Check for updates

Received: 13 November 2020 | Revised: 10 September 2021 | Accepted: 19 November 2021

DOI: 10.1111/1748-8583.12430

Human Resource Management Journal WILEY

RESEARCH ARTICLE

Scientism as illusio in HR algorithms: Towards a framework for algorithmic hygiene for bias proofing

Joana Vassilopoulou[1] | Olivia Kyriakidou[2] | Mustafa F. Özbilgin[1,3,4] | Dimitria Groutsis[5]

[1]Brunel University London, Brunel Business School, Uxbridge, UK

[2]Deree – The American College of Greece, Graduate School, Department of Psychology, Athens, Greece

[3]Université Paris Dauphine, Paris, France

[4]Koç University, Istanbul, Turkey

[5]University of Sydney, University of Sydney Business School, Work and Organisational Studies, Sydney, Australia

Correspondence
Joana Vassilopoulou, Brunel Business School, Brunel University London, Uxbridge UB8 3PH, UK.
Email: Joana.Vassilopoulou@brunel.ac.uk

Abstract

Human Resource (HR) algorithms are now widely used for decision making in the field of HR. In this paper, we examine how biases may become entrenched in HR algorithms, which are often designed without consultation with HR specialists, assumed to operate with scientific objectivity and often viewed as instruments beyond scrutiny. Using three orienting concepts such as scientism, illusio and rationales, we demonstrate why and how biases of HR algorithms go unchecked and in turn may perpetuate the biases in HR systems and consequent HR decisions. Based on a narrative review, we examine bias in HR algorithms; provide a methodology for algorithmic hygiene for HR professionals.

KEYWORDS
artificial intelligence, HR algorithmic bias, illusio, performativity, scientism, sociomateriality

1 | INTRODUCTION

Algorithmic decision-making has become an increasingly important consideration for Human Resource (HR) professionals and organisations. Algorithms are defined as "precise recipes that specify the exact sequence of steps required to solve a problem" (MacCormick, 2012: 2). They are used by HR professionals to take advantage of the huge volume of data generated in contemporary work life to augment various HR decisions. Some scholars have emphasised the potential of algorithms to make better predictions in, for instance, facilitating value creation (Mayer-Schönberger &

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. Human Resource Management Journal published by John Wiley & Sons Ltd.

Practitioner notes

What is currently known:

- HR algorithms are now widely used, and they are often assumed to be bias free, however, examples of their usage have emerged proving the opposite.
- There is often a lack of consultation with HR professionals in design of HR algorithms with the assumption that they are bias free
- Algorithmic hygiene is recognised as a way forward for ethical and responsible development of HR algorithms

What this paper adds:

- This paper demonstrates that HR algorithms have an allure, a pseudo-scientific quality, which prevents them from being scrutinised for biases.
- The paper scrutinises this allure as the illusio of scientism, by which HR professionals may show blind faith in the HR algorithms falsely assuming that they are objective, scientific and therefore bias free tools.
- The paper shows that these assumptions of bias free algorithms do not hold true.

Study findings for practitioners

- HR practitioners need to break away from this illusio of scientism in order to bias proof HR algorithms.
- The paper provides a step-by-step methodology for algorithmic hygiene for HR professionals.
- The paper shows the challenges running algorithmic hygiene for HR professionals, who should be included in the design of HR algorithms.

Cukier, 2013), automating repetitive and simple administrative tasks (Langer et al., 2019), saving costs and increasing productivity (McDonald et al., 2017) and transforming organisational culture (Schildt, 2020).

While the business benefits are apparent, it is perhaps unsurprising that the use of algorithms in Human Resource Management (HRM) operations, processes, and practices (Cheng & Hackett, 2021) has also come under increasing scrutiny by critical studies within the field of HRM (e.g., Hmoud & Laszlo, 2019; Leicht-Deobald et al., 2019; Ong, 2019). Critical research has not only considered the lack of regulatory measures (Ajunwa, 2020) or "good" employment data (Citron & Pasquale, 2014), but also the implications of algorithmic decision-making on employee control, surveillance, ethics, and discrimination, raising questions around the governance of such phenomena (Ajunwa, 2020; Mittelstadt et al., 2016; Parry et al., 2016). Moreover, it has highlighted how human biases can be inscribed into the code of the HRM algorithms embedding and sustaining inequalities while assuming a veneer of objectivity (Raghavan et al., 2020). The use of algorithmic pre-employment assessment exemplifies such inscription. For example, basing algorithms on historical employment data where men primarily hold management roles, may result in the conclusion that women do not and will not pursue such roles. The result is the exclusion of women from the job post. Biased historical data therefore perpetuates an ongoing gender bias (Devlin, 2017).

On the back of such evidence, critical scholars have called for an understanding of the processes through which HR algorithms may mask inequality and discrimination, replicate social and organisational inequalities and in some instances even amplify human bias (Kellogg et al., 2020; Köchling et al., 2021; Leicht-Deobald et al., 2019). Drawing on Bourdieu's notion of 'illusio' (Bourdieu & Wacquant, 1992), the theory of sociomateriality (Orlikowski, 2007), and performativity (Butler, 1990, 1993), we argue that algorithmic decision-making in HRM, whilst attractive, demands reflexive and flexible reasoning. Accordingly, the aim of the paper is to develop a reflexive and critical inquiry into the assumed objectivity and bias-free antecedents of algorithmic decision-making in HR, based on a narrative review of the literature, underscored by a theoretically informed orientation towards algorithmic implementation (Hovorka & Peter, 2018; Lepri et al., 2018). Throughout the paper, algorithmic decision-making is framed not just as a process that could be enhanced with "better" data, but rather, as a process that involves many actors embedded in subjective

perceptions of scientism (Haack, 2011; Kaufman, 2020). Such a conceptualisation advances the opportunities found in responsible approaches to algorithmic decision-making and governance pathways in HR. The latter are presented in the closing section where we place a social equity model at the centre of our analysis, generating a governance pathway which challenges algorithmic bias.

