

The Birth of AI-driven Nudges

Nyman, Stig

Document Version
Final published version

Published in:
Proceedings of the 56th Hawaii International Conference on System Sciences

DOI:
[10125/103276](https://doi.org/10.125/103276)

Publication date:
2023

License
CC BY-NC-ND

Citation for published version (APA):
Nyman, S. (2023). The Birth of AI-driven Nudges. In T. X. Bui (Ed.), *Proceedings of the 56th Hawaii International Conference on System Sciences* (pp. 5252-5261). Hawaii International Conference on System Sciences (HICSS). <https://doi.org/10125/103276>

[Link to publication in CBS Research Portal](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 09. Dec. 2023



The Birth of AI-driven Nudges

Stig Nyman
 Copenhagen Business School
snc.digi@cbs.dk

Abstract

AI methods allow for a multitude of new forms of managerial control. One is algorithmic nudging, in which organizations use AI methods to control workers through targeted recommendations. Drawing upon Michel Foucault's analytical strategies, the paper examines the intellectual heritage and ideological roots of AI-nudges. Scholars have commented on the resemblance between algorithmic nudging and Taylorist scientific management. However, as this paper shows the discourse of AI-nudges also shares significant linages with other subsequent opposing managerial paradigms. Building on the analysis of AI-nudges linages, the paper discusses how their use implies three contestable presumptions 1) that work can be codified, 2) that workers require autonomy over their work, and 3) that there is no existing conflict of interest between workers and the organization.

Keywords: Nudging, Algorithmic control, Algorithmic management, Genealogy, Critical IS research

1. Introduction

Influential managerial outlets have started to promote the idea that organizations should use so-called “artificial intelligence (AI)-driven nudges” to manage their workforce. For instance, McKinsey’s website features an article describing how:

Companies are combining behavioral-science insights with the latest AI technologies to create a new kind of tailored coaching experience for their employees. [...] these new “smart AI coaching” systems combine multiple sources of data [...] to select timely, context-specific nudges, delivered via the employee’s computer or hand-held device, with intelligent, deliberate repetition.” (Amar et al., 2022)

The notion of AI-driven nudges (AI-nudges) describes a new type of human-machine configuration where software assumes the role of a human manager,

directing workers towards organizational goals. Research on algorithmic management practices has documented how software tools using machine-learning and AI methods have created the basis for the emergence of a plurality of new algorithmic control practices (for a review of this topic, see Koukouvinou & Holmström, 2022). At platforms and incumbent firms, AI methods and tools are used to monitor, sanction and reward worker’s behavior (Kellogg et al., 2020). However, as McKinsey’s article illustrates, AI methods and tools are increasingly also being used by organizations to direct their workforce in more subtle and informal ways (Burr et al., 2018), offering workers targeted recommendations on how to work and live effectively (Kellogg et al., 2020).

Information systems (IS) and organizational scholars caution that such AI-nudges can be manipulative. AI-nudges risk overriding the worker’s own ideas about their well-being and can encourage them to act in ways that are not in their own interest (Burr et al., 2018; Gal et al., 2020; Kellogg et al., 2020; Zuboff, 2019). The crowdwork platform Uber has delivered one of the most iconic cases of an ethically questionable use of AI-nudges. In 2016, Uber received critical attention because an article in the Atlantic revealed how the firm had harnessed insights from nudge theory to manipulate workers to drive longer work hours (Gino, 2017). Furthermore, scholars warn that AI-nudges promulgate a neo-liberal and exploitative work ethic (Moore, 2018; Wajcman, 2019). Moreover, scholars contest that the opacity of learning algorithms behind AI-nudges introduces issues of discrimination, mistrust, and lack of accountability and skill development (Berente et al., 2021; Faraj et al., 2018; O’neil, 2016).

Yet, AI-nudges “does not have to be unethical”, as Möhlmann (2021) contends in a Harvard Business Review article. Recognizing the challenges and pitfalls of AI-nudges, Möhlmann advocates that companies might be able to leverage these systems to create beneficial circumstances for workers and the organization. The tensions outlined by Möhlmann display how the debate on AI and algorithmic-management practices has tended to oscillate between two poles of dystopian and utopian views of technology (Koukouvinou & Holmström, 2022).

This paper brings forward an affirmative critique that seeks to provide nuances to the dystopian and utopian views of the managerial practices involving AI-nudges. The paper draws on a problematization lens (Nyman et al., 2021) to examine the intellectual heritage and ideological roots of AI-nudges. This historical perspective allows us to ask: What are the assumptions that proponents are implicitly committed to when arguing for the benefits of AI-nudges.

The paper contributes to the growing research streams on how AI-human symbiosis reconfigures work and organizational practices (Koukouvina & Holmström, 2022). The literature of algorithmic control has often commented on the resemblance between AI-nudges and concepts from Taylorist scientific management (e.g. Altenried, 2022; Jarrahi et al., 2021; Kellogg et al., 2020; Wang et al., 2020). However, as this current paper shows, managerial practices using AI-nudges also share significant ideological and intellectual lineages with other subsequent managerial paradigms, which oppose the principles of scientific management.

Informed by genealogical accounts of the evolution of managerial thought, this paper analyzes seminal texts related to the discourse of AI-nudges. Through this work, the paper traces the differences and similarities between the discourse of AI-nudges and the four managerial paradigms that have dominated managerial thinking in the twentieth century; that is, scientific management, human relations, systems rationalism, as well as organizational culture and quality (Barley and Kunda, 1992).

