INSIGHTS FROM *THE MACHINE STOPS* TO BETTER UNDERSTAND RATIONAL ASSUMPTIONS IN ALGORITHMIC DECISION-MAKING AND ITS IMPLICATIONS FOR ORGANIZATIONS

Dirk Lindebaum
Cardiff Business School
mail@dirklindebaum.EU
www.dirklindebaum.EU

Mikko Vesa
Hanken School of Economics, Helsinki
mikko.vesa@hanken.fi

Frank den Hond
Hanken School of Economics, Helsinki
and VU Amsterdam
fdenhond@hanken.fi

review essay - no abstract and keywords required

Disclaimer: This essay represents the synthesis of a long processes of deliberation and intellectual struggle amongst the authors about the wider topic of algorithmic decision-making at work and in society. Following Parkinson (2019), we underline that this essay is our original work and that it was produced without the help of an algorithm.

Acknowledgements: The editor Jean Bartunek and the three reviewers provided most insightful and stimulating comments during the review process. In addition, Christine Moser's comments, and the feedback from the MOTI seminar series at Grenoble Ecole de Management, helped us consolidate the arguments presented in the essay. We are indebted to all of them for the intellectual generosity and guidance. A final 'thank you' (from DL) goes to Virginia Brown for having left Forster's short story as small gift on his desk one day. It was only a small gift, but its inspiration for this essay was extraordinary.

Cite as:
Lindebaum, Vesa, & den Hond (2019) Insights from the Machine Stops to better understand rational assumptions in algorithmic decision-making and its implications for organizations. *Academy of Management Review*. (Check AMR 'in press' section for updates)

> *We created the Machine, to do our will, but we cannot make it do our will now. It has robbed us of the sense of space and of the sense of touch, it has blurred every human relation and narrowed down love to a carnal act, it has paralysed our bodies and our wills, and now it compels us to worship it. The Machine develops – but not on our lines. The Machine proceeds – but not to our goal.*

These are the words of Kuno, the protagonist in E. M. Forster's much-acclaimed short story *The Machine Stops*.[1] In this review essay, Forster's story is a source of inspiration for us to reflect on and problematize (i) the assumptions of rationality underlying algorithms and algorithmic decision-making, (ii) the possible role of organization and management theory (OMT) in the rapid rise of algorithmic decision-making, and (iii) the implications of algorithmic decision-making for organizations.

Decision-making is, of course, central to organizations, and artificial intelligent (AI) algorithms are increasingly being used in organizational decision-making. As Pasquale notes, "authority is increasingly expressed algorithmically" (2015: 8). This development is applauded by some and criticized by others; the popular debate is whether or not algorithms will become 'more intelligent' than humans and relegate them to insignificance (Barrat, 2013), or turn out to be a blessing for humanity and management (Amabile, 2019; Bostrom, 2014). However, the debate is somewhat misguided in the first place; the assumptions of rationality that underlie algorithmic decision-making remain neglected (one example is Vallor, 2016). This neglect comes at a cost, since these assumptions of rationality do not remain mere assumptions; they translate into regularities and patterns of action in social life (Kalberg, 1980). Algorithms have real bearing on decision-making in organizations, as we shall outline.

[1] The story was initially published in *The Oxford and Cambridge Review* in November 1909, but appeared in book form only in 1928 in Forster's collection of short stories titled *The Eternal Moment and other Stories* (1928/2011). All direct quotes in this essay are taken from Forster (1928/2011). The introductory quote is taken from pp. 33-34.

By means of concept clarification, we define algorithms as "processes or sets of rules to be followed in calculations or other problem-solving operations, especially by a computer" (*Oxford Living Dictionary*). On that base, we refer to algorithmic decision-making as *automated decision-making based on predefined rules and/or goals*. Algorithmic decision-making comes with a set of assumptions, and Forster's story made us realize that a reflection on these assumptions is warranted to understand better what algorithmic decision-making is and how it may affect organizations. Specifically, in this review essay, we offer a discussion and interpretation of Forster's story, in which the *Machine* becomes the embodiment of algorithmic decision-making pushed to the limit. The result is a technology-enabled totalitarian system (see also Helbing et al., 2017), whose rise to domination and later demise, when the *Machine* stopped, were both associated with its predominant (or perhaps its sole) reliance on formal rationality. Forster offers us a glimpse of the state of humanity in that dystopian society by allowing us to peer into the lives of Kuno and his mother, Vashti.

We see a strong parallel in the emergence of the *Machine* and the contemporary rise of AI-based algorithmic decision-making. As Kuno's opening words show, 'we, humans, created the *Machine*'. Likewise, 'we, humans' now push the introduction and use of algorithmic decision-making in ever-more domains of our society of organizations (Andrews, 2018; Friedland, in press). Driven by our normative commitment, we believe it is crucial to appreciate fully how serious, extensive, and deep the impact of AI-based algorithmic decision-making is already across many spheres of social life. For example, algorithms are used to obtain what is held to be the *correct* medical diagnosis in a much faster and more efficient way (Froomkin, 2016), and the *best possible way* to process financial transactions and the revenues it generates (Hendershott, Jones, & Menkveld, 2011). In China, by way of facial recognition, traffic

offenders are issued fines for jaywalking in *real time*, which they receive as text messages on their phones (Baynes, 2018). The role of algorithms in warfare (cyber or 'traditional') is increasingly apparent (Tarnoff, 2018). In 2018, Google announced that its search engine can now provide answers to questions before they have even been raised (De Vynck & Grant, 2018). Relevant for management and organization studies are cases in which organizations seek to map the 'brain' of successful CEOs onto an algorithm for more efficient decision-making (Copeland & Hope, 2016), screen job applications through algorithms (Economist, 2018), and install an artificial intelligent (AI) system as a member of the Board (Libert, Beck, & Bonchek, 2017). Amidst these and other applications of AI-based algorithmic decision-making are also numerous cases of 'data harm', some of which can be intentional while others are unintended (for examples, see O'Neill, 2016; Redden & Brand, 2018). Taken together, these examples point to growing concern about the 'automation of society' (Helbing et al., 2017).

We draw upon Max Weber's (especially 1915/1946; 1922/1978) notions of formal and substantive rationality to scrutinize the assumptions underlying algorithmic decision-making and its implications. Decision-making premised on formal rationality is based on following abstract and formal procedures, rules and laws, which are taken as unproblematic and legitimate fixed ends. Formal rationality encompasses means-end calculation to the extent that the calculus itself is a strict procedure toward optimization or maximization of outcomes. It rejects arbitrariness in decision-making that flows from the personal qualities and experiences of individual decision-makers or from situational idiosyncrasies.

We refer to AI-based algorithms as *supercarriers* of formal rationality. Algorithms are *carriers* of formal rationality because they can only operate through logical and mathematical procedures, typically for purposes of optimization. But AI-based algorithms are more than that.

They have become *super*carriers of formal rationality in two ways. For one, they can improve themselves in performing their tasks – in this sense, algorithms can 'learn'. Given constraints and/or boundary conditions, they work to find the logically or mathematically 'correct' solution in a dataset, and upon being given more data, they are able to develop a 'more correct' solution. A predefined criterion for optimization is given in the case of 'supervised learning' (i.e., when software engineers prescribe what the optimum or goal is), but not so in the case of 'unsupervised learning' (i.e., when the goal is a 'meta' goal about another goal, for instance, when algorithms estimate individuals' preferences in order to act on their behalf).[2] Algorithms are also supercarriers because of the unprecedented speed – even in real time as shown before – of processing vast volumes of 'big' data to produce or calculate outcomes, although in ways that are often intractable and opaque, even to their programmers (Knight, 2017). If algorithms are supercarriers of formal rationality, then decision-making based on AI algorithms reflects the aspiration toward 'enhanced decision-making capability' or 'augmented rationality' (Cohen, 2007: 504), in management and organization, and beyond.