The paper begins with the presentation of our methodology and the review of relevant literature on the use of algorithmic decision-making in HRM outlining five identified ways through which HR algorithms could influence the perpetuation of inequalities in organisations. The discussion then proceeds to focus on scientism as the main reason for the unreflective implementation of HR algorithms. Scientism is understood as the unfounded prioritisation of scientific logic turning algorithms into "carriers of rationality" (Cabantous & Gond, 2011) resulting in assumed "correct" and "unbiased" decisions. Following, the paper turns to explore how scientism is attributed to algorithms through the notions of a theoretical tripod that includes illusio, sociomateriality, and performativity. Guided by Jonsen and Ozbilgin's (2014) maturity model for organisational interventions on diversity, we conclude with recommendations to address the lack of governance in the process of algorithmic decision making, which has unintended consequences for HR professionals and workers alike. As such, we make a practical contribution, developing a methodology for algorithmic hygience which can be adapted by HR professionals.

2 | METHOD

Informed by a narrative review, the study undertook a process of organising and synthesising otherwise disparate concepts from multiple fields to shed light on identified gaps and a pathway forward to address a specific problem (Hodgkinson & Ford, 2014; Jelf, 1999). The key problem identified in this paper is that there has been a neglect of possible biases that the growing use of algorithms may bring to the practice of HR in organisations. To uncover the multi-layered terrain in which the problem resides, we draw on the concepts of illusio, scientism, performativity and sociomateriality. For the purposes of our research paper, the narrative approach allows us to fulfil three important elements.

First, it allows us to explain and understand the qualitative nuances underscoring our engagement with a critical investigation of digital inclusion. Notably, rather than providing a measure of the blind spot effects of algorithmic decision making, a narrative review allows us to understand the multi-faceted nature of this phenomenon. Guided by this we undertook a deep reading and review of key publications relevant to our critical investigation of bias in HR algorithms, in the business and management, HRM, organisational behaviour, sociology, information systems (IS), social psychology and critical diversity fields (Jones & Gatrell, 2014). To this end, we conducted manuscript searches for the keywords: algorithm, bias, discrimination, inequality, diversity, HRM, OB, and IS separately in the Web of Science, which generated over a million papers for each keyword. Then we examined combined outcomes for the keywords: algorithm and bias/discrimination/diversity, which generated more than 10,000 papers for each instance. Finally, we ran a search for algorithm, bias, and HR together, which generated 144 papers. Of these papers, we eliminated 84 papers, following a reading of abstracts which deemed the papers were beyond the focus of our study outside the scope of the study. We read the remaining 60 papers with adequate numbers for each thread (approximately 20) in our line of inquiry. Guided by our theoretical framework and following, our team discussion we identified papers which fell outside the scope of our study, due to an apparent lack of relevance to the central theme of algorithmic bias in HR in content, despite evidence of keywords in text. In line with the narrative review literature (Dixon-Woods et al., 2006) we included the sources which are of analytical significance to the paper. We have also read and cited a much wider range of sources including books and research reports from the fields mentioned earlier. Table 1 outlines our guided narrative review approach.

Second, our narrative review approach foregrounds our conceptual framework, which allows us to study the gap in and subsequently the way forward to a bias-free approach in algorithmic decision making. Setting up the key lines of inquiry to investigate our central problem, involved an iterative process of 'review and (re)constitution' (c.f.,

TABLE 1 Literature review process

Search engine	Web of science
Key disciplines	Business and management, HRM, organisational behaviour, sociology, information systems (IS), social psychology and critical diversity areas
Key terms	Algorithm, bias, discrimination, inequality, diversity, HRM, OB, and IS
Total number of papers (none combined)	Over a million papers
Combined terms: Algorithm and bias/discrimination/diversity	10,000 papers
Narrow combination: Algorithm, bias and HR	144 papers
Elimination method: Reading	84 deemed outside the scope of the study
Final selection	60 manuscripts were reviewed, and half (30) are included

Rhodes & Pullen, 2018: 485). This phase of the narrative review process fed into the construction and reconstruction of our conceptual model while also identifying the gap to be addressed in our investigation. As we note, there is a lack of inquiry into the potential/actual bias and discrimination of HR algorithms, which hides behind a facade of objectivity and scientificity. This gap informed our core assumptions: the purported bias-free, scientifically derived, and objective assumptions underpinning algorithmic-decision making. We coupled this investigation with real life and HR examples, carefully examining the narratives surrounding the process and outcomes of algorithms in these examples.

Third, in so doing, a narrative review informs the construction of recommendations to address the central problem and future research. Accordingly, we develop a bias-free methodology in the HR decision making assemblage.

In the following section we present five ways through which HR algorithms could influence the perpetuation of inequalities in organisations.

3 | FIVE WAYS THROUGH WHICH HR ALGORITHMS COULD INFLUENCE THE PERPETUATION OF INEQUALITIES IN ORGANISATIONS

There are a myriad of ways in which HR algorithms could perpetuate societal inequalities and ongoing biases in workplaces. In this section, we propose five ways through which HR algorithms could influence the perpetuation of inequalities in organisations. We suggest that hidden behind these technical limitations is a neglect of the assumptions and values on which algorithmic decision-making is established.