From this analysis of the intellectual and ideological lineages of AI-nudges, this paper articulates three assumptions that proponents of AI-nudges implicitly commit. These assumptions are formulated as questions that managers must accept as points of departure for the assessment of the meaningfulness and ethical defendability of AI-nudges. These questions can be used by researchers and managers to undertake a critical evaluation of how and when to promote or discourage the use of AI-nudges.

2. Background

Thaler and Sunstein (2009) first coined the term ‘nudge’ to describe how public and private institutions could adopt methods and theoretical insight from behavioral economics as an alternative to traditional regulatory approaches. They illustrated this alternative by detailing how even mundane choices such as the organization of a buffet in a canteen can influence what people eat. Thus, if you want people to eat healthier, you do not need to ban unhealthy food or incentivize healthy eating, you could simply start by

placing the salad first in the buffet. According to Thaler and Sunstein, the canteen case illustrates that when authorities (public or private) pay attention to how environments shape decisions and actions, they can “influence people’s behavior”, while they simultaneously “maintain or increase freedom of choice” (ibid. p. 5).

Soon after the publication of *Nudge*, the concept started to gain traction among managerial scholars and practitioners. Google was one of the early firms to explore how to adopt nudging as a managerial strategy. In his account of Google’s innovative HR practices, Google’s former head of people operations, Lazlo Bock (2015), detailed how Google used nudges to promote healthy life-style choices for their employees, and improve security protocols and onboarding processes.

Since Google’s early nudge experiments, the digital transformation of work and technological advances has widened the potential scope and efficiency of managerial nudges. Mele et al. (2021) describe how wearables, machine learning platforms and conversational agents allow for the incorporation of real-time data and dynamic personalized predictions and recommendations into the design of a nudge. In their influential review of algorithmic management practices, Kellogg et al. (2020) draw on the concept of nudging to characterize the form of managerial control exercised when organizations use AI methods to prompt the workers with targeted recommendations. Following a general tendency in the literature on algorithmic management practices and algorithmic control (e.g. Gal et al., 2020, Möhlmann, et al., 2021), Kellogg et al. (2020) use the concept of nudging to highlight the ethical and tensions concerns introduced by controlling workers with algorithmic recommendation systems.

Despite the criticisms, as McKinsey’s article indicates, the concept of nudging continues to serve as inspiration for the development of a wide range of algorithmic recommendation systems in the workplace. According to Wajcman (2019) and Zuboff (2019) nudge theory has provided the overarching ideology to promote and legitimate algorithmic recommendation systems and the novel form of control that these systems afford. The critical aim of this paper is to articulate the assumptions that are being made when using nudge theory to promote and legitimate the use of algorithmic recommendation systems to control workers.

3. Methodology

3.1 Towards an affirmative critique of AI-nudges

Critical research that challenges prevailing assumptions in the development, use and impact of digital technology has emerged as an important research agenda in the IS community (Myers & Klein, 2011; Rowe, 2018). We have learned from an extensive body of IS research that discourse plays a central role in the diffusion of emerging technologies because it provides a common sense understanding for the emerging technology's reason d'être and legitimizes the deployment the technology (Swanson and Ramiller, 1997). This paper uses the problematization lens as a critical IS research strategy to challenge the prevailing assumptions underpinning the use of AI-nudges in the workplace.

The problematization lens has been developed from Michel Foucault's analytical strategies, particularly his genealogical method (see Nyman et al., 2021 for an elaborate introduction to the problematizations lens). From the works of Foucault (1984; 1990), we learn that when we engage with technology, we implicitly commit to certain historical contingent assumptions concerning our philosophical worldviews. Consequently, when we promote and use a given technology, we inevitably commit to certain assumptions of: 1) what we can and cannot know, 2) what we believe constitutes legitimate and praiseworthy modes for governing (or managing) others, and 3) how we assume ourselves and others to be moral subjects. This paper aims to shed light on the discursive assumptions that inevitably are made when promoting and using AI-nudges.

Similar to other lines of critical IS research, the problematization lens has an emancipatory aspiration (Myers and Klein, 2011). However, its critical aim is not to challenge the status quo of technological applications and unmask its harmful consequences. Instead the critical aim is of the problematization lens affirmative. It does not aim to provide answers and solutions but rather aims to redescribe and ask new questions. Paraphrasing Foucault (2008), this paper does not aim to say, "look how oppressive AI-nudges is," but instead it aims to bring to light the conditions that had to be met for a discourse on AI-nudges to be possible. Drawing inspiration from Foucault's analytical approach, this paper articulates the conditions that one needs to commit oneself to for the use of AI-nudges to constitute a meaningful and acceptable managerial practice.

As outlined in the introduction, the debate on AI and AI-nudges have tended to oscillate between two

poles of a dystopian and utopian view of the impact of algorithmic recommending systems. It is hoped that by engaging in an affirmative form of critique, we may be able to recognize the challenges of AI-nudges while maintaining that the fact that since a discourse on AI-nudges continue to prevail suggests that the concept has its merits.