Yet, formal rationality as expressed in algorithmic decision-making may reduce the possibility to decision-making based on substantive rationality; that is, the capacity of humans for value-rational reflection and action (Føllesdal, 1994). Substantive rationality orders action

[2] Algorithms generally require goals to be set before any action is taken, and the examples of supervised learning and unsupervised learning help explain this point. Algorithms can 'optimize' when they are fed more data (as is the case with ad preferences and retailer recommendations, or automatic image recognition). In addition, so-called 'deep-learning' technologies have facilitated the discovery of intricate structures in large data sets by employing the backpropagation algorithm to suggest how a machine should alter its internal parameters used "to compute the representation in each layer from the representation in the previous layer" (LeCun, Bengio, & Hinton, 2015: 436). Deep learning computational models are composed of multiple processing layers to acquire representations of data with multiple levels of abstraction. In terms of its characteristic, deep learning basically refers to an "autonomous, self-teaching system in which you use existing data to train algorithms to find patterns and then use that to make predictions about new data" (Marr, 2018). Depending on the complexity of the algorithm (see Tutt, 2016), a self-learning algorithm can thus become more 'intelligent' by 'optimizing' its programming *en route* to goal fulfilment. This development blurs the distinction between 'code' and 'data'. We are grateful to Nikos Gorogiannis from Facebook for clarifying these differences for us.

into patterns through appeals to a cluster of past, present, or potential value postulates that vary in content, comprehensiveness and consistency (Kalberg, 1980). There is a vast, perhaps even infinite, scope for value postulates, and hence action can be ordered into patterns in numerous different ways. Therefore, substantive rationality relies on 'valuation' (Dewey, 1939), in the sense of an effort to develop "genuine and grounded propositions about the purposes, plans, measures, and policies which influence human activity whenever the latter is other than merely impulsive or routine" (Dewey, 1939: 57). Substantive rationality contains the possibility to see normatively "the world as it might be" (Suddaby, 2014: 408), involving 'what is', 'what can' and 'what ought to be' in empirical, moral and aesthetic terms. In our essay, the issue of morality in the context of algorithmic decision-making is particularly relevant and will be revisited later again.

Received wisdom suggests that formal and substantive rationalities are often treated as antagonistic or opposites (Weber, 1922/1978); they inform in radically distinct ways the conscious regularities of action that are needed to "master *(beherrschen)* fragmented and disconnected realities" (Kalberg, 1980: 1148). However, we advocate a more complex and dynamic understanding of the relationship between the two forms of rationality. First, one would lose mastery of these fragmented realities when one type of rationality is privileged over, or imposed upon, the other. This is illustrated in Kuno's opening words of *The Machine Stops,* which suggest a tale of formal rationality dominating over substantive rationality, while his final words in the story (as quoted in the final paragraph of this review essay) suggest a resurfacing need for substantive rationality.[3] Second, there is the possibility of a transformation of

[3] The story ends with Kuno's final words, which signal that the domination of formal over substantive rationality is coming to an end. We do not know, however, what comes next. Forster never wrote a sequel to the *Machine Stops*. But if such sequel were to be written, and if characters therein were to learn from prior mistakes, it probably would

substantive into formal rationality through formalization. Substantive rationality becomes formal rationality when it is expressed in a universal, absolute rule, thus reifying its value orientation. In the story, that is evident when a moral principle exhorts individuals to be 'mechanical' in their thinking, feeling and action.

Building on what we already suggested, we foreshadow two key arguments in our essay. First, as supercarriers of formal rationality, algorithms do not only rely on formal rationality, but also reinforce formal rationality in their domain of application (e.g., in organizational decision-making). Concurrently, because formal rationality informs algorithms, this development can have two consequences: (i) it may reinforce the suppression of substantive rationality through formal rationality, or it may result in a transformation of substantive rationality into formal rationality through formalization. Both consequences may well imply the end of choice. Second, the advancement of algorithms in organizational decision-making is perpetuated by the striving for an 'ideal' state of reality that is profoundly influenced by the quest for 'perfect' rationality in decision-making. Algorithms provide the technological solution through which to attain this glorious state of affairs. To the extent that organizational theory adopts or privileges the same assumption of formal rationality for decision-making (e.g., Simon, 1947/1997), there is a possibility that organizational theory becomes complicit, even facilitates, the end of choice through legitimizing algorithmic decision-making.

Our review essay progresses as follows. First, we offer a general overview of the plot of *The Machine Stops* and tease out specific arguments to illuminate key theoretical themes. Toward the end of that overview, we move from description to interpretation and show more closely how our interpretation of the story informs our theorizing about the rational assumptions

---

entail a pluralist approach in giving a more balanced weight to substantive and formal rationality, or making formal rationality support substantive rationality.

underlying AI-based algorithmic decision-making, the possible role of OMT in the rise of such decision-making, and the implications thereof for organizations. In our interpretation, the *Machine* and its followers, including Vashti embody formal rationality, whereas Kuno represents the capacity for substantive rationality. Second, we briefly discuss these Weberian notions of formal and substantive rationality to support the argument that algorithms are supercarriers of formal rationality. However, Weber was of course not the only scholar preoccupied with rationality. Therefore, in the third section, it is relevant to show that OMT features strands that are predominantly grounded in formal rationality (e.g., Simon, 1947/1997), and this grounding helps account for the remarkable rise of algorithmic decision-making in organizations. Having said that, as in the Weberian view on rationality, there are of course other strands in OMT grounded in a more pluralist tradition that defend the significance of substantive rationality alongside formal rationality and argue for formal rationality to be supportive of substantive rationality (e.g., Mintzberg, 1994). Fourth, we consider the implications of algorithms as supercarriers of formal rationality for organizational decision-making, especially their potential to represent enhanced decision-making capability. Finally, we synthesize the contents of our essay and offer a generic discussion on theoretical and practical ramifications of AI-based algorithmic decision-making. At this point, we acknowledge that the genre of essay writing enables us both to signal clearly our normative concerns about the topic, while being mindful of the complex nature of the phenomena under discussion.

THE MACHINE STOPS – BACKGROUND AND PLOT

Forster's short story is cast in a dystopian future, in which the *Machine* controls all aspects of life. The *Machine* is an all-encompassing decision-making system of apparent super-human qualities, omnipotent and all-knowing: in short, a technology-based totalitarian system. Because

the Earth's surface has become inhabitable ("One dies immediately in the outer air", p. 6), humans have become dwellers of subterranean hexagonal chambers that resemble bee cells. The *Machine* assigns each of them a chamber. The logic of the design is that the rooms contain very little, but their dwellers enjoy full service. Neither window nor lamp provide light, yet a soft radiance engulfs the rooms. No overt provision is made for ventilation, yet the air is fresh. There is no furniture in the room, except for an armchair and a reading desk by its side. A bed – "of the same dimension all over the world" (p. 8) – can be 'summoned', and a medical apparatus falls from the ceiling in case its dweller has a health complaint. The eternally humming *Machine* thus caters for all the physical needs of the members of the social system it supports. All but a few of its members are convinced that their present civilization is much advanced over the previous one, which foolishly "had mistaken the functions of the system, and had used it for bringing people to things, instead of for bringing things to people" (p. 10).

The *Machine* also ensures the provision of cultural, spiritual and even educational needs. There are no instruments in the rooms, yet upon a dweller's demand, his/her room is "throbbing with melodious sounds" (p. 1). Communication, too, is arranged through the *Machine* in the form of instant messaging and video conferencing (like present-day's affordances of Facetime and Skype, for example), leading users to claim that they know thousands of people (like present-day users of social media such as Facebook and Twitter). In its reliance on these technologies, the "clumsy system of public gatherings had been long abandoned" (p. 7), so that ancient customs and practices, such as going out to meet other people, have been made obsolete. Owing to the *Machine*, people no longer touch each other, and never do they have to leave their cells – apart from the 'carnal act', performed upon the *Machine's* command. Crucially, in parallel with these changes in customs and practices, societal attitudes toward life have also changed, including

those toward newborn life. Given that physical activity, such as a displacement to another location, is no longer needed and actually discouraged, it has become the conviction that being muscular is a demerit. Therefore, every newborn child is examined at birth, and "all who promised undue strength were destroyed" (p. 24). Forster comments dryly: "humanitarians may protest," but nevertheless explains how "it would have been no true kindness to let an athlete live; he would never have been happy in that state of life to which the Machine had called him" (p. 24). The imposition and acceptance of the *Machine's* will has thus created a new reality, one that is hostile to the kind of human life that many – including ourselves – would like to live.

The story is told in three sections. At the beginning of the first section – "the Airship" – Vashti sits on her armchair, with a face "white as a fungus" (p. 1). She is the other character in the story, mother to Kuno, who lives in another continent. Vashti is one of many so-called lecturers, whose task it is to ceaselessly reproduce the approved, official, and indisputable version of knowledge. She is preparing for her 10-minute lecture on the history of music. The plot opens with Vashti being irritated by yet another interruption of her preparations. Somebody seeks to contact her. It turns out to be Kuno, beseeching his mother to come and 'see' him. "But," she objects flatly, "I can see you!" Kuno, a sensualist and rebel in the making, is adamant: "I want to see you not through the Machine . . . I want to speak to you not through the wearisome Machine." With vague shock, Vashti responds that Kuno "mustn't say anything against the Machine." Baffled with incomprehension, Kuno replies:

> *You talk as if a god had made the Machine . . . I believe that you pray to it when you are unhappy. Men made it, do not forget that. Great men, but men. The Machine is much, but not everything. I see something like you in this plate, but I do not see you. I hear something like you through this telephone, but I do not hear you. That is why I want you to come*. (p. 3).