Programed for Bias. Algorithms use data to find associations between inputs and outcomes (Gillespie, 2014). When the training data is biased, due to, for example, the underrepresentation of specific subgroups or real-world discrimination, the algorithmic predictions could reflect these shortcomings and perpetuate ongoing inequity (Crawford, 2013). For instance, employers very often wish to model what distinguishes best performers from their lower-rated colleagues and they use performance evaluations as a definition of success. If these performance evaluations are themselves biased, privileging men for instance (Rivera, 2015), then the algorithm might predict that men are more likely to perform better than women.

Recruitment algorithms that promise to target jobseekers with interest and a skills/job match (Datta et al., 2018), could actively reify biases in the employment market. They may limit the number of underrepresented groups who are alerted by specific employment opportunities and restrict others from being seen by recruiters (Burke et al., 2018). Further, recruitment algorithms may optimise the delivery of job ads based on the behaviour of the users (Chen et al., 2018). As such, when a recruiter looks online at the resumes of male programmers for web design opportunities, the algorithm will present them with more male programmers and more male programmers will be alerted for these job opportunities than women, whose resumes were initially screened out (Bogen & Rieke, 2018).

Screening algorithms that are used for the automatic review of applicants' resumes may also produce discriminatory outcomes when they are based on historically biased hiring decisions (Barocas & Selbst, 2016). For example, when screening algorithms create a match between a candidate's resume and preferred qualifications derived from the organisation's past hiring decisions, this may result in the possible reproduction of past individual, organisational, and structural biases (Martin, 2018). Finally, hiring algorithms could also perpetuate biases using chatbots that assess applicants' qualifications using natural language processing (NLP) tools to extract information which may discourage candidates from applying for a job if they detect "poor fit" (Lai, 2016). Sutton et al. (2018) found that NLP tools connected African American names with negative sentiments and female names with domestic rather than professional or technical occupations (Bogen & Rieke, 2018). Limited training data may also provoke the unfair exclusion of minority applicants who have strong accents and are not native language speakers (Blodgett et al., 2016).

Proxies. Criteria that serve as proxies for group membership (Barocas & Selbst, 2016) could be implicated in algorithmic decision-making biases, producing patterns based on flawed motifs of causation (Chandler, 2017). For example, assessing employment gaps as a criterion in hiring might inadvertently influence women jobseekers as they exit the workplace in greater numbers due to caring responsibilities (Ajunwa, 2020). Proxies for "interest" may also be very powerful in reproducing cognitive biases (Bogen & Rieke, 2018). For instance, if an experienced woman tends to look for lower-level positions online because she lacks confidence in her qualifications, over time she will be targeted by lower paying and lower status positions (Bogen, 2019). Hiring algorithms could also optimise the level of salary, bonus, and benefits offered to candidates in order to increase their possibility of acceptance (Burke et al., 2018). Such recommendations however may augment gender or racial pay gaps since HR data include numerous proxies that could be reflected in salary recommendations (Porter & Jones, 2018). The existence of proxies thus indicates that eliminating identifying variables, such as race or gender, may not deter algorithmic models from mirroring patterns of past bias.

Facial variables may also act as proxies to limit the job opportunities of particular groups (Kramer & Ward, 2010). Bias of this nature can occur even where algorithms are deemed technically robust (Barocas & Selbst, 2016). Buolamwini and Gebru (2018) demonstrated that facial recognition algorithms cannot reliably detect the gender of African American women and Feldman et al. (2019) suggested that people have difficulties deducing emotions from facial expressions in cross-cultural contexts, and for people with disabilities (Guo et al., 2019). Most importantly, facial expressions may not be causally related to workplace performance and therefore are irrelevant proxies for hiring decisions (Buolamwini & Gebru, 2018).

Algorithmic specification of fit. A widely used criterion in personnel selection is "culture fit", described as the shared values and behaviours between a candidate and an organisation (Rivera, 2012). To achieve job and organisational matching, recruiters consider several qualities and variables as approximations of fit (Bye et al., 2014). In many respects, this process is subjective and difficult to achieve. However, when measures of culture fit are being used in algorithmic decision-making, they may become hard rules. Such rules may operate to exclude rather than create a culture of inclusion within which the candidate 'fits'.

Segregation of individuals. Algorithms tend to segregate individuals into groups, drawing conclusions about how groups behave differently (Citron & Pasquale, 2014) and their common characteristics: an action that perpetuates stereotypes (Crawford, 2013). Hiring algorithms, for example, are based on the traits that differentiate high from low performers within a company (Hart, 2005), outcomes subsequently used to recommend certain applicants in hiring and promotion decisions. Such traits, however, even when inferred accurately by the algorithms, may not be causally related to performance and could even be quite random, features which may unjustifiably allocate specific applicants, especially people with disabilities (Trewin, 2018) to lower status positions. In addition, the classification of individuals into specific identity categories, such as "male" and "female" and "cisgender" could result in the marginalisation of non-binary and transgender people, while race could act as a political classification (Bowker & Leigh Starr, 1999), signalling a status inequality (Keyes, 2018).

Technical design. Many of the HR algorithms operate on specific platforms for recruiting, hiring, and managing employees which demand their engagement with them. Engagement is dictated solely by the terms programed into

the algorithm (Ajunwa, 2018), controlling the matching process and the information users have about one another (Levy & Barocas, 2016). For example, jobseekers must accept the data demands of hiring platforms if they want to receive information for job opportunities. Applicants surrender full control to the platform which presents their application to employers without their approval. In addition, hiring platforms may rank job applicants based on numerical scores, presenting the rankings to employers while creating an objective impression of a hierarchy of choice. More importantly, it can affect employers' perceptions of job applicants during the rest of the selection process (Ajunwa & Green, 2019).