3.2 Scoping the concept of inquiry

Before, further detailing the analytical strategy of the paper, a brief remark on its empirical scope. Thaler and Sunstein's original work defines nudging broadly stating the term refer to: "*any* aspect of choice architecture that alter people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives" (2009, p. 6, emphasis added). Notably, this broad definition has allowed the concept of nudging to be applied to a wide range of context and purposes. In this paper, we are interested in nudge theory only in so far as nudge theory is being used to support the development and implementation of AI based systems in the workplace with the specific aim to steer workers behavior using target recommendations (Kellogg et al., 2020). Similar to nudging, the concept of AI is rather broad. According Berente et al. (2021): "AI is not a single, discernable thing" nor set of technologies" but a concept that is used in public and academic discourse to describe the "frontier of computational advance" (p. 1435). Consequently, we approach AI-nudges not as a single technology or set of technologies but as a discourse that is used to promote and legitimize the use AI based software, with the aspiration of steering workers behavior but without using rewards or sanctions. As such, the paper's empirical concern is the discursive logic that underpins managerial control practices that involve the use of computational techniques to prompt workers with targeted recommendations.

3.3. Strategies for training the analysis

The paper elicits the assumptions embedded in the discourse of AI-nudges through in-depth readings of seminal- and key texts related to AI-nudges (as advocated by Alvesson and Sandberg, 2011). By the nature of the analytical approach, the selection of texts is eclectic (Koopman, 2019). The texts have been selected for their ability to display the discourse of AI-nudges. They have been identified by looking at references in existing research (particularly commentaries and review articles), blog posts and Wikipedia. In addition, texts have been identified from teaching activities and informal talks with colleagues.

The body of seminal texts that are chosen for analysis includes Thaler and Sunstein's original works (2003; 2009), texts introducing nudging into the managerial and digital domain (e.g., Bock, 2015; Möhlmann, 2021), as well as literature reviews and commentaries discussing AI-nudges and its related concepts (e.g. Kellogg et al., 2020; Mele et al., 2021).

The texts have been read repeatedly. Informed by the problematization lens, the paper asks, according to the discourse on AI-nudges: 1) what kind of knowledge is required for AI-nudges to be possible? And, what kind of knowledge does AI-nudges implies a skepticism towards? 2) When is AI-nudges a praiseworthy and legitimate form of control? And, when is it not? 3) What notion of human kind underpins AI-nudges? (see Hacking, 2007, for an elaboration on the concept of human kind)?

Unraveling the assumptions that have shaped our theoretical imagination and our theoretical reasoning requires critical reflection (Rowe, 2018). Commentators such as Land (2010) and Stahl (2011) encourage us mobilize historical knowledge to cultivate our capacity for critical reflection. Taking inspiration from the historical orientation in Foucault's work, we used a historical perspective to map the assumptions of AI-nudges. Thus, our analytical work has been supported by records of the history of behavioral economics and nudge theory (e.g., Bruni & Sugden, 2009; Thaler, 2016). Moreover, as we are interested in the use of AI-nudges for managerial purposes, the analytical work has been guided by modern classics that conceptualize the evolution of managerial thought (Barley & Kunda, 1992; Boltanski & Chiapello, 2018; Zuboff, 1988) as well as recent genealogical investigations of workplace datafication (Moore, 2018).

Building upon these works, the next section analyzes similarities and differences in the discourse of AI-nudges with past lineages of managerial thinking to elicits its assumptions.

4.1 Historical linages of AI-nudges

Drawing upon Barley and Kunda's (1992) analysis of the evolution of managerial thinking, Moore (2018) shows how this shift in managerial discourse merges the assumptions and persuasions of several periods in the history of management. Using Barley and Kunda (1992) as a guiding structure, the following section details how managerial practices involving the use of AI-nudges shares significant ideological and intellectual lineages with the four managerial paradigms that have dominated managerial thinking in the twentieth century.

4.1.1 Scientific management Commentators often compare algorithmic management practices to the managerial thinking of Taylorism and scientific management that proliferated in the early-twentieth century (e.g. Altenried, 2022; Jarrahi et al., 2021; Kellogg et al., 2020; Wang et al., 2020). As a form of algorithmic management, managerial practices using AI-nudges shares an interest in using technological advancements to measure and modulate worker's behavior with the pioneers of scientific management. The founding father of scientific management, Frederic Taylor, argued that managerial analysis of work processes was needed because workers could not both observe their work and operate the machine simultaneously (Taylor, 1911). Therefore, workers needed managers to teach them how to carry out their work in the best possible way. AI-nudges enters today's workplace with a similar conviction. For instance the McKinsey article, featured in the introduction, states:

Complex performance management systems with numerous metrics make it difficult for workers and their managers to identify and focus on the issues that matter most. (Amar et al., 2022)

According to the discourse of AI-nudges, modern work has become so complicated that humans need machines to assist them in figuring out how they should carry out their work. To guide workers on how to carry out their work efficiently, the human engineer in scientific management and today's algorithmic management practices both relies on codifications of work. Zuboff (1988) describes how scientific management was driven by a vision of truth, placing the managerial function on a footing of objectivity. This vision was made possible by technology. In scientific management, stopwatches and cameras allowed a transfer of workers' implicit knowledge to systematic list, flowcharts, and measurements.