Reluctantly, Vashti agrees to venture the journey, take the 'air-ship' and visit her son on the other side of the world. It is a journey permitted by the *Machine*, though traveling is generally discouraged. Because direct social encounters are so rare, the horror of direct experience hits Vashti full force as she queues to enter the battered air-ship. She has to endure the direct glances from the other passengers, which causes her profound discomfort. Once aboard, all passengers keep to their own abhorred by the discomfort caused by direct, first-hand experience; they even shut the blinders to prevent them from accidentally glancing out of the air-ship's windows.

Upon her arrival – at the beginning of the second section, "The Mending Apparatus" – Kuno confides in her his unbearable disenchantment with the sanitized and mechanical world they inhabit, sustained and governed solely by the *Machine*. Even though the notion of 'exploration' – and the associated sense of curiosity – is contrary to the spirit of the time, Kuno had already ventured to the surface of the Earth. He tells Vashti about his clandestine visit, how he miraculously survived, that he saw other human beings that live outside the realm of the *Machine*, and that he was recaptured and returned to his room by long white worms: the *Mending Apparatus*. After his visit, Kuno was threatened with 'homelessness', which entails being banished from the safe underground life to likely death on the surface of the Earth. Vashti reacts with shock about her son's actions, which for her amounts to dangerous madness. Ashamed of ever having giving birth to him (and later disavowing him altogether!), she returns to her part of the world and for some time continues her normal life.

Two significant developments occur in the years after Kuno's escapade. They are recounted in the third section, "The Homeless." For one, the life support facilities called 'respirators', needed to survive a visit to the outer world and one of which Kuno brought with him on his venture, are decommissioned – a development that most people welcome. For

another, religion is re-established, although "the word 'religion' is sedulously avoided" (p. 42). By now, people appreciably (and overtly) worship the *Machine*:

> *The Machine . . . feeds us and clothes us and houses us . . . in it we have our being. The Machine is the friend of ideas and the enemy of superstition: the Machine is omnipotent, eternal; blessed is the Machine* (p. 42).

This development, however, implies that the people forgot that they themselves invented the *Machine*. Worse still, they have come to treat the *Machine* as some kind of a mystical deity whose needs matter more than their own. There is now the *Book of the Machine* to provide comfort and answers to all questions, morphed as it had from being a mere technical manual to the *Machine* into a vicarious artefact of worship, like any holy scripture. Those who do not worship are considered 'unmechanical' and risk the punishment of homelessness.

Then, the *Machine* starts to show defects and, moreover, the device designed to fix these problems, the *Mending Apparatus*, turns out to be in need of repair itself, as it is acknowledged. Reassured that the problems with the *Machine* are not caused by sabotage but simply its whims, "men spent their strength praying to their Books, tangible proofs of the Machine's omnipotence" (p. 50).

At this stage, Kuno is transferred to a cell close to Vashti's. Being a sensualist, he warns that the *Machine* is breaking down. Cryptically, he tells Vashti "The Machine stops" (p. 45), but Vashti continues her life. The defects in the *Machine* become more serious and the situation continues to deteriorate; having become wholly dependent on and subservient to the *Machine*, people had forgotten how to fix it. More specifically:

> *The better a man knew his own duties upon it, the less he understood the duties of his neighbor, and in all the world, there was no-one who understood the monster as a whole. Those master brains* [who developed the Machine] *had perished*. (p. 43)

The plot then draws to a close; the *Machine* collapses, and the entire subterranean 'civilization' perishes with it. There are explosions in the subterranean system, killing many people. Unable to handle the experience of other people's sufferings, many die outright at the face of it. In the final scene, Vashti and Kuno, both injured, weep "for humanity . . . not for themselves" (p. 54). They crawl over dead bodies to find themselves, and in a way to find humanity in themselves. "Quicker," Kuno gasps, "I am dying – but we touch, we talk, not through the Machine," adding "we have come back to our own" (p. 55). The so-called 'homeless' people on the surface of Earth are now to rebuild humanity's future.

Merely describing the storyline of Forster's short story does not absolve us from the need to offer an interpretation of the story, and how that relates to the theorizing we present here. Wikipedia suggests that Forster's short story is visionary in having anticipated the arrival of many concurrent technologies, including instant messaging and the internet.[4] Our interpretation of *The Machine Stops* goes beyond such technological novelties to consider the possibly broader context of Forster's concerns, namely, the emergence, growth, and eventual demise of a technology-enabled totalitarian system. The *Machine* is totalitarian in being "a system of government that is centralized and dictatorial and requires complete subservience to the state" (*Oxford Living Dictionary*). In prohibiting opposition and exercising tight control over public and private life through its centralized organs, the *Machine* anticipates and parallels characteristics of historical totalitarian systems. The *Central Committee* (as described in the story), for example, is similar in name and function to the later *Zentralkommitee* in the former German Democratic Republic that emulated parts of the political structure of other countries under Soviet-style communist rule. As in that system, the *Machine* requires complete

[4] See https://en.wikipedia.org/wiki/The_Machine_Stops, accessed on 19 Nov 2018.

subservience of its people; any failure to do so provokes retaliation from the system, as well as horror among its subjects (e.g., on the part of Vashti when she encounters the blasphemous disrespect that Kuno displaced toward the *Machine*). We can take recourse to Arendt (1958: viii) to underline the cost of that subservience:

> *The totalitarian attempt at . . . total domination has been the destructive way out of all impasses. Its victory may coincide with the destruction of humanity; wherever it has ruled, it has begun to destroy the essence of man.*

However, Forster takes the analysis of a centralized totalitarian society a step further by showing how, in his story, it can be *technology-based*. Wikipedia's view that Forster's story merely anticipated present day innovations is naïve in light of the totalitarian nature of the *Machine's* society. Although it may not have been created *a priori* to this end, technology has morphed into an enabler of the *Machine's* totalitarian society. How did this happen? And why did the *Machine* stop?

It has to be recognized how the rise of the *Machine* is premised on humanity's fall into decadence. Initially, people invented the *Machine* for their convenience. Then they relied on it and became increasingly dependent upon it. Finally, they worshipped it, only to stare into the abyss they were bound to fall into, when the *Machine* collapsed, due them no longer being able to make decisions about the organization of their society and themselves (see the introductory quote by Kuno). The *Machine* can be seen as a mirror image of our own use of technology in general, and of AI-based algorithmic decision-making in particular. We thus propose that machines originate in the tendency of humans to ease their lives, to make them more comfortable and to reduce effort, largely through technological innovations. Machines were invented, first, to make it easier to 'bring people to things', and then, through further invention to 'bring things to people,' as Forster suggested. To the extent that such machines rely on science and technology, a

program (i.e., an algorithm) governs their development, production, and use (Flusser, 1983/2000); they rely on formal rationality. In their reliance on formal rationality, they demand of their users, too, the adoption of formal rationality and their further socio-technical integration can only proceed on the base of that same formal rationality. In this way, as the number of machines increases, formal rationality pervades ever more spheres of life. In Forster's story, machines became the *Machine* and its decision-making reflected the formal rationality in the programming of its algorithms.

It is no stretch to imagine a further escalation toward formal rationality as it justifies adherence to the rules it expresses: formal rationality breeds formal rationality. The eventual demise of the *Machine* can then be attributed to its predominant, if not sole, reliance on formal rationality at the expense of substantive rationality. The *Machine* became so pervasive that it crowded out the possibility to make decisions based on substantive rationality.[5] It is this inflexibility that renders decision-makers unable to deal with contingencies (i.e., events that deviate from universal rules that govern social life). People had forgotten how to fix the *Machine*, the master brains who developed it had perished. Time thus had taken its toll, as had the ever-growing complexity of the *Machine*. But more importantly, people had forgotten where to look for the source of the problem, the *Machine* had taught them to be helpless (Moore, 2018) in the face of contingencies. They were no longer able to think for themselves. Yet, the ability to deal with contingencies is vital for the survival of societies as it helps manage the multiplicity of value postulates that maintains rather than destroys the essence of humanity (Dewey, 1922; Elster, 1994; Kalberg, 1980).

[5] A related, complementary argument has been made by Clegg and his colleagues (2008), when they pointed out the allure of formal rationality for decision-makers that wish to feel in control: a sense of ambiguity may render them likely succumb to its promise of certainty.

Forster's story lends itself exquisitely for the purpose of our analysis, for it is rife with vivid and imaginative insights that feature significant parallels with ongoing debates on AI-based algorithms and algorithmic decision-making. Therefore, we have woven relevant sections of the story into the academic literature that follows (see also Lindebaum & Courpasson, 2019).