Evidently, algorithmically determined decision-making in HR makes salient specific elements of the social world that may have unintended and largely negative consequences for the success of diversity, equality, and inclusion initiatives in organisational settings. Accordingly, we argue that rather than merely proposing ways to build more accurate and fair algorithms, we should interrogate the assumptions and values that are given precedence by the turn to algorithmic decision-making in HR.

We proceed with an examination of the theoretical insights that form the building blocks to a bias-free pathway to algorithmic decision making. We focus on scientism as the main reason for the unreflective implementation of HR algorithms, and we explore how scientism is attributed to algorithms through the notions of illusio, sociomateriality and performativity that strengthen its effects.

4 | SCIENTISM, ILLUSIO AND HR ALGORITHMIC DECISION-MAKING

The unquestionable pursuit of scientism has been identified as the "orthodoxy" and the predominant research perspective of information systems (Klein & Lyytinen, 1985; Orlikowski & Baroudi, 1991). Scientism is often defined as the unfounded prioritisation of scientific method over and above other moral and reasoned arguments (Haack, 2011), historical context and relevance of results (Klein & Lyytinen, 1985). It is associated with the positivist method seeking to guide HRM towards deductively derived theories, models, and hypotheses analysed and tested with advanced statistical methods (Kaufman, 2020).

The development of HR algorithms is based on a scientist's orthodoxy that construes their recommendations as a true and objective form of knowledge. This impartial logic is considered to improve decision-making through formal reasoning and empiricism in data collection (Finkel et al., 2012). Moreover, the assumption that information is derived from objective data independently of the data scientists who develop the algorithms, leads to the belief that reality exists independently and a priori. In this sense, algorithm development is considered as an "engineering" process that could be improved by technical expertise and 'good' data (Klein & Lyytinen, 1985).

However, an over-reliance on scientism may cause significant problems since the development of algorithms as merely technical artefacts ignores the fact that the creation of meaning is a socially constructed process, evolving continuously based on human interactions and interpretations. In this sense, questions are raised about meaning perceived as something that can be measured as an objective fact and not as the outcome of conflict of interest and human intentionality (Winograd, 1980). In addition, the development of algorithms is based upon the decontextualised and quantified interpretation of the social world cultivating the view that the real dimensions of a context are indeed quantifiable (Bloch, 1986). In this way, the scientific approach elevates organisational laws (e.g., HR policy) and social conventions to a given reality (Klein & Lyytinen, 1985) turning something abstract into a material and concrete entity.

These processes are further highlighted by taking a closer look at algorithmic performance assessments and reward recommendations designed to sort employees based on merit (Espeland & Vannebo, 2007). To do that, evaluation systems deduce neat scores and metrics of merit and worth, turning different individuals into comparable entities (Espeland & Stevens, 1988). Transforming, however, a multidimensional construct into a tangible attribute based on an objective and rational measurement process may reify merit making it look like an objective employee quality (Accominotti & Tadmon, 2020). Reifying performance could lead to the legitimisation of a merit hierarchy (Accominotti, 2021) amplifying even further reward and wage differences (Accominotti & Tadmon, 2020) The allure

of the success of scientism behind algorithms has largely prevented the scientific community from scrutinising it for biases and differential impact on different communities. One fundamental mechanism by which this phenomenon can be explained is the notion of illusio, as coined by Bourdieu and Wacquant (1992). Illusio is defined as the appeal of a game which draws players together, while in turn stripping them of the possibility of questioning its rules and stakes, even when the game harms the players themselves. Illusio also enables the absence of responsibility and accountability, which is, as we see it, a fundamental problem characterising algorithms. Crucially, we note that there is an illusio associated with algorithms that prevents a reflexive and critical inquiry into the design and implementation of HR algorithms.

5 | SOCIOMATERIALITY IN HR

Accepting that the rules of the game are embedded in a visible and invisible assemblage made up of human and non-human actants leads us to question the outcomes of algorithmic decision making. The theory of sociomateriality is built upon the intersection of technology, work and organisation and as such draws attention to the constitutive entanglement of the social and material, the visible and invisible in everyday organisational life (Bader & Kaiser, 2017; Orlikowski, 2007). HR algorithms are embedded in a socio-technical assemblage which is characterised by a liminal space and time (Kitchin, 2017). The decisions steered by and derived from these algorithms are seen by some as carrying significant weight, defined as the 'power brokers of society' (Diakopoulos, 2014: 2).

Latour (2005) suggests that actants - both human and non-human are a meshed assemblage with agentic qualities creating movement, action, and influence. Tanya Li (2007: 266) aptly defines assemblage as a 'gathering of heterogeneous elements consistently drawn together as an identifiable terrain of action and debate'. Drawing therefore on sociomateriality permits the understanding of the entanglement between algorithms and the human element as real, seamless, and ongoing. There is no artificial separation between matter, meaning and experience and there is no privileging of one over the other (Orlikowski, 2007).

Employing the language and instruction of Latour's actor network theory and using algorithms as the part that circulates with other parts of the assemblage, we find that the algorithm is not simply a cognitively formulated equation, but it is also a social entity. That is, as it weaves itself into and as part of the assemblage, it generates effects and becomes a visible entity, a material artefact with outcomes and social impact. As Latour (2011, p. 797) points out: 'Take any object: At first, it looks contained within itself with well-delineated edges and limits; then something happens, a strike, an accident, a catastrophe, and suddenly you discover swarms of entities that seem to have been there all along but were not visible before and that appear in retrospect necessary for its sustenance.' This holds for HR decisions and policies, where the assemblage remains invisible, the outcome contained (and then not), where appearance only emerges under interrogation and in retrospect. That is, the rules of the game remain unquestioned and accepted as fact.