Similarly, the application of nudge theory in the managerial domain implies a systemization of implicit knowledge. For instance, nudging design frameworks (e.g. OECD, 2019) specify that the first step for developing a nudge is defining measurable behaviors that lead to attractive outcomes. When nudge theory is combined with AI, managers use sensors and computational technology to transfer knowledge from a lived work environment into data, detailing behaviors leading to attractive outcomes for the organization (Kellogg et al., 2020).

However, an important difference between the discourse of AI-nudges and scientific management is their different visions for how these codifications should be used. Whereas the codifications of work are

used to develop more fair pay schemes in scientific management, managerial practices involving AI-nudges aims to steer the workers' behavior by re-configuring the environment in which the decisions are made. The difference can be ascribed to contrasting understandings of what constitutes human nature. Taylor experienced great resistance from workers in implementing the changes that he invented to optimize work processes. He observed how the traditional piece-rating systems of his time involved "a permanent antagonism between employers, and a certainty of punishment for each workman who reaches a high rate of efficiency" (cited in Rudin, 1972, p. 76). Influenced by his training as an engineer, Taylor viewed workers as calculative actors with instrumental orientations to work. According to this logic, managers can only overcome the workers' resistance to technological change if the interests of workers and managers are aligned.

4.1.2 Human relations At Taylor's contemporary, new advanced reward systems did not vanquish workers resistance (Kunda & Barly, 1992). In the 1920s, the rational view of human nature came into question. Elton Mayo raised one of the most influential critiques based on his iconic Hawthorne studies. From observing anomalies in attempts to rationalize work processes, Mayo (1935) concluded that social dynamics is a chief determinant of industrial outputs. Thus, "if explicit attention is not given to this informal organization, the plans of management will miscarry" (ibid. p. 214). The Hawthorne studies were part of a greater shift in managerial thinking. They became the example per excellence in the human relations movement. They illustrated why managerial attention should be geared towards influencing the social dynamics of workplace (Barley & Kunda, 1992).

The discourse of AI-nudges rehashes Mayo's critique of the rational view of human kind. The academic discipline of behavioral economics, which nudge theory is based upon, emerged from researchers observing a series of anomalies in individual decision-making processes, firm behaviors, and market prices (Thaler, 2016). These researchers argued that the anomalies question the fundamental assumptions of the rational view of human kind that underpins classical economic thinking. According to nudge theory, the rational view of human kind is limited because it fails to recognize how the worker's behavior is not only influenced by economic incentives but also by social norms, emotions and so on.

Translated into the debate of algorithmic control, nudge theory implies skepticism towards practices involving ratings, rewards and monitoring of

employees (Cram and Wiener, 2020), which find their legitimization in the rational view of human kind. In contrast, control practices, using AI-nudges, positions software based on AI methods "to play a collaborative role" (Winikoff, 2020), indicating that the view of human nature assumed in discourse of AI-nudges aligns closer with Mayo than Taylor.

4.1.3 Systems rationalism Despite Mayo's and the behavioral economists' critiques, the rational view on human kind continued to influence some trajectories of managerial thinking. In the early-1950s, the rational view of human nature started to proliferate once again as a diverse group of intellectuals and managerial thinkers started to revitalize ideas from scientific management. Barley and Kunda (1992) refer to this shift in managerial thinking as *systems rationalism*. The notion describes a broad group of thinkers—spanning from cybernetics, operations research, and management science. Proponents of system rationalism adopted metaphors from the emerging technology of computers to envision organizations as systems which managers are able to control by manipulating its structures. Workers were conceived as programmable entities and the manager's job was to ensure that the workers were "programmed" to work in ways that contributed to the organization's overall goals.

Peter Drucker is one of the influential representatives of systems rationalisms. He observed how manufacturing work was increasingly being automated and how work in general became more knowledge intensive. According to Drucker (1954), the rising complexity of work generated new demands for workers to engage in the coordination of work and become self-controlling—or as he later framed it, self-managing. To be self-managing require that workers are awarded certain degrees of autonomy over their work (Drucker, 1999). For a self-managing work workforce to be effective, Drucker (1954) argued that workers need to clearly understand the organizational goals and that workers should be measured based on their contributions to these specified goals. He labelled this idea management by objectives (MBO). In many ways, MBO can be understood as an abstracted version of Taylor's piecemeal system. MBO follows Taylor's ambition of aligning interests between workers and the organization. However, instead of specific measurements of concrete work tasks, MBO holds workers accountable for the overall outcome by measuring contributions and performance on an abstract level. This abstraction grants workers autonomy because it is now up to the workers rather than the managers to decide how to perform their work as it decentralizes decision making and managerial authority (Boltanski & Chiapello, 2018).

When platforms and firms use algorithmic ratings to control workers, they draw a direct lineage to MBO. Cast as an alternative to algorithmic distributed ratings, rewards and sanctions, the paper argues that the discourse AI-nudges is committed to a different logic of control than MBO. Moreover, the discourse of AI-nudges also differs from MBO in its similarity to Taylorism, due to the attention given to the worker's specific behaviors over abstracted measures of contribution. Bruni and Sugden (2009) have argued that we can view the project of behavioral economics as the attempt to reverse the shift in economic thinking towards abstract, mathematical, and theoretical modelling. Their analysis suggests that we may view practices involving AI-nudges as representing a reversal of decentralized decision-making starting with systems rationalism.