THE RATIONALITY OF ALGORITHMS FROM A WEBERIAN PERSPECTIVE

On a general level, Weber (e.g., 1915/1946; 1922/1978) saw rationality as a conceptual scheme through which to study regularities and patterns in social life. Different types of rationality, and the rationalization processes that follow from them, inform decision-making and thereby influence what kinds of actions are pursued. In this sense, something is not 'rational' or 'irrational' in and of itself, but becomes so only once examined from the point of view of a particular type of rationality (Weber, 1920/2002); what may be considered 'rational' under one type of rationality may not be considered so under another type of rationality.

In terms of its characteristics, formal rationality legitimizes means-end calculation and reliance on abstract and universally valid rules. To the extent that brute calculation reigns with regard to abstract rules, decisions are arrived at "without regard to persons" (Kalberg, 1980: 1158). The orientation of action toward compliance with formal rules and laws is synonymous with a denial of randomness and arbitrariness *ex post* (whereas it can accommodate *ex ante* variability and uncertainty). In other words, decision-making premised on formal rationality is not based upon the qualities of the individual concerned – neither the judgment of the decision-maker, nor the specific conditions of the decision-taker – but predicated upon universalism and calculation with reference to formal rules and regulations (Kalberg, 1980). It thus represents an orientation to decision-making that is inhospitable to the diversities, complexities, spontaneity, vicissitudes and richness of human life, and may generate a kind of human life that many would

reject. The *Machine* embodies this state of affairs, in which its 'mechanical' ruling eventually led to the decline of a civilization that considered itself, as it should be remembered, highly 'advanced.' Because, as a machine, the *Machine* is programmed through algorithms and – as we proposed in the introduction – (AI-based) algorithms are (super)carriers of formal rationality, the *Machine* is the embodiment of formal rationality.

As to substantive rationality, it calls on decision-makers to tolerate considerable ambiguity; because there are myriad values, substantive rationality can be configured in many different ways (Weber, 1922/1978). The large number of value postulates underlines a critical characteristic of Weber's substantive rationality, namely, that of radical perspectivism (Kalberg, 1980). For Weber, this radical perspectivism is associated with his view that values are not justifiable through the methods of science (see e.g., Weber, 1921/1946). In addition, there is no constellation of values that represents a persistent 'standard' for 'rational' decision-making and associated rationalization. Therefore, substantive rationality constitutes a domain in contemporary life where the "autonomous individual confronts his 'own demons'" (Kalberg, 1980: 1157) and where one has to think for oneself. In *The Machine Stops*, Kuno appears to be the only one with a surviving sense of substantive rationality, whereas all others have systematically 'unlearned' it in the absence of first-hand experiences, a thread that runs consistently through the short story. The notion of radical perspectivism is of intrinsic significance in the context of *Machine* and algorithms. As supercarriers of formal rationality, algorithms may erase the condition of radical perspectivism and thereby usher the end of choice

We underline two crucial points in our understanding of Weber's concepts of rationality. Both imply that substantive rationality is often unable to stand firmly against the impersonal character of formal rationality. First, and in spite of "the most direct antagonism" between formal

and substantive rationality and the "particularly fateful role" that "the recurrent conflict of these types of rationality has played . . . in the unfolding of rationalization processes in the West" (Kalberg, 1980: 1157), a diversity of rationalities is needed to 'master fragmented and disconnected realities' and prevent the possible dysfunctions resulting from a situation in which one rationality comes to dominate, or suppresses, other rationalities. This antagonism shines acutely through in Weber's work when he writes about the dominance of an impersonal formal rationality in spheres of social life that gained particular status in the industrialized era, and whose status has remained dominant to the present day (Kalberg, 1980; Weber, 1921/1946). In *The Machine Stops*, it is evident in the norm to be 'mechanical' versus Kuno's conviction that "Man is the measure" (p. 25), respectively.

Second, and implicit in our writing so far, formal rationality is not just another type of rationality; it is also understood as a historical accomplishment (Kalberg, 1980), the result of a process of formalization of substantive rationality. The ambiguity and radical perspectivism associated with substantive rationality pose the question how groups agree on an identifiable constellation of values that informs the direction of their rationalization? If there is no constellation of values that represents a persistent 'standard' for 'rational' decision-making, how is it that, nevertheless, in the words of Kalberg (1980: 1155) "small groups, organizations, institutions, political entities, cultures, and civilizations are . . . ordered in terms of specifiable value postulates"? Kalberg suggests that the ordering of individual preferences to specific ultimate values relies on socialization, but that supposes that these ultimate values are already a given. Others argued that people develop a shared value orientation through negotiation and debate, thus turning differences in values into workable propositions and actions that rest on compromise and creativity. Dewey (1916) called this process democracy, whereas Habermas

(1984) referred to it as deliberation. This process may be perverted, however, when one (or a limited set of) value orientation(s) is privileged over a broader orientation on values and then reified through formalization; that is, through programming it in the binary language of algorithms to allow a machine to make decisions. Formalization of a value orientation may entail the very demise of substantive rationality because it relegates the substantive to the margins. The net result is that formalization of substantive rationality may arrive at a point at which it never sought to arrive and which it, in fact, wanted to avoid with all might; that is, the transformation of substantive into formal rationality. It is at this instance that the tremendous historical influence of formal rationality becomes evident; the appeal it exercises in decision-making makes it reside and reign over substantive rationality, because any reification of value or value clusters (even the ones with moral connotations) means it is transformed into a calculus.

OMT AND THE RISE OF ALGORITHMIC DECISION-MAKING

Applied to our essay, while people are boundedly rational, they desire to create technology, machines and algorithms that are not thus constrained. The aim is to achieve a status of *un*bounded rationality, which according to some commentators is bound to culminate at the very moment of 'singularity,' the moment when AI-based algorithms will have become more 'intelligent' than human beings (Barrat, 2013). For these commentators, algorithms are a way of bringing rationality to the world of people, a rationality that would otherwise be unattainable. In the process, whichever substantive value orientation prevails, it is reified in formal rationality through Simon's procedure. That we may well abolish ourselves in that same process is a deplorably under-represented possibility in the current advocacy on algorithmic decision-making (for exceptions, see Barrat, 2013; Friedland, in press; Moore, 2018).

As indicated earlier in this section, some currents within OMT have challenged and resisted the impetus from Simon to bring substantive rationality under the reign of formal rationality. To begin with, and more generally, some organizational theorists are deeply suspicious of 'simplified' theorizing as it leaves out what matters most, "especially, context, uniqueness, process, and time" (Tsoukas, 2017: 134). A specific example is Mintzberg's (1994) work on strategic planning. For him, planning represents a "formalized procedure to produce articulated result, in the form of an integrated system of decisions" (p. 13). He added that planning itself may be the very cause of the 'costly misadventures' associated with it, rather than human factors as is expressed in a quote by Kurt Vonnegut Jr. in *Player Piano*:

> *'If it weren't for the people, the god-damn people,' said Finnerty, 'always getting tangled up in the machinery. If it weren't for them, the world would be an engineer's paradise.'* (Hackman & Oldham, 1980: iii)

Mintzberg's (1994) work is crucial in the context of our essay, not because of an all-out dismissal of formal rationality or strategic planning, but because of its dismissal of a sole reliance on formal rationality in strategic planning. He highlighted some of the assumptions of strategic planning, and argued that its formal rationality can be appropriate in decision-making as decision support and under the conditions of transparency and due recognition of its limitations and assumptions. That is, algorithmic support of decision-making may be helpful if (i) data are not incorrect, (ii) no irrelevant data are left out, (iii) data have future relevance, and (iv) the structure and parameters of the algorithm are validated for the decision situation. However, as previous studies show (for examples, see O'Neill, 2016; Redden & Brand, 2018), as well as recurrent media reports of AI-associated scandals, such assumptions are often violated. For example, regarding (i), initial reports on the cause of the two Boeing 737 MAX 8 crashes (i.e., Lion Air in 2018 and Ethiopian Airlines in 2019) suggest that the AI flight control system was

not only fed incorrect data from a dysfunctional sensor, but also that the software continually overrode the pilots' actions to stabilize the plane (Suhartono & Dahrul, 2018). Regarding (iv), Hutson (2018) reported on evidence suggesting that the outcomes of AI-based algorithms are not independent from their system settings (parameters), and thereby invalid. There is probably no escape from violating these assumptions, for example, because one can only know *after* the fact whether data were correct, that no irrelevant data were left out, or that the model was validated for the situation.