The workplace is a key site in what is an assemblage of events, activities, and processes. Crucially, the assemblage is composed of both human and non-human actants or stakeholders, shaping its processes through interactions (Savage & Lewis, 2018). Agency is gained and lost as a result of the relational interactions between actants (Baker & McGuirk, 2017). Of this, Baker and McGuirk (2017: 15) state: 'Despite different perspectives on the agential status of the other-than-human, there is wide acceptance of the need not only to unpack how agency is distributed across humans, but how human agencies are intertwined and enabled by a host of materials essential to labours of assembling.' We see value in this position as the policies and practices shaping decisions and outcomes also shape actant agency while simultaneously having agentic qualities.

For instance, how HR algorithms are embedded in a socio-technical assemblage is captured by looking at the process of performance management and rewards. Mercer's global performance management survey (2019) which collected information from HR professionals found that only 2% felt their performance management system delivered value. McKinsey (2018) reported that the major issue with employee driven performance management systems has

been the lack of objective measures and equity, the heavy-laden bureaucracy, and the lack of ongoing measures. AI powered data sets are said to provide up to the minute performance information, suggesting interventions for gaps in performance (Fisher, 2019). The benefits are said to be significant with HR professionals taking on the role of 'coach' rather than rating and ranking the performance of the team. The formula that drives the AI and shapes the algorithm is however left unquestioned. Instead, the 'human' element is seen to be messy, judgemental, and cumbersome. We fail to consider what Kitchin notes (2017, 18): 'that creating an algorithm unfolds in context through processes such as trial and error, play, collaboration, discussion and negotiation. They are ontogenetic in nature (always in a state of becoming), teased into being edited, revised, deleted and restarted, shared with others, passing through multiple iterations stretched out over time and space'. At the heart of this is the agency of the human and non-human actants which is difficult to disentangle.

Agency is considered a performative process and actants as negotiators translate, co-construct and reconstruct the social through movement and voice, making others/other parts of the assemblage do things: that is, generating rules by which the interactors abide. As such, the movement of the social becomes visible. For instance, we have: (i) the human who develops the algorithm, (ii) the algorithm, (iii) the human who enters the data, (iv) the AI, using algorithms, that mines and processes the data providing an outcome and (v) the human who arrives at a decision based on the data outcome. As outcomes are translated from one form to another form, they become stabilised as material artefacts. It is these material artefacts which are visible and by nature of their visibility become influential markers of organisational arrangements; feeding into organisational structures; and shaping organisational and HR experience and outcomes - over, within and at given points in time. Algorithmic decision-making is a performative game and as p such algorithmic recommendations enact the phenomena they depict in their code, shaping the phenomena they depict.

6 | PERFORMATIVITY IN HR

Algorithms do not only become visible as social-material artefacts, but they are also the products of and feed into performative acts. Drawing on both Austin (1962) and Butler (1990, 1993), performativity focuses on the interactive power of discourse to produce the phenomena it regulates and co-create what seems to be the objective external reality it describes (Simpson et al., 2019). In this way, social norms and structures may be reproduced through performative acts and institutionalised through repetition. Algorithms, as discourses, have the capability to develop a picture of the future enabling specific modes of acting in the present while restricting others. They enact the reality they intend to describe in their code, configuring the phenomena they depict. A training algorithm, for example, that recommends content can lead individuals to accept the recommendation. A job matching algorithm that pre-selects applicants acts directly on the real situation (Roscoe & Chillas, 2014). Hacking also demonstrates how specific group classifications can generate a loop effect, through a performative process of "shaping people" (Hacking, 1995).

Algorithms, as artefacts, are best understood as entangled within the broader assemblages of theories, artefacts, actors, and practices (D'Adderio et al., 2019). In this way, they "enact objects of knowledge and subjects of practice in specific ways" (Introna, 2016, p. 26) that might entrench biases and legitimate systems of discrimination. The outcomes of algorithms therefore might be invalid because they are tightly linked to specific contexts, parameters, and theories. The discourse surrounding the algorithm might also reveal something of the wider political dynamics of which they are a part since promises and ideas are projected onto the code itself.

It is important also to note that an algorithm differs significantly from other types of artefacts such as physically written standard operating procedures, which often function as guidelines, and include, for instance, hiring instructions outlined in the manual of a human resources department. While actors can readily ignore such procedures and guidelines, algorithms compute these decisions by following rules designed prior to the routine's performance. These rules are materially and socio-technically embedded in the algorithmic assemblage, such as for instance the operational procedures encoded in enterprise resource planning software. In other words, once rules and procedures

are algorithmically embedded in software, they become "locked doors in that they truly constrain action" (Pentland & Feldman, 2008: 242) even though actants that are part of the assemblage might decide to look for alternatives (see for instance, D'Adderio, 2008).

Perceiving algorithms as assemblages provides an understanding of the values and assumptions underlying algorithmic decision-making that might perpetuate biases. Algorithms are envisioned to promote certain values and forms of formal rationality and they are considered supercarriers of rationality (Lindebaum et al., 2020). Formal rationality is predicated on abiding by abstract and formal processes, rules, and laws, which are perceived as unproblematic and legitimate fixed ends (Kalberg, 1980) for the purposes of optimization or maximisation of outcomes (Bolan, 1999). HR algorithms are restricted to formal rationality since they operate through logical and mathematical procedures and they are automated to achieve a particular goal such as for instance to suggest the employees who are going to quit within the next three months, or to identify a number of suitable candidates for a job role. The personal qualities and lived experiences of HR professionals or specific contextual and employee specificities are rejected since they are assumed to cause randomness in decision-making for the sake of calculation and universalism (Kalberg, 1980).

