on the established empirical theory, we leverage the "algorithmic agency" proposed by ANT and contribute to the existing scholarly conversation on the algorithmic management of workers in professional work settings. We also enhance the academic dialogue on trust by demonstrating how this recursiveness affects the key dimensions of trust.

Indeed, our findings are quite significant for practitioners: leaders who extensively rely on managing people with algorithms need to understand the nuanced relationships between employees and algorithms and how the emergence of trust between them affects employee productivity.

Our research has limitations, however. The triangulation method we used was limited to accessing publicly available artifacts, whereas it would be preferable to use site observations and focus group interviews if full site access can be obtained. Also, we applied the ANT in a focused way – just enough to demonstrate that algorithms become equal players along with other actors. Third, we focused on the three classical dimensions of trust (ability, benevolence, and integrity), although trust theory offers a broader (and deeper!) interpretation of this construct. Finally, we relied on retrospective employee recollections while lacking longitudinal data and quantitative evidence, which limited our ability to assess the evolution and longer-term impacts of algorithmic management.

Additional research would need to test our framework "at scale" by expanding to several case studies using, perhaps, the Eisenhardt methodology for qualitative research. We also stayed committed to identifying and explaining the phenomena of recursive effects, thus leaving outside of the scope a complete and exhaustive list of all known effects of algorithms on trust; instead, we focused only on those that emerged from our research. Conducting a more comprehensive, holistic study and developing a complete taxonomy would be advantageous both from an academic and practitioner's viewpoint.

6. Conclusion

When professional organizations adopt algorithmic management, algorithms may become full actors, assuming the roles of decision influencers and decision-makers. This leads to complex and multidimensional impacts on trust that could adversely affect worker productivity instead of enhancing it. Whereas the existing theories may not fully capture the nuances of these relationships, our framework identifies specific effects and explains their recursive nature to allow scholars and practitioners alike to understand these influences better and address the dynamics of professional workplace relationships. We see this as an essential prerequisite for mitigating adverse effects on worker productivity and enabling more effective integration of algorithmic management in the workplace.

7. References

Anthony, C., Bechky, B. A., & Fayard, A. L. (2023). "Collaborating" with AI: Taking a system view to explore the future of work. *Organization Science*, *34*(5), 1672-1694.

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, *45*(1).

Benton, A. A., Gelber, E. R., Kelley, H. H., & Liebling, B. A. (1969). Reactions to various degrees of deceit in a mixed-motive relationship. *Journal of Personality and Social Psychology*, *12*(2), 170.

Burke, C. S., Sims, D. E., Lazzara, E. H., & Salas, E. (2007). Trust in leadership: A multi-level review and integration. *The Leadership Quarterly*, *18*(6), 606-632.

Butler Jr, J. K. (1991). Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of Management*, *17*(3), 643-663.

Charmaz, K. (2014). *Constructing Grounded Theory.*

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review*, *46*(3), 534-551.

Dirks, K. T., & Ferrin, D. L. (2002). Trust in leadership: meta-analytic findings and implications for research and practice. *Journal of Applied Psychology*, *87*(4), 611.

Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2020). Algorithmic management and app‐work in the gig economy: A research agenda for employment relations and HRM. *Human Resource Management Journal*, *30*(1), 114-132.

Gerring, J. (2006). *Case study research: Principles and practices*. Cambridge University Press.

Gioia, D. A., & Pitre, E. (1990). Multiparadigm perspectives on theory building. *Academy of Management Review*, *15*(4), 584-602.

Gutiérrez, J. L. M. (2023). On actor-network theory and algorithms: ChatGPT and the new power relationships in the age of AI. *AI and Ethics*, 1-14.

Jain, N., Gupta, V., Temperini, V., Meissner, D., & D'angelo, E. (2024). Human machine interactions: from past to future-a systematic literature review. *Journal of Management History*, *30*(2), 263-302.

Jarrahi, M. H., Möhlmann, M., & Lee, M. K. (2023). Algorithmic management: The role of AI in managing workforces. *MIT Sloan Management Review*, *64*(3), 1-5.

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.

Latour, B. (1987). Science in action: How to follow scientists and engineers through society. *Harvard University Press. Cambridge, Mass*.

Leavitt, K., Barnes, C. M., & Shapiro, D. L. (2024). The Role of Human Managers within Algorithmic Performance Management Systems: A Process Model of Employee Trust in Managers through Reflexivity. *Academy of Management Review*, (ja), amr-2022.

Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015, April). Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 1603-1612).

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2022). The challenges of algorithm-based HR decision-making for personal integrity. In *Business and the Ethical Implications of Technology* (pp. 71-86). Springer Nature Switzerland.

Likert, R. (1967). *The human organization: its management and values.* McGraw-Hill.

Lumineau, F., Long, C., Sitkin, S. B., Argyres, N., & Markman, G. (2023). Rethinking control and trust dynamics in and between organizations. *Journal of Management Studies*, *60*(8), 1937-1961.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709-734.

McGregor, D.M. (1967). *The professional manager.* McGraw-Hill.

Mellinger, G. D. (1956). Interpersonal trust as a factor in communication. *The Journal of Abnormal and Social Psychology*, *52*(3), 304.

Möhlmann, M., Zalmanson, L., Henfridsson, O., & Gregory, R. W. (2021). Algorithmic Management of Work on Online Labor Platforms: When Matching Meets Control. *MIS Quarterly*, *45*(4).

Pakarinen, P., & Huising, R. (2023). Relational expertise: What machines can't know. *Journal of Management Studies*.

Palanski, M. E., & Yammarino, F. J. (2007). Integrity and leadership: clearing the conceptual confusion. *European Management Journal*, *25*(3), 171-184.

Patriotta, G. (2020). Actors and actorhood in institutional theory. *Journal of Management Studies*, *57*(4), 867-872.

Read, W. H. (1962). Upward communication in industrial hierarchies. *Human Relations*, *15*(1), 3-15.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, *23*(3), 393-404.

Waldkirch, M., Bucher, E., Schou, P. K., & Grünwald, E. (2021). Controlled by the algorithm, coached by the crowd–how HRM activities take shape on digital work platforms in the gig economy. *The International Journal of Human Resource Management*, *32*(12), 2643-2682.