All this is to insinuate that, to the extent that organizational theory adopts the same assumptions of formal rationality in a manner that is unreflective of its associated limitation and problems, it can be considered as clearing the ground for the further adoption of algorithmic decision-making organizations and thereby as becoming complicit in the end of choice.

IMPLICATIONS OF ALGORITHMS AS SUPERCARRIERS OF FORMAL RATIONALITY FOR ORGANIZATIONAL DECISION-MAKING

Decision-making is, of course, central to organizations. Luhmann (2000), for instance, noted that the quintessence of decision-making is choice amongst alternatives. Thus, decision-making reduces uncertainty through the elimination of alternative options, but it simultaneously increases uncertainty through the implication that the choice might have been different; there is a degree of arbitrariness in the selection of one option over others, such that "decision-making takes on a certain air of randomness" (Brunsson & Brunsson, 2017: 44). As detailed above, Simon considered such randomness and arbitrariness intolerable for 'good decision-making'.

The mounting use of AI-based algorithms has the potential to change this state of affairs fundamentally. Such algorithms carry with them the promise of aligning decision-making with the ideal of rationality as embodied in Simon's work, due to their ability to analyze ever-bigger

amounts of data at ever-higher processing speeds. Yet, they are restricted to formal rationality. They are programmed, after all, to reach a particular goal (e.g., to drive this car from place A to place B, to suggest the next song that most likely appeals to the listener's preferences, or to identify and eliminate a particular terrorist cell) under pre-defined constraints (e.g., the priority order among the prevailing traffic rules, only songs without explicit content, or no collateral damage). Such constrained goal-orientation is performed through optimization made possible by the availability of highly advanced correlational, clustering and regression analyses and other techniques of pattern recognition. Therefore, *because* algorithms operate through a calculus that necessarily produces one outcome, they *preclude* the idea of choice. It should be noted that humans often experience the outcomes of algorithmic operations as if they are decisions and accordingly act upon them. This is because of the high status and legitimacy associated with the presumed rationality and objectivity of algorithmic procedures, which renders it difficult to contest and not follow their outcomes (Brunsson & Brunsson, 2017). We might say that algorithms produce 'calcucisions'[6] and thereby imply the imminent possibility of the end of choice through transforming or subsuming substantive under formal rationality.

However, there are serious limitations to AI-based algorithms that turn into liabilities for decision-making. Prominently among them is that they are not transparent. It is not clear, for example, what data an algorithm has found and used, how it has abstracted and generalized from them, and what are the consequent triggers in the algorithm's performance goals. Neither can we any longer understand how exactly the algorithm goes about in fulfilling its performance criteria; it is a 'black box' in its processing of data (see Al-Amoudi & Latsis, 2019). The calculus that informs its output is often even incomprehensible to their creators (Knight, 2017).

[6] A play of words, the contraction of 'calculation' and 'decision'.

This lack of transparency is a liability to decision-making in light of two vulnerabilities in algorithmic data processing: one relates to the quality of input data (input data should be correct and meaningful), and another to the possibility of automated decision-making 'running wild', for example, due to escalating cycles of response to self-reinforcing feedback.[7] The first vulnerability has been dubbed 'algorithmic selection error' (Andrews, 2018). It is based on the amplification of (hidden) bias in the data that feed the algorithm. There are many examples of incidents following from this vulnerability, including the UK Health Secretary blaming a computer algorithm for mistakes in cancer screening in 2018; the South Wales Police conceding that the facial technology system they had employed had thrown up thousands of 'false positives'; and Amnesty International criticizing the Metropolitan Police's Gang Violence Matrix database as being racially discriminatory (Andrews, 2018). So-called 'flash crashes' are examples of the second vulnerability. On 6 May, 2010, anomalies in high-frequency trading caused considerable fluctuation in the NYSE Dow Jones index. Over a 25 minutes period, 2bn shares, worth about $56bn, changed ownership, often at irrational prices (Bridle, 2018). The likely scenario for this case is that small fluctuations in automated economic valuations of stock prices reinforced each other, thus triggering an avalanche of trading orders. So, while algorithms follow strict rules based upon formal rationality in processing data to execute their goal function, they are not fail-safe. Only after their vulnerabilities have materialized in mistake, error, or worse, can they be detected and corrected, but substantive rationality is required for this to happen; the detection and correction of errors depends on an escape from the strictures of formal rationality.

A further consequence of the lack of transparency in algorithmic decision-making thus pertains to issues of responsibility and accountability. The imperative of responsibility and

[7] We forego a discussion of the many ways by which humans can use algorithms in immoral and/or illegal ways.

accountability holds that agents – whether humanly or artificially intelligent – should always be able to give reasons for actions when asked for explanation (Føllesdal, 1994). The point has been raised in the legal context about the nonlinearity and unpredictability of robots' actions (which often run on algorithms to process sensed inputs) and the challenges this creates to determine culpability (Froomkin, 2016). Likewise, fundamental functions of organizations, such as the exercise of control over processes (Ahrne & Brunsson, 2011), are at risk of escaping the direct supervision of human decision-makers. No longer being able to understand algorithmic decision-making (and their rationales along the way) that sustains and informs organizations, it becomes less likely that we can keep control over the outcomes to which organizations aspire. Forster's novel is a clear testament to this possibility. Deprived of the ability to decode the warning signs that the *Machine* was about to collapse, Kuno shouts at Vashti: "Cannot you see . . . that it is we that are dying, and that down here the only thing that really lives is the Machine?" (p. 33).

Kuno's words indicate a third implication of formal rationality having crowded out substantive rationality; it may imply that we become "machines ourselves" (Friedland, in press), that AI-based algorithms induce us to a condition of 'learned helplessness' (Moore, 2018). This condition is implied in Friedland's (in press) cautions about the perils of technological convenience, which can lead individuals and society to become 'disengaged' through: (i) increased passivity (i.e., accepting assistance for task completion), (ii) emotional detachment (through diminished participation), (iii), decreased agency (i.e., reduced power to make own decisions because of lack of awareness of alternative choices), (iv) decreased responsibility (i.e., ceding control over a decision-making process implies loss of accountability for consequences), (v) increased ignorance (i.e., our needs/wants are translated into algorithmic shorthand or mechanical processes that eventually may function differently than we would ourselves), and (vi)

deskilling (i.e., dependency on algorithms for routine task completion can numb trained skills needed to interact with the social and physical world around us). In a way, it seems this learned helplessness is both a result of, and a condition for, the functioning of technology-based totalitarian systems; it renders individuals within such systems more readily controllable. We do wonder, however, whether there is a willingness to listen to such concerns in the various countries and organizations that wish to be at the forefront of the technology. Just as the *modus operandi* of the *Machine* is strictly designed around the assumptions of formal rationality at the expense of substantive rationality, so do ethical issues hardly seem to matter when major players in the domain of AI and algorithms focus first and foremost on privacy invasion for economic gain and/or behavioral control of citizens (Mader, 2018).[8]

The quest for efficiency (and thus economic gain) is a central postulate in Simon's adoption of formal rationality. When brought into our current capitalist context, it seems that this quest for efficiency is reified in organizational activity in two forms. First, in an *economic sense*, it is reified in the search for constantly high yields on capital, and second, in a *technological sense*, in the search for constantly higher computational power and automated (i.e., algorithmic) decision-making that supposedly neutralize the 'damaging' potential of human limitations and biases in maximizing efficiency. Taken together, capital and technology form the basis for rational organizational action (Cabantous & Gond, 2011) and 'digital capitalism' (Fuchs & Chandler, 2019; Schiller, 1999), whereby enhanced decision-making seeks to emulate Simon's ideal of unbounded rational decision making. It does so by way of algorithms that potentially bring about the end of choice and the end of clearly identifiable accountability.

[8] These developments are especially salient in the USA and China. As of yet, European countries hardly play a role in them judged by levels of investment in the technology, number of relevant patents submitted, and submissions to leading conferences (Mader, 2018).

The end of choice is a result of decision-making having become automated, the outcome of an algorithmic process: the algorithm optimizes toward a given value. But the sharp contraction of decision 'alternatives' that follows impoverishes our ability to use a broader range of value orientations. Simon's ideal of unbounded rational decision-making constrains us, in our deliberations, to explore, weigh, consider and evaluate several decision alternatives along multiple substantive rationalities (cf. Habermas and Dewey, as mentioned before), including issues of morality and concern for others. We can further unpack the need for considering a broader range of rationalities in decision-making with reference to Boden (1978). She follows Weber (1915/1946) in arguing that reasoning (including moral reasoning) is prior to action. Therefore, the more one is capable of 'flexible reasoning',[9] the greater the freedom of action available. Boden adds that free action is self-determined, in reference to a person's image of its ideal (or actual) moral self (i.e., do I want to be the kind of person who does that?). Exercising moral responsibility is a complex task that requires knowledge of moral principles, the world to which they are applied, and oneself, both as a person now and as how one wishes to be. The combination of these domains of knowledge is key to moral decision-making and the ability to justify one's actions (Solomon, 1993).