Algorithms that produce decisions that restrict discussion of alternative outcomes and arrangements might restrict the choices of HR professionals and impoverish their ability to use a wider variety of value choices including issues of morality and concern for others (Healy et al., 2010). Such a worldview is unwelcoming to the diversities and richness of organisational life. This becomes problematic since the detection of biases and errors depends upon the implementation of substantive rationality, escaping from the strictures of formal rationality.

Finally, when HR professionals use algorithms in their decision-making, they do not just work conjointly with the algorithms, rather they learn from and internalise them (Galloway, 2006). Once a recommendation report on hiring, for example, is verified as legitimate, it tends to be internalised by the people it is supposed to assist. Consequently, HR professionals might develop a particular understanding of themselves as "carriers of formal rationality" rather than subjects of bounded rationality with their own expertise and (flawed) judgement. That is, they attribute to themselves the characteristics of the algorithms. However, studies on stereotyping and cognitive biases have indicated that when people are led to feel that they are unbiased or objective, they are more likely to behave in inequitable ways (Monin & Miller, 2001; Uhlmann & Cohen, 2007). Castilla and Benard (2010) reported that an uncritical belief in a meritocratic organisational culture resulted in biases towards low-status groups of employees. When people feel objective, they believe in the validity of their thoughts, and are thus more likely to enact them (Uhlmann & Cohen, 2007). This suggests HR professionals have limited reflective capacity and power for independent decision-making because they lack awareness of different choices, they are emotionally detached and have reduced accountability for consequences.

The algorithmic decision-making process verifies its recommendations, proposing solutions with a strong correlation to the way it was constructed in the first place. The proposed solutions are not legitimised through expert knowledge and socio-technical negotiation between different stakeholders, but rather through supposedly objective and scientific development processes - founded on equations - which are at the heart of algorithms. In this sense, algorithmic recommendations as discourses and as part of an assemblage have the power to guide the subjects who make a specific decision, steering their interpretation of issues and outcomes according to pre-existing suggestions.

7 | DEVELOPING A BIAS PROOFING METHODOLOGY FOR ALGORITHMIC HYGIENE FOR HR PROFESSIONALS

We acknowledge that future proofing HR algorithms for possible bias based on social diversity is a complex, multifaceted and ambitious project. Casting algorithmic decision making as an assemblage of human and non-human actants offers us the opportunity to reimagine HR algorithms: as bias-free by reflecting on who produces them, based on which theories, for which purposes, under what conditions of accountability and for whose benefit. Thus, any critical

engagement with bias-proofing HR algorithms should start with questioning the assemblage in which algorithms are produced.

We propose a system by which moral and ethical responsibility is brought into the design and implementation of algorithms; and that regulatory arrangements are built into the assemblage, ensure compliance with laws and social justice requirements.

To reimagine a bias-proof future, we place the human element and social diversity at the centre of our analysis. Accordingly, we draw on Jonsen and Ozbilgin's (2014) maturity model for organisational interventions which includes several phases: awareness raising activities whereby the problem is identified, structural remedies, and deep level learning which challenges fundamental assumptions.

In terms of the first phase, we are guided by awareness raising in identifying the unquestionable association between scientism and the use of HR algorithms. Involvement of scientists and technicians provides the illusio of a scientific product which is supposedly generated by objective methods, which could yield innocuous impact without bias, and which is then ultimately performed. As we explained in our earlier examples, algorithms in fact suffer from biases. Mergen and Ozbilgin (2021) explain that for illusio to be shattered, individuals should experience a cognitive dissonance and a moral dilemma, about the illusio and the reality of the situation. The first step then towards tackling the illusio and scientism associated with HR algorithms is to address the awareness gap among HR professionals with a view to challenge their trust in HR algorithms as innocuous devices or 'rationality carriers'. Through such awareness raising an initial step could be taken towards bias proofing HR algorithms. Only then could organisational stakeholders start collecting data and evidence about the current and future use of HR algorithms in terms of their purpose, and possible impact on diverse communities.

Turning to HR practitioners, we suggest the importance of self-questioning and reflection in order to critically probe the implementation and reproduction of algorithmic recommendations in terms of exclusion. To this end, we recommend the development of reflection about their professional identities, the assumptions they use, and their definition of what is at stake. HR practitioners should recognise that they must widen their interpretative repertoire, complementing minimally, the ethic of efficiency, effectiveness and objectivity with an ethic that emphasises equality, diversity, and inclusion. Awareness raising processes should stretch HR professionals' sense of who they are and who they might become by encouraging them to engage in an open dialog with assumptions that do not emerge from formal rationality theory.

Awareness could be raised by cultivating HR practitioners' openness to the unexpected facts that result from their involvement in organisational life. These might involve instances that depart from normative assumptions and discourses (Spicer et al., 2009). The aim is to try to destabilise formal rationality assumptions through connecting with organisational conditions and contradictions while maintaining an eye for paradoxes, unexpected reactions, or points of confusion within the organisation. Our suggestion for reflexivity has multiple sources and drivers. One of these is HRM education as an evidence-based practice. We hope that HR journals such as HRMJ provide such a basis. Further, more practically, HR practitioners could use their professional development tools such as education and training to integrate bias proofing and algorithmic hygiene in their programmes. Keeping an eye on power relations, agency, ideologies, and organisational norms inscribed in algorithmic recommendations is vital for awareness raising. The realization that algorithms do not represent the complexity of organisational reality, the plurality of actors and interests and the richness of different logics may shift HR practitioners' focus from the objectivity of data to the marginalised voices in organisations, which are excluded as outliers in algorithmic recommendations. These marginalised voices, if included, may disrupt the status quo, and create new meanings (Spicer et al., 2009). Awareness also involves the engagement of HR practitioners with potentialities, moving beyond critiquing the outcomes of algorithmic recommendations and trying to construct a sense of what the organisation could be.