Yet, the reversal is not a complete return to Taylorism. An integrated element of nudge theory is an insistence on the possibility that organizations can shape behavior in a paternalistic way without limiting individual autonomy. Prior to publishing *Nudge*, Thaler and Sunstein (2003) presented this idea in a conference paper titled: *Libertarian Paternalism Is Not An Oxymoron*. In this paper, they argue that libertarianism and paternalism are often thought of as juxtapositions in the sense that *libertarianism* is the freedom to choose and *paternalism* is limiting the freedom and autonomy of the choice. However, citing research from behavioral economics, Thaler and Sunstein argue that this juxtaposition is not necessarily the case. Thaler and Sunstein (2003; 2009) elaborate on two principles for these juxtapositions to be possible. First, the intervention (i.e., the nudge) "will make choosers better off, *as judged by themselves*" and the intervention "must be easy and cheap to avoid" (2009: pp. 5-6). When Google first introduced nudging to the managerial domain, they celebrated the concept's ability handle the inherent tension in knowledge work between the need to steer workers and preserve their autonomy. "The goal is not to supplant decision-making, but to replace [...] poorly designed structured [...] without limiting freedom (Brock, 2015, p. 292). As such, we argue that the appraisal of nudge theory in the managerial domain derives from the nudging's ability to offer an alternative strategy for managing self-managing workers.

The principles of libertarian paternalism inherent to the concept of nudging have been extensively contested in the debate on the ethics of nudging. Critics characterize nudges as manipulative while its proponents continue to defend nudging by referring to the principles of libertarian paternalism as guidelines for ethical nudges. According to this defense: if a

nudge is manipulative in the sense that it is exploitive or overrule people possibility to choose, then it does not live up to the principles of libertarian paternalism and thus by definition is not a nudge (for details on this argument, see Wilkinson, 2013).

However, when we enter the arena of the workplace, the question of what makes people better off takes a different tone. If we understand the project of scientific management and systems rationalism as exercises in aligning workers' and organizational interests, then the question for an ethical, non-manipulative use if nudge theory becomes: how is a managerial intervention in the interests of the workers? Proponents of AI-nudges may argue that they should only be used in situations where the interests between organizations and workers are aligned. This logic of argumentation mobilizes managerial ideas stemming from the wave of managerial thinking that replaced systems rationalism.

4.1.4 Organizational culture and quality From around the 1970s, managerial discourse started to shift. Endeavors to rationalize organizational processes started to be conceived as creating inhumane machines. Control mechanisms such as MBO were critiqued for being too static and unable to deliver the required flexibility called for by rapid technological developments and waves of outsourcing (Boltanski & Chiapello, 2018). This led managerial thinkers to re-explore ideas from human relations. Firms started to intentionally engineer corporate culture as a mean to ensure the worker's full commitment, offering elaborate scripts for the cognitive and emotional life of the employee (Barley & Kunda, 1992). This period continued the trajectory of decentralizing managerial authority because these elaborate scripts freed workers from the bureaucratic snares of MBO. Rather than crafting structures, the managers were told to become coaches and leaders. The managerial job description included task such as helping workers to discover their full potential and formulate inspiring visions, ensuring the worker's commitment without recourse to compulsion (Boltanski & Chiapello, 2018).

Portrayed as a tailored coaching experience (Amar et al., 2022), the discourse of AI-nudges owes to this period of managerial thinking the idea that the worker's full commitment could be guaranteed, even in situations where there is no direct financial incentive. According to nudge theory, the roots of many organizational problems is not a lack of motivation, but because environmental and contextual conditions prevent workers from working optimally. For instance, Dan and Chip Heath (2010) argue that corporate change programs often fail because they neglect that "what looks like people problem is in fact

a situation problem” (p. 183). The discourse of AI-nudges, embody this logic. One example nudging inspired software Microsoft MyAnalytics, described as:

An emerging class of intelligent tools [...that] aims to help knowledge workers make better-informed decisions around their productivity and work-life balance. (Winikoff et al., 2020)

The idea that managerial technology can work as a helping hand only makes sense when assuming that an organization already have a dedicated workforce interested in improving their productivity. As such, if AI-nudges should not be deemed manipulative, their use rests upon the presumption that some sort of alignment of interests has already taken place, justifying that a change of behavior can be both in the interest of the worker and in the interest of the organization.

4.2 Discursive commitments of AI-nudges

According to Barley and Kunda (1992) managerial thinking have developed in waves alternating between rational (scientific management and systems rationalism) and normative (Human relations and organizational culture) managerial discourses. Next, we use the three analytical questions introduced in section 3.3 to summarize the findings and show how the discourse of AI-nudges transcends the historical established dichotomy between rational and normative managerial discourse.

First, we have seen that AI-nudges, similar to scientific management, assumes that workers’ implicit knowledge can be codified. Second, whereas scientific management utilize codifications of work to design better incentive structures, the discourse of AI-nudges position itself as an alternative to managerial control practices involving formal mechanisms of algorithmic control. Similar to Organizational culture and quality the discourse of AI-nudges portray managers and AI-based software as coaches. Thus, assuming that the worker’s full commitment can be guaranteed, even in situations where there is no direct financial incentive. Third, the difference between formal forms of algorithmic control and AI-nudges, derives from two competing views of what constitutes human nature. Whereas system rationalism assumes organizational problems is rooted in a conflict of interest between workers and managers due to its rational view on human nature, the discourse of AI-nudges, aligns closer with the view of human relations and Organizational culture and quality. Similar to these paradigms, AI-nudges expand the area of managers

attention to social and structural dynamics of the workplace.