In Forster's story, the *Machine* determined these knowledges, leaving humans vulnerable to the approaching apocalypse. Knowledge of the world was suppressed through the rejection of primary or first-hand experience. "Beware of first-hand experiences", exclaimed an advanced lecturer. "First-hand ideas do not really exist. They are but physical impressions produced by love and fear, and on this gross foundation, who could erect a philosophy?" (p. 40). Continuing

[9] For as long as substantive rationality implies a value rationality in which 'value' is *not* reified (cf. Weber's notion of radical perspectivism), Boden's flexible reasoning is closely related to pragmatist conceptualizations of deliberation and inquiry (e.g., Dewey, 1922).

to dismiss primary ideas, the lecturer claimed that "let your ideas be second-hand, and if possible tenth-hand, for then they will be far removed from that disturbing element – direct observation" (p. 40). Knowledge of the person one wishes to be was in reference to one overarching 'moral' principle: to be 'mechanical'. Being mechanical was the "proper thing [to do]" (p. 25). How to be mechanical was specified in the *Book of the Machine,* with "instructions against every possible contingency" (p. 8).

If moral responsibility and moral decision-making are associated with the reasoned choice between alternatives (Boden, 1978), then the 'advanced' civilization of the *Machine* had suppressed morality through the formalization of substantive rationality. It did so by reducing the number of permissible alternatives for action to those prescribed by/described in the *Book*. Prophetically, the lecturer exclaimed: "In time . . . there will come a generation that has gone beyond facts, beyond impressions, a generation absolutely colourless, a generation seraphically free from taint of personality" (p. 41). To create a generation along these lines also implies a generation of individuals unable to be reflexive, defined as the "regular exercise of the mental ability, shared by all normal people, to consider themselves in relation to their (social) contexts and vice versa" in order to have an informed "internal conversation" (Archer, 2007: 65). But not only that; lacking reflexivity also created a generation incapable of disobedience. For Fromm (1981/2010), the ability to disobey signifies a critical juncture between the future of mankind and the end of civilization. Thus, apparently all people in Forster's story, apart from Kuno, had already lost the abilities to be reflexive and disobey, having internalized the doctrine of the *Book*.

Since formal rationality as espoused in algorithmic decision-making lacks sensitivity for moral concerns (unless the maximization of some utility or the absolute interdiction of precisely circumscribed actions are defined as moral concerns), it may just as well imply an inability to

question the objectivity of the calculus. Yet, it is such skepsis that feeds deliberation and creates the possibility to act as moral decision-makers. The suppression and formalization of substantive rationality eventually leads to the end of choice with its various implications. Having discussed these implications of algorithms as supercarriers of formal rationality for organizational decision-making, we can now draw this essay to a close.

DISCUSSION

Inspired by *The Machine Stops*, our essay reflected on and problematized the assumptions of rationality underlying algorithms and algorithmic decision-making, the role of OMT in the rapid rise of algorithmic decision-making, and (iii) the implications of algorithmic decision-making for organizations. Along the way, we discussed Weberian notions of formal and substantive rationality, respectively, to make the case of algorithms as supercarriers of formal rationality. We went on to show that some strands within organization theory feature a firm commitment to formal rationality (e.g., Simon, 1947/1997), which, in turn, helps account for the remarkable rise of algorithmic decision-making (i.e., because it justifies the turn to formal rationality to enhance decision-making). We also briefly contrasted Simon's with Mintzberg work (e.g., 1994) to show that the latter defends the significance of substantive rationality alongside formal rationality. Fourth, we considered the implications of algorithms as supercarriers of formal rationality for organizational decision-making, especially their potential to represent enhanced decision-making capability. In what follows, we offer a discussion of theoretical and practical ramifications of algorithmic decision-making for management and organization studies.

From a theoretical perspective, assumptions around the theoretical parameters of formal and substantive rationality do not remain assumptions alone; whichever type of rationality is dominant, it crystalizes into conscious regularities of action over time (Kalberg, 1980). Thus, the

notion of parity between the two types of rationality is a hypothetical construct. As we have shown at length, formal rationality exercises a much more powerful appeal than substantive rationality. We proposed that the pursuit of 'perfect' or 'augmented' rationality may be a major driver for the advancement of algorithmic decision-making. In fact, as supercarriers of formal rationality, algorithms do not only rely on formal rationality, but also reinforce formal rationality in organizational decision-making.

However, 'rationality' thus conceived only professes to offer a superior version of rationality; in reality, it turns a blind eye on the fragmented and multifaceted realities that each version of rationality helps to master. Algorithmic decision-making, infused as it with formal rationality, may imply the end of choice, not only through the suppression of substantive rationality, but also though the transformation of substantive rationality into formal rationality through formalization. The formalization of substantive rationality is like the proverbial wolf in sheep skin. It sets in motion a process driven by best intentions (e.g., a value orientation like fairness principles to stop disadvantaging minorities in the population), but in doing so it is transforms from a heuristic that informs what one can do best in a particular situation into a rule that invariantly applies to all situations. The latter has been referred to as 'restrictive ethics' (Kjonstad & Willmott, 1995). This transformation progressively numbs our human capacities for reflexivity and disobedience that can be applied to more positive sense of ethics, one that can inspire and empower (see Kjonstad & Willmott, 1995) us normatively toward a future that is worth living.

The formalization of substantive rationality is what is at stake in attempts to develop 'good AI'. If we were to teach AI-based algorithms human morality, we would first have to define – in universal fashion – what this morality is and do so in a way that algorithms can

process it (i.e., having clear and quantifiable rules about what morality is, see Hao, 2019; Polonski & Zavalishina, 2017). In *the Machine*, this is reflected in the moral principle to be 'mechanical', as indicated before. This entails, however, a reduction of the essence of moral behavior and ethics – understood as a problem of dealing with *aporetic,* or puzzling, situations to which there are no simple and straightforward answers and that cannot be captured unequivocally and unambiguously in an explicit rule (e.g., Dewey, 1930) – into the following of a formalized rule. Yet, as James (1891: 330) famously stated, "there is no such thing possible as an ethical philosophy dogmatically made up in advance."

Despite these profound concerns, AI-based algorithms appear as the ultimate means through which the apex of rationality can be reached; a state in which bounded rationality is upended and rationality has become *unbounded* for the sake of efficiency and control. As Habermas (1970/1976: 340) reminds us, rationality is a value in and of itself, such that indeed the critique of ideologies' preparation for rational conduct "recommends rationality as the preferred – if not exclusive – means for the realization of values, because it guarantees the 'efficiency' or 'economy' of procedures". At some future moment, there might be the ultimate test to assess whether or not a state of unbounded rationality has been reached. That is, when a court of law rules that a medical doctor has caused the death of a patient by *not* following the recommendations of an AI-based diagnostic algorithm (see Al-Amoudi & Latsis, 2019). This is the moment when the abolition of human decision-making is initiated.

In our analysis, the imminent possibility of enhanced organizational decision-making is rooted in the quest for perfectly rational decision-making that presumes to have cast off, and liberated itself from, the frailties and biases in human decision-making. It has been our contention that the pursuit of the perfect decision, the 'one single best' as algorithmically

determined, progressively encroaches upon us and regresses our ability to question and disobey. Therefore, there are fundamental questions about the accountability for such decision-making and for whose benefits and purposes such decision-making is pursued. To reinvoke Kuno's words from the opening of our essay, "the Machine develops – but not on our lines. The Machine proceeds – but not to our goal." Thus, among the many possible lessons to be learned from *The Machine Stops* is the premonition that there are risks involved if 'data' and 'automated decision-making' become *more* credible and obtain *more* authority than human experience. Seen in this way, perfect rationality, in the narrow sense of formal rationality based on universal rules, would be a disaster. Berlin (1947/2013: 15) warns that:

> *In a society in which all the same goals are universally accepted, problems can be only of means, all soluble by technological methods. That is a society in which the inner life of man, the moral and spiritual and aesthetic imagination, no longer speaks at all.*

In a society thus conceived, there will be no deviance, no disobedience, no opportunity to learn from exploring something new or unexpected, and no scope for innovation, whereas the opportunity to learn from mistake will be severely reduced. This society would deny us the possibility to make sense of social life through any other device but formal rationality. The development toward it would progressively reduce our ability to act spontaneously and to be creative, reflexive and informed in judging how we engage with the world (Al-Amoudi, 2017). Ultimately, it may undermine our ability to influence courses of action.