The predominant aim of this undertaking is to encourage micro-emancipation for organisational decision makers (Alvesson & Willmott, 1992). For instance, an HR practitioner neglecting to follow the exact recommendations of HR algorithms might harm short-term efficiency but may be involved in the transformation of power relations and the increase of equality, diversity, and inclusion in the workplace in the medium to long term. Micro-emancipations may

also involve focussed efforts to generate micro-transformations of facets of organisational life that reinforce equality, diversity, and inclusion. These might include HR practitioners' perceptions of professional self, norms of interaction, and the bodies of knowledge and skills they use.

Once a level of awareness is attained, and possible problems are identified in the organisation, the second phase of bias proofing would involve building structures to address the problems identified. First, this involves the establishment of governance arrangements around HR algorithms, incorporating an exploration of possible inappropriate use, bias, discrimination, and other adverse ethical effects. A working party could be set up to draft the HR algorithm policy and monitor its appropriate and bias free use within processes and across the organisation. Organisations should also consider building accountability structures for commissioning, procuring, adopting, using, and retiring HR policies. Such accountability structures should include lines of authority and reporting arrangements. Finally, for appropriate and bias free use of HR algorithms, modifications should be made to the algorithms to eliminate adverse impact on diverse communities and individuals. As most organisations lack expertise and they may suffer from blind sight in bias proofing their own HR algorithms, they may seek research and advice externally. Also, for most organisations to conduct such an intervention, better regulation at the national and sectoral levels could be helpful. Such a process would also involve accountability of HR with the task of bias proofing algorithms.

The final phase for bias proofing HR algorithms requires deeper level insights. Professionals should be tasked with highlighting organisational values such as equality, diversity, inclusion, and ethical conduct to inform the design of future HR algorithms. This phase would involve a more integrated approach to the assemblage of HR algorithms, embedding the values early so that future HR algorithms arrive in the organisation already checked for biases. As this final stage is the most mature stage in organisations, we can expect that the HR professionals would work with the scientific team that designs the HR algorithms to co-create them in a way that is aligned with organisational values. We propose that organisations which lack such core values as inclusion and diversity, may consult the Sustainability and Development Goals of the United Nations for inspiration.

HR algorithms can be bias proofed retrospectively. However, in an ideal world, HR professionals should be involved in the process of co-creating algorithms. In their role as co-designers, they could ensure that bias proofing is embedded within the algorithm and not only as an afterthought. For this to happen experts need training enabling them to recognise the links between the technical, social, and material dimensions and the HR side of algorithm creation and application.

8 | CONCLUSIONS

Human biases can be inscribed into the code of the HRM algorithms sustaining inequalities while assuming a veneer of objectivity underscored by perceptions of scientism (Raghavan et al., 2020). Algorithmic decision making is perceived not just as a process that could be enhanced with “better” data, but rather, as a process that involves many human and non-human actants. This entanglement has meant that HR professionals have been prevented from consciously scrutinising bias and differential impacts on different workplace communities.

Drawing on Bourdieu's notion of ‘illusio’ (Bourdieu & Wacquant, 1992), the theory of sociomateriality (Orlikowski, 2007), and performativity (Butler, 1990, 1993), we argued that algorithmic decision-making in HRM, whilst attractive, demands some reflexive and flexible reasoning, in order to question the objectivity and scientism of algorithmic decision-making, to reinforce a more theoretically informed orientation towards algorithmic implementation, and to establish a clearer governance pathway around algorithmic decision making.

Algorithms do not only become visible as social-material artefacts, but they are also the products of and feed into performative acts. As artefacts, they are understood as being entangled within broader assemblages of theories, artefacts, actors, and practices (D’Adderio et al., 2019). Algorithms do not simply execute instructions, but they are enacted by a diverse assemblage of actants (Barad, 2007) who produce the actions they are assumed to be doing. In this way, they might entrench biases and legitimate systems of discrimination. The illusio effect illustrates that once

rules and procedures are algorithmically embedded in code, they may constrain action and choice, even though actors that are part of the assemblage might decide to look for alternative recommendations and actions. Algorithms are envisioned to promote certain values and forms of formal rationality and they are considered as supercarriers of rationality. The supposed rationality of algorithmic processes might render it difficult for HR professionals to reflect critically upon them, experiencing the recommendations as if they are decisions (Brunsson & Brunsson, 2017). Knowledge of the world is thus hindered through the dismissal of first-hand experience. The performativity of algorithms may then not just be in the code, but in the way that it becomes part of a discursive and performed understanding and assemblage of desirability, objectivity, efficiency, and rationality. In this way, the further adoption of algorithmic decision-making in HR is reinforced, consolidating the restriction of choice (Lindebaum et al., 2020).

ACKNOWLEDGEMENTS

No funding was received for this study.

CONFLICT OF INTEREST

None.

DATA AVAILABILITY STATEMENT

Data sharing not applicable–no new data generated, or the article describes entirely theoretical research.