The summarization suggests that the discourse of AI-nudges merge ideas from seemingly competing linages of managerial discourse in paradoxical ways.

5. Discussion: Placing a price tag on AI-nudges

Informed by genealogical accounts of the evolution of managerial thought, the paper has provided an analysis of the some of the conditions that had to be met for a discourse on AI-nudges to be possible. This analytical work does not predicate whether we should accept or reject AI in the workplace. It merely allows us to elucidate the assumptions that we commit to when using AI-nudges for managerial purposes. Metaphorically, we can describe the aspiration of this form of critique as stating ‘the price paid’ controlling workers using AI-nudges (Nyman et al., 2021). The paper concludes by presenting three ‘price tags’ on AI-nudges and discusses how critiques of AI-driven nudges can be understood as an unwillingness to pay this price.

5.1 Are you willing to codify behavior?

In their literature review on algorithmic aversion, Burton, Stein and Jensen notice that “a significant proportion of existing research that addresses algorithm aversion and augmented decision making has uncritically adopted the view of the heuristics-and-biases program” (2019, p. 226). The authors argue that this has restricted the development of AI-decision aid systems from including other views of human decision-making such as the notion of ecological rationality advocated for by for example Gigerenzer.

Gigerenzer has been one of the central critical commentators of nudge theory. His work questions the assumption of nudge theory that complex computational models outperform the human capacity to make accurate inferences outside of the experimental designs (Goldstein & Gigerenzer, 2002). Moreover, he cautions that nudging “requires a technocracy of experts” (Gigerenzer, 2015, p. 366). Thus, for AI-nudges to be possible, we must assume that technical experts can in fact effectively codify behaviors for how work should be done. Even though computational components arguably allow managers to partly outsource the work involved with systemizing implicit knowledge to technology, AI methods continues to depend on a human writing the first lines of code. As pointed out in the analysis, the discourse of AI-nudges shares the assumption that

codifications of work are possible with scientific management. Hence, AI-nudges can be cast as a reversal of the trajectory towards decentralizing managerial authority that have been dominating managerial thinking since the 1950s.

This reversal is not unproblematic. We may read the growing literature on algorithmic aversion as resistance towards the codification of workers behavior. Faraj et al. (2018) express this concern of codification when describing organizational practices using AI methods as reductionist. Increasingly, we have come to understand that codifications are never perfect since biased training data may lead to discriminatory models and false recommendations (O'Neil, 2016). On that backdrop, the opacity of AI is highlighted as one of the key concerns for algorithmic management practices (Berente et al., 2021; Lebovitz et al., 2022). We can understand the scepticism towards the opacity of algorithms in the context of work as a violation of our conception that it is workers, not managers who are capable of knowing how work should be planned and carried out. In that light, the problem of the opacity makes it impossible for workers to follow the logic and rationale of the nudge. Hence, since practices involving AI-nudges inevitably implies that a human manager needs to provide some form of input to the machine, we can view scepticism towards AI-nudges as directed towards limitations in managers' ability to codify behaviors and procedures.

5.2 Are you willing to preserve workers' autonomy

Somewhat paradoxically, in the discourse of AI-nudges, the trajectory towards decentralizing managerial authority is never fully reversed. According to the principles of libertarian paternalism, to count as a nudge, workers need to be given the opportunity to opt-out of the nudge. Hence, managerial practices using AI-nudges are positioned as an alternative to practices using AI to restrict information and behaviors of workers. For instance, this happens when platforms use AI methods to narrow available jobs or when organizations create artificial limitations for the available time to conduct a given task (Kellogg et al., 2020). Since nudges do not forbid or mandate actions, practices involving AI-nudges commits to a different set of managerial norms then controlling workers with algorithmic restrictions. As suggested in the analysis, one way to read the emphasis of autonomy in the discourse of AI-nudges relates to the managerial norms of a self-managing workforce.

However, it is not so that the use of AI-nudges are the only form of algorithmic control affording the

autonomy that self-managing practices require. In the analysis, we further described how algorithmic rating practices draw inspiration from system rationalism. Algorithmic rating practices relies grant workers autonomy by rewarding workers based on abstract measurements of contribution rather than specific actions. Nevertheless, rating practices have also been a contested form of control (e.g. O'Neil, 2016). With its image of AI as a coaching tool (Amar et al., 2022), the discourse of AI-nudges aligns to the critique of rating practices being too static to deliver organizational flexibility; originally presented in the discourse of organizational culture and quality. An important take-away of differentiating managerial practices involving AI-nudges from algorithmic control through restrictions and ratings is that the concerns of these practices are not identical. Kellogg et al. (2020) caution that when organizations follow Taylor's principles, breaking down jobs into 'micro' tasks and then use AI methods to restrict, sanction and reward workers, the consequences are likely a shift towards more precariat job roles. Since nudging per definition never involves the use of sanctions and rewards, the use of AI-nudges does not directly contribute to the shift towards more precariat work conditions. Rather the concerns of AI-nudges center around a question of whose interest the nudge nurtures.