Lacking reflexivity, and being faced with the end of choice, it is no wonder that Forster's story recounts the distasteful practice of healthier and more athletic infants being destroyed after birth. Recall that in ancient Greece the Spartans exercised a similar gruesome practice, only that they discarded children with physical defects. In Forster's story, the attitudes toward life and social interaction had increasingly valued the 'bringing of things to people'. Under the superior

rule and technology of the *Machine*, practices such as social gatherings, explorations, and physical activity had become redundant and were, therefore, discouraged. Once these attitudes were accepted unquestionably, it is no stretch to argue that it is unethical to allow an athletic child to be raised in such an environment; it would never be able to make use of its inherited physical abilities. Such a child, the logic goes in the story, would be confined to a life of frustration and unhappiness. In fact, the child would be better off not to live – not to speak of society at large, having ridden itself of a potential insurgent.

Perhaps this particular practice may not materialize in our generation or in some societies, but a fundamental shift in attitudes and practices may emerge over time as the cumulative effect of small changes here and there, barely noticeable as they come in small doses and with small steps. An emerging real-world example of this very development can be observed in China's 'social credit' system. This system is designed to monitor and shape the behavior of citizens and businesses (Munro, 2018). It does so through a range of big data and algorithmic processes that bestow individuals and businesses a social credit score based upon their social, political and economic behavior. A low score may result in being blacklisted or banned from accessing services, such as travel, health, or other amenities, whereas a high score bestows privileges (Helbing et al., 2017). Indeed, recent news indicated that China has already banned more than 20 million people from travelling by train or plane (Wurzel, 2019). The system also affects foreign companies, who need a license to operate within China. These companies are added to the social credit system and have to comply; otherwise, they cannot operate in China. As reported in the news, Chinese Authorities used their social credit system to pressurize US and Australian airlines to using China's preferred terminology when referring to Taiwan and Hong

Kong (Munro, 2018). Consequently, the system dictates an economic reality for these companies by deciding if they can trade in China or not.

There are, therefore, also significant practical implications that follow from our analysis. Although Weber and others were acutely aware that formal rationality represents a set of theoretical assumptions that, in a way, would delight an engineer's heart, or for that matter the heart of a bureaucrat, or an administrator – because formal rationality implies decisions are made 'without regard to persons' – these assumptions fare poorly once applied to the social world. Weber was explicit that social life solely based on formal rationality is impossible. Kalberg, in his introduction to the *Protestant Ethic* (in Weber, 1920/2002: xv), acts as a *porte parole* for Weber in rhetorically pondering:

> *What 'type of person' will inhabit this new universe* [that of an industrialized society]*? How, amidst the overwhelmingly material and pragmatic character of everyday life in industrialized societies, will persons be able to orient their lives to ethical values? . . . will not the sheer instrumental-rational calculations typical of the modern capitalist economy push aside all ethical values?*

This is exactly what we implied earlier when we cautioned against a premature dismissal of our normative commitment. It is simply not only a question of what ought to be and what not, but a question of what kind of rational assumptions are required to make social life possible. Not all that is technically possible is ethically justifiable; not just because of any insistence on particular values that can be contested, but because of 'rational' assumptions that pose a challenged to social life, as Forster's story underlines.

The Machine Stops is thus a hortatory story of the consequences for social systems when the sources dry up that nourish substantive rationality within communities, when the ability is suppressed to gain first-hand experience in a concrete social situation involving individuals capable of reflexivity and disobedience, when no one any longer brings in different ethical

values that must be negotiated. Vallor (2016: 2, original emphasis) has a point when asserting that "ethics and technology are connected because technologies invite or *afford* specific patterns of thought, behavior, and valuing; they open up new possibilities for human action and foreclose or obscure others." Yet, the possibility and need for developing 'technomoral' choices and wisdom (Vallor, 2016) depend on keeping open the possibility of deliberation based on substantive rationality. The invitation and affordance of technology in the form of algorithmic decision-making – Forster's *Machine* – is to follow its rational assumptions; it thus carries with it a serious risk of formal rationality suppressing substantive rationality and its invitation and affordance may in the end be little more than a narrow, one-dimensional track. The sensualist Kuno, a curious character thirsting for first-hand experiences in a sensory-deprived environment, resisted this straight jacking of social life. In a way, Kuno was a tragic "historical subject . . . which embodied Western civilization's highest ideal: the autonomous and free individual whose actions were given continuity by their reference to ultimate values" (Kalberg, 1980: 1176). We can only speculate what might have happened to the civilization of which he and his mother were part, if *more* people had his ability for ethical reflexivity. Would the *Machine* have ever reached its status as a deity to be worshipped? Would the civilization have collapsed?

We reserve the concluding thought of this review essay for Forster. At the end of the story, Vashti notes, with resignation, that "some fool will start the Machine again, to-morrow", to which Kuno replies "Never . . . Humanity has learnt its lesson" (p. 55). We are not convinced that the right lessons are being learned at present (given the examples we offered), but neither do we subscribe to the idea of a pending apocalypse. Thus, our essay did not aim to demonize formal rationality and technology *in toto*, but to show the limits and dangers of its unreflexive use. We believe that there is vast space in between these scenarios for careful examination and

profound debate amongst users, pressure groups, policy makers, tech companies, legislators, financial institutions, and scholars in the natural and social sciences, including organizational theorists. The absence of examination and debate in the leading management journals is an anomaly that requires urgent attention. It is our sincere hope that this essay helps redress this imbalance, and sparks future theorizing, empirical advances, and perhaps even social movements around the perils and promises of algorithmic decision-making.

REFERENCES

Ahrne, G., & Brunsson, N. 2011. Organization outside organizations: The significance of partial organization. ***Organization***, 18(1): 83-104.

Al-Amoudi, I. 2017. Reflexivity in a just morphogenic society: A sociological contribution to political philosophy. In M. S. Archer (Ed.), ***Morphogenesis and human flourishing***: 63-91. Cham: Springer International Publishing.

Al-Amoudi, I., & Latsis, J. 2019. Anormative black boxes: Artificial intelligence and health policy. In I. Al-Amoudi, & E. Lazega (Eds.), ***Post-human institutions and organizations: Confronting the matrix***. London: Routledge.

Amabile, T. 2019. Guidepost: Creativity, artificial intelligence, and a world of surprises. ***Academy of Management Discoveries***.

Andrews, L. 2018. Public administration, public leadership and the construction of public value in the age of the algorithm and 'big data'. ***Public Administration***.

Archer, M. S. 2007. ***Making our way through the world: Human reflexivity and social mobility***. Cambridge, UK: Cambridge University Pres.

Arendt, H. 1958. ***The origins of totalitarianism***. New York: Meredian Books.

Barrat, J. 2013. ***Our final invention: Artificial intelligence and the end of the human era***. New York: Thomas Dunne.

Baynes, C. 2018. Chinese police to use facial recognition technology to send jaywalkers instant fines by text. (Retrieved from https://www.independent.co.uk/news/world/asia/china-police-facial-recognition-technology-ai-jaywalkers-fines-text-wechat-weibo-cctv-a8279531.html on 19 November 2018).

Berlin, I. 1947/2013. ***The crooked timber of humanity: Chapters in the history of ideas***. Princeton, NJ: Princeton University Press.

Boden, M. A. 1978. Human values in a mechanistic universe. In G. Vesey (Ed.), ***Human Values***: 135-171. Sussex: Harvester Press.

Bostrom, N. 2014. ***Superintelligence: Paths, dangers, strategies***. Oxford, UK: Oxford University Press.

Bridle, J. 2018. Rise of the machines: Has technology evolved beyond our control? . (Retrieved from https://www.theguardian.com/books/2018/jun/15/rise-of-the-machines-has-technology-evolved-beyond-our-control- on 26 June 2018).

Brunsson, K., & Brunsson, N. 2017. ***Decisions: The complexities of individual and organizational decision-making***. Cheltenham, UK: Edward Elgar.

Cabantous, L., & Gond, J. P. 2011. Rational decision making as performative praxis: Explaining rationality's éternel retour. ***Organization Science***, 22(3): 573-586.

Clegg, S., Kornberger, M., & Pitsis, T. 2008. ***Managing and organizations***. London: Sage.

Copeland, R., & Hope, B. 2016. The world's largest hedge fund is building an algorithmic model from its employees' brains. (Retrieved from https://www.wsj.com/articles/the-worlds-largest-hedge-fund-is-building-an-algorithmic-model-of-its-founders-brain-1482423694 on 31 October 2018).