ORCID

Joana Vassilopoulou https://orcid.org/0000-0002-8854-3135
Mustafa F. Özbilgin https://orcid.org/0000-0002-8672-9534
Dimitria Groutsis https://orcid.org/0000-0002-4890-8216

REFERENCES

Accominotti, F. (2021). Consecration as a population-level phenomenon. *American Behavioral Scientist*, *65*(1), 9–24.
Accominotti, F., & Tadmon, D. (2020). How the reification of merit breeds inequality: Theory and experimental evidence. Working Paper (42). International Inequalities Institute, London School of Economics and Political Science, London, UK.
Ajunwa, I. (2018). Facebook users aren't the reason Facebook is in trouble now. The Washington Post, March 23.
Ajunwa, I. (2020). The paradox of automation as anti-bias intervention. *Cardozo Law Review*, *41*. https://doi.org/10.2139/ssrn.2746078
Ajunwa, I., & Green, D. (2019). Platforms at work: Automated hiring platforms and other new intermediaries in the organisation of work. In S. P Vallas & A. Kovalainen (Eds.), *Work and labor in the digital age*. Emerald Publishing.
Alvesson, M., & Willmott, H. (1992). On the idea of emancipation in management and organisation studies. *Academy of Management Review*, *17*(3), 432–464.
Austin, J. L. (1962). *How to do things with words*. Oxford University Press.
Bader, V., & Kaiser, S. (2017). Autonomy and control? How heterogeneous sociomaterial assemblages explain paradoxical rationalities in the digital workplace. *Management Revue*, *28*(3), 338–358.
Baker, T., & McGuirk, P. (2017). Assemblage thinking as methodology: Commitments and practices for critical policy research. *Territory, Politics, Governance*, *5*(4), 425–442.
Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.
Barocas, S., & Selbst, A. (2016). Big data's disparate impact. *California Law Review*, *104*, 671–732.
Bloch, E. (1986). Scientific and technological literacy: The need and the challenge. *Bulletin of Science, Technology & Society*, *6*(3), 138–145.
Blodgett, S. L., Green, L., & O'Connor, B. (2016). Demographic dialectal variation in social media: A case study for African American English. Proceedings of the 2016 Conference on Empirical methods in natural language processing. https://www.aclweb.org/anthology/D16-1120.pdf
Bogen, M. (2019). All the ways hiring algorithms can introduce bias. *Harvard Business Review*, *6*. https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias
Bogen, M., & Rieke, A. (2018). Help wanted: An exploration of hiring algorithms, equity, and bias. Technical report, Upturn.

Bolan, R. S. (1999). Rationality revisited: An alternative perspective on reason in management and planning. *Journal of Management History*, *5*(2), 68–86.
Bourdieu, P., & Wacquant, L. J. D. (1992). *An invitation to reflexive sociology*. Polity Press.
Bowker, G., & Leigh Starr, S. (1999). *Sorting things out. Classification and its consequences*. MIT Press.
Brunsson, K., & Brunsson, N. (2017). *Decisions: The complexities of individual and organisational decision-making*. Edward Elgar.
Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 1–15.
Burke, R., Sonboli, N., & Ordonez-Gauger, A. (2018). Balanced neighborhoods for multi-sided fairness in recommendation. In Conference on fairness, accountability, and transparency.
Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.
Butler, J. (1993). *Bodies that matter: On the discursive limits of 'sex'*. Routledge.
Bye, H. H., Herrebroden, H., Hjetland, G. J., Royset, G. O., & Westby, L. L. (2014). Stereotypes of Norwegian social groups. *Scandinavian Journal of Psychology*, *55*(5), 469–476.
Cabantous, L., & Gond, J.-P. (2011). Rational decision making as performative praxis: Explaining rationality's eternel retour. *Organisation Science*, *22*(3), 573–586.
Castilla, E. J., & Benard, S. (2010). The paradox of meritocracy. *Administrative Science Quarterly*, *55*(4), 543–676.
Chandler, S. (2017). The AI chatbot will hire you now. Wired. https://www.wired.com/story/the-ai-chatbot-will-hire-you-now/
Chen, L., Ma, R., Hannak, A., & Wilson, C. (2018). Investigating the impact of gender on rank in resume research engines. In Proceedings of the 2018 chi conference on human factors in computing systems (pp. 1–14).
Cheng, M. M., & Hackett, R. D. (2021). A critical review of algorithms in HRM: Definition, theory, and practice. *Human Resource Management Review*, *31*, 100698. https://doi.org/10.1016/j.hrmr.2019.100698
Citron, D. K., & Pasquale, F. A. (2014). The scored society: Due process for automated predictions. *Washington Law Review*, *89*, 1–33.
Crawford, K. (2013). The hidden biases in big data. *Harvard Business Review*, 1 April. http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/
D'Adderio, L. (2008). The performativity of routines: Theorizing the influence of artefacts and distributed agencies on routine dynamics. *Research Policy*, *37*(5), 769–789.
D'Adderio, L., Glaser, V., & Pollock, N. (2019). Performing theories, transforming organisations: A reply to Marti and Gond. *Academy of Management Review*, *44*, 676–679.
Datta, A., Datta, A., Makagon, J., Mulligan, D. K., & Tscantz, M. C. (2018). Discrimination in online advertising: A multidisciplinary inquiry. In Proceedings of the 1st conference on fairness, accountability, and transparency (pp. 2–18).
Devlin, H. T. G. (2017). AI programs exhibit racial and gender biases, research reveals. Accessed on September 2020. https://www.theguardian.com/technology/2017/apr/13/ai-programs-exhibit-racist-and-sexist-biases-research-reveals
Diakopoulos, N. (2014). Algorithmic accountability reporting: On the investigation of black boxes. A Tow/Knight Brief. Tow Center for Digital Journalism, Columbia Journalism School. Accessed on June 2021. https://towcenter.columbia.edu/news/algorithmic-accountability-reporting-investigation-black-boxes
Dixon-Woods, M., Bonas, S., Booth, A., Jones, D. R., Miller, T., Sutton, A. J., Shaw, R. L., Smith, J. A., & Young, B. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, *6*(1), 27–44.
Espeland, W. N., & Stevens, M. L. (1988). Commensuration as a social process. *Annual Review of Sociology*, *24*, 313–343.
Espeland, W. N., & Vannebo, B. I. (2007). Accountability, quantification, and