5.3 Are you willing to accept that there is no conflict of interest?

As described in the analysis, the central problem for managerial thinking in the twentieth century has been how to ensure a dedicated workforce. According to scientific management and systems rationalism, a lack of alignment between workers' and the organizations' interests is the key driver for organizational problems. According to human relations and organizational culture and quality, organizational problems arise when workers are not aligned with organizational community, goals and values. However, according to nudge theory, the alignment interests and organizational goals and values is not enough. Viewed through the lens of behavioral economics, organizational problems are argued to be rooted in a misaligned context. Workers need nudges, not because incentives are not aligned nor because they are not aligned with organizational goals and values. Instead workers need nudges because the context in some way or the other prevents them from acting in ways that aligns with their own and the organization's interest.

From the analysis, we learn that the assumption that workers' and organizations' interests are aligned

is contestable. When we today accept that there is no inherent conflict between workers and organizations, this is to a large extent the product of the shift in managerial discourse taking place under discourse of organizational culture and quality. However, as demonstrated by a long tradition of critical work this rendering commit to a neoliberal ideology, hiding the often exploitive nature of contemporary capitalism (e.g. Boltanski & Chiapello, 2018). Following a similar argumentation, recent investigations of algorithmic management practices show that AI-nudges do not always promote the interest of workers. One example is Altenried (2022), noticing that when the parcel deliver firm UPS's started to use AI to nudge drivers work more effectively, it led to an intensification of their work. The paper suggests that Altenried's critique can be conceptualized as a concern for the lack of alignment of interest between workers and the organization.

6. Concluding remarks

The paper has taken a discursive view to provide an affirmative critique of managerial practices involving the use of AI-nudges. Importantly, this discursive view does not consider how AI-nudges are being used and resisted. Rather it is an examination of the ideology underpinning its use (see Nyman et al. for an elaboration of the analytical limitations of the problematization lens). The paper has examined the intellectual heritage and ideological roots of AI-nudges. This work contributes to the literature on algorithmic control by showing how managerial practices involving AI-nudges transcends the historical established dichotomy between rational and normative control. Consequently, when examining algorithmic management practices, researchers must be careful not to reduce these practices to a simple form of digital Taylorism.

Developed from the historical examination, the paper presents some of the assumptions that the use of AI-nudges implies. These assumptions may offer a starting point for future research. For instance, future research may examine how the tensions between codification, autonomy and alignments of interest are handled in meetings between human workers and managerial recommending machines. The paper presents the identified assumptions as three questions which can help managers better decide if AI-nudges are relevant to their practice. The three questions can be seen as price tags in the sense that they help managers understand the tacit assumptions that they must be willing to accept for AI-nudges to constitute a meaningful and acceptable control technology. Managers need to be aware of these assumptions since

unwillingness to commit to them foreshadow the inherent concerns related to the use of AI-nudges.

7. References

- Alvesson, M., & Sandberg, J. (2011). Generating Research Questions Through Problematization. *Academy of Management Review*, 36(2), 247–271.
- Altenried, M. (2022). *The Digital Factory: The Human Labour of Automation*. The University of Chicago Press
- Amar, J., Majumder, S., Surak, Z., & von Bismarck, N. (2022). *How AI-driven nudges can transform an operation's performance*. McKinsey & Company. <https://www.mckinsey.com/business-functions/operations/our-insights/how-ai-driven-nudges-can-transform-an-operations-performance>
- Barley, S., & Kunda, G. (1992). Design and Devotion Surges of Rational and Normative Ideologies of Control in Discourse. *Administrative Science Quarterly*, 37, 363.
- Berente, N.; Gu, B.; Recker, J.; and Santhanam, R. (2021). Special Issue Editor's Comments: Managing Artificial Intelligence. *MIS Quarterly*, 45(3), 1433–1450.
- Beshears, J., & Gino, F. (2015). Leaders as Decision Architects. *Harvard Business Review Digital Articles*. <https://hbr.org/2015/05/leaders-as-decision-architects>
- Bock, L. (2015). *Work rules! - insights from inside Google that will transform how you live and lead*. John Murrey.
- Boltanski, L., & Chiapello, E. (2018). *The New Spirit of Capitalism*. Verso (trans. G. Elliot. 2nd ed.)
- Bruni, L., & Sugden, R. (2007). The road not taken: how psychology was removed from economics, and how it might be brought back. *The Economic Journal*, 117(516), 146–173.
- Burr, C., Cristianini, N., & Ladyman, J. (2018). An Analysis of the Interaction Between Intelligent Software Agents and Human Users. In *Minds and Machines* (Vol. 28, Issue 4). Springer Netherlands.
- Burton, J. W., Stein, M., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*.
- Cram, W., & Wiener, M. (2020). Technology-mediated control: Case examples and research directions for the future of organizational control. *Communications of the Association for Information Systems*, 46, 70–91.
- Drucker, P. F. (1954). *The Practice of Management*. Harper.
- Drucker, P. F. (1999). Knowledge-Worker Productivity: The Biggest Challenge. *California Management Review*, 41(2), 79–94.
- Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, 28(1), 62–70.
- Foucault, M. (2008). *The birth of biopolitics : lectures at the Collège de France, 1978-79* (M. Senellart, F. Ewald, & A. Fontana (eds.)). Palgrave Macmillan.
- Foucault, M. (1984). "What is Enlightenment?", in Paul Rabinow (eds), *The Foucault Reader* (1st ed.).