De Vynck, G., & Grant, N. 2018. Google wants to answer the questions you haven't even asked yet. (Retrieved from https://gadgets.ndtv.com/internet/features/google-wants-to-answer-the-questions-you-havent-even-asked-yet-1921664 on 19 November 2018).

Dewey, J. 1916. ***Democracy and education***. New York: Macmillan.

Dewey, J. 1922. ***Human nature and conduct***. New York, NY: Henry Holt and Company.

Dewey, J. 1930. Three independent factors in morals. In J. A. Boydston (Ed.), ***The later works, 1925-1953, Vol. 5*** 279-288. Carbondale, IL: Southern Illinois University Press.

Dewey, J. 1939. ***Theory of valuation***. Chicago, IL: Chicago University Press.

Economist. 2018. How an algorithm may decide your career. ***The Economist*** (Retrieved from https://www.economist.com/business/2018/06/21/how-an-algorithm-may-decide-your-career on 9 of January 2019).

Elster, J. 1994. Nature and scope of rational-choice explanation. In M. Martin, & L. C. McIntyre (Eds.), ***Readings in the philosophy of social science***: 311-322. Cambridge, MA: MIT Press.

Føllesdal, D. 1994. The status of rationality assumptions in interpretation and in the explanation of action. In M. Martin, & L. C. McIntyre (Eds.), ***Readings in the philosophy of social science***: 299-310. Cambridge, MA: MIT Press.

Forster, E. M. 1928/2011. ***The Machine Stops***. London: Pengion. Originally published in E. M. Forster's '***The eternal moment and other stories***' by Harcout Brace & Company.

Friedland, J. in press. Activating moral agency by design: A model for ethical AI development. ***MIT Sloan Management Review***.

Fromm, E. 1981/2010. ***On disobedience: Why freedom means saying "no" to power***. New York: HarperCollins.

Froomkin, A. M. 2016. Introduction. In R. Calo, & I. Kerr (Eds.), ***Robot law***: x-xxiii: Edward Elgar Publishing.

Fuchs, C., & Chandler, D. 2019. Introduction. In D. Chandler, & C. Fuchs (Eds.), ***Digital objects, digital subjects: Interdisciplinary perspectives on capitalism, labour and politics in the age of big data***: 1-20. London, UK: University of Westminster Press.

Habermas, J. 1970/1976. Theory and practice in a scientific civilization. In P. Connerton (Ed.), ***Critical Sociology***: 330-347.

Habermas, J. 1984. ***The theory of communicative action (Volume 1): Reason and the rationalization of society***. Boston, MA: Beacon Press.

Hackman, J. R., & Oldham, G. R. 1980. ***Work redesign***. Reading, MA: Addison Wesley.

Hao, K. 2019. Giving algorithms a sense of uncertainty could make them more ethical. ***MIT Technology Review***(Retrieved from https://www.technologyreview.com/s/612764/giving-algorithms-a-sense-of-uncertainty-could-make-them-more-ethical/. on 21 March 2019).

Helbing, D., Frey, B. S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., van den Hoven, J., Zicari, R. V., & Zwitter, A. 2017. Will democracy survive big data and artificial

intelligence? ***Scientific American*** (Retrieved from https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/ on 23 March 2017).
Hendershott, T., Jones, C. M., & Menkveld, A. J. 2011. Does algorithmic trading improve liquidity? ***The Journal of Finance***, 66(1): 1-33.
Hutson, M. 2018. Artificial intelligence faces reproducibility crisis. ***Science***, 359(6377): 725-726.
James, W. 1891. The moral philosopher and the moral life. ***International Journal of Ethics***, 1(3): 330-354.
Kalberg, S. 1980. Max Weber's types of rationality: Cornerstones for the analysis of rationalization processes in history. ***American Journal of Sociology***, 85(5): 1145-1179.
Kjonstad, B., & Willmott, H. 1995. Business ethics: Restrictive or empowering? ***Journal of Business Ethics***, 14(6): 445-464.
Knight, W. 2017. The dark secret at the heart of AI. ***MIT Technology Review***(Retrieved from https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/ on 18 April 2017).
Libert, B., Beck, M., & Bonchek, M. 2017. AI in the boardroom: The next realm of corporate governance. (Retrieved from https://sloanreview.mit.edu/article/ai-in-the-boardroom-the-next-realm-of-corporate-governance/ on 21 February 2019).
Lindebaum, D., & Courpasson, D. 2019. Becoming the next Charlie Parker: Rewriting the role of passions in bureaucracies with whiplash ***Academy of Management Review***, 44(1): 227–239.
Luhmann, N. 2000. ***Organisation und Entscheidung***. Wiesbaden: VS Verlag.
Mader, F. 2018. Künstliche Intelligenz: Wie Deutschland seinen Wohlstand aufs Spiel setzt [AI: How Germany jeopardizes its prosperity], ***Report Muenchen (Bayerischer Rundfunk)***. Retrieved from https://www.ardmediathek.de/tv/report-M%C3%9CNCHEN/K%C3%BCnstliche-Intelligenz-Wie-Deutschland/Das-Erste/Video?bcastId=431936&documentId=56646544 on 27 March 2019.
Mintzberg, H. 1994. Rethinking strategic planning part I: Pitfalls and fallacies. ***Long Range Planning***, 27(3): 12-21.
Moore, P. 2018. Book review: Artificial intelligence: What everyone needs to know. ***Organization Studies***.
Munro, K. 2018. China's social credit system 'could interfere in other nations' sovereignty' ***Guardian, 27 June***.
O'Neill, C. 2016. ***Weapons of math destruction: How big data increases inequality and threatens democracy***. New York: Crown Publishers.
Parkinson, H. J. 2019. AI can write just like me. Brace for the robot apocalypse. ***Guardian***(Retrieved from https://www.theguardian.com/commentisfree/2019/feb/15/ai-write-robot-openai-gpt2-elon-musk on 21 February 2019).
Pasquale, F. 2015. ***The black box society: The secret algorithms that control money and information***. Cambridge, MA: Harvard University Press.
Polonski, V., & Zavalishina, J. 2017. 3 ways to teach robots right from wrong. (Retrieved from https://www.weforum.org/agenda/2017/11/3-ways-to-build-more-moral-robots on 21 March 2019).
Redden, J., & Brand, J. 2018. Data harm record. (Retrieved from https://datajusticelab.org/data-harm-record/ on 19 November 2018).

Schiller, D. 1999. ***Digital capitalism. Networking the global market system***. Cambridge, MA: MIT Press.

Simon, H. 1947/1997. ***Administrative behavior: A study of decision-making processes in administrative organizations (4th edition)***. New York: Free Press.

Solomon, R. 1993. ***Ethics: A short introduction***. Dubuque, IA: Brown & Benchmark.

Suddaby, R. 2014. Editor's comments: Why theory? ***Academy of Management Review***, 39(4): 407-411.

Suhartono, H., & Dahrul, F. 2018. Lax safety, faulty systems cited in Indonesian crash of jet. (Retrieved from https://www.bloomberg.com/news/articles/2018-11-28/lion-air-crash-probe-findings-focus-on-carrier-s-safety-culture on 1 December 2018).

Tarnoff, B. 2018. Weaponised AI is coming. Are algorithmic forever wars our future? ***Guardian***(Retrieved from https://www.theguardian.com/commentisfree/2018/oct/11/war-jedi-algorithmic-warfare-us-military on 19 November 2018).

Tsoukas, H. 2017. Don't simplify, complexify: From disjunctive to conjunctive theorizing in organization and management studies. ***Journal of Management Studies***, 54(2): 132-153.

Tutt, A. 2016. An FDA for algorithms? ***Administrative Law Review***, 69 (Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2747994 on 25 June 2018).

Vallor, S. 2016. ***Technology and the virtues: A philosophical guide to a future worth wanting***. Oxford, UK: Oxford University Press.

Weber, M. 1915/1946. Religious rejections of the world and their directions. In H. Gerth, & C. Wright (Eds.), ***From Max Weber: Essays in sociology***: 323-359. New York: Oxford University Press.

Weber, M. 1920/2002. ***The protestant ethic and the spirit of capitalism*** (translated by S. Kalberg). Los Angeles: Roxbury Publishing Company.

Weber, M. 1921/1946. Science as a vocation. In H. Gerth, & C. W. Mills (Eds.), ***From Max Weber***. New York: Oxford University Press.

Weber, M. 1922/1978. ***Economy and society: An outline of interpretative sociology***. Berkeley, CA: University of California Press.

Wurzel, S. 2019. Reiseverbot für Millionen Menschen (Travel ban for millions of people). (Retrieved from https://www.tagesschau.de/ausland/sozialkredit-ranking-china-101.html on 8 March 2019).