- Pantheon Books.
- Foucault, M. (1990). *The history of sexuality, Vol. 2: The Use of Pleasures*, R. Hurley, (eds). Vintage Books
- Gal, U., Jensen, T. B., & Stein, M. K. (2020). Breaking the vicious cycle of algorithmic management: A virtue ethics approach to people analytics. *Information and Organization*, 30(2).
- Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology*, 6(3), 361–383.
- Gino, F. (2017). Uber Shows How Not to Apply Behavioral Economics. *Harvard Business Review Digital Articles*. <https://hbr.org/2017/04/uber-shows-how-not-to-apply-behavioral-economics>
- Goldstein, D. & Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychol. Rev.* 109(1), 75-90.
- Hacking, I. (2007). Kinds of people: Moving target. *Proceedings of the British Academy*, 285–318.
- Hansen, P. (2016). The Definition of Nudge and Libertarian Paternalism: Does the Hand Fit the Glove? *European Journal of Risk Regulation*, 7(1), 8
- Heath, D., & Heath, C. (2010). *Switch: How to change things when change is hard*. Crown Business.
- Jarrah, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data and Society*, 8(2).
- Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*, 14(1), 366–410. <https://doi.org/10.5465/annals.2018.0174>
- Koopman, C. (2019). *How we became our data : a genealogy of the informational person* (1st ed.). University of Chicago Press.
- Koukouvouinou, P., & Holmström, J. (2022). AI Management Beyond Narratives of Dystopian Nightmares and Utopian Dreams. *ECIS 2022 Research Papers*, 137.
- Land, F. (2010). The use of history in IS research: An opportunity missed. *Journal of Information Technology*, 25(4), 385–394.
- Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 1–45.
- Mayo, E. (1935). The blind spot in scientific management. *Proceedings of the Development Section. Sixth Annual Congress for Scientific Management*.
- Mele, C., Russo Spena, T., Kaartemo, V., & Marzullo, M. L. (2021). Smart nudging: How cognitive technologies enable choice architectures for value co-creation. *Journal of Business Research*, 129, 949–960.
- Möhlmann, M., Zalmanson, L., Henfridsson, O. & Gregory, R. (2021). Algorithmic Management of Work on Online Labor Platforms: When Matching Meets Control. *MIS Quarterly*, (45: 4), pp.1999-2022.
- Möhlmann, M. (2021). Algorithmic Nudges Dont Have to Be Unethical. *Harvard Business Review Digital Articles*. <https://hbr.org/2021/04/algorithmic-nudges-dont-have-to-be-unethical>
- Moore, P. V. (2018). *The Quantified Self in Precarity Work, Technology and What Counts*. Routledge
- Myers, M. & Klein, H. (2011). A Set of Principles for Conducting Critical Research in Information Systems. *MIS Quarterly*, 35(1), 17-36.
- Nyman, S., Gudmand-høyer, M., Jensen, T. B., & Gudmand-høyer, M. (2021). Exploring the Archive : A Problematization Lens for Conducting Critical IS Research. *ICIS 2021 Proceedings*. 7., 0–17.
- O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.
- OECD. (2019). *The BASIC Toolkit: Tools and ethics applied for Behavioural Sights*. <https://www.oecd.org/regreform/tools-and-ethics-for-applied-behavioural-insights-the-basic-toolkit-9ea76a8f-en.htm>
- Rowe, F. (2018). Being critical is good, but better with philosophy! From digital transformation and values to the future of IS research. *European Journal of Information Systems*, 27(3), 380-393.
- Rudin, B. (1972). Industrial betterment and Scientific Management as Social Control. *Berkley Journal of Sociology*, 17, 59–77.
- Stahl, B. C. (2011). Critical Social information System Research, in Galliers, R. & Currie, W. (ed). *The Oxford Handbook of Management Information Systems: Critical Perspectives and new directions*, 199-228. Oxford University Pres
- Swanson, E. B., & Ramiller, N. C. (1997). The Organizing Vision in Information Systems Innovation. *Organization Science*, 8(5), 458–474.
- Taylor, F. (1911). *Principles of scientific management*. Harper and Brothers.
- Thaler, R. (2016). *Misbehaving: The making og Behavioral Economics*. W. W. Norton & Company.
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian Paternalism. *The American Economic Review*, 93(2), 175–179.
- Thaler, R., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth and happiness*. Penguin Books.
- Wajcman, J. (2019). The Digital Architecture of Time Management. *Science Technology and Human Values*, 44(2), 315–337.
- Wilkinson, T. M. (2013). Nudging and Manipulation. *Political Studies*, 61, 341–355.
- Wang, B., Schlagwein, D., Cecez-Kecmanovic, D., & Cahalane, M. C. (2020). Beyond the factory paradigm: Digital nomadism and the digital future(s) of knowledge work post-COVID-19. *Journal of the Association for Information Systems*, 21(6), 1379–1401.
- Winikoff, M., Cranefield, J., Li, J., Richter, A., & Doyle, C. (2021). The Advent of Digital Productivity Assistants : The Case of Microsoft MyAnalytics. *54th Annual Hawaii International Conference on System Sciences (HICSS)*.
- Zuboff, S. (1988). *In the age of the smart machine : the future of work and power*. Basic Books.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism*. Profile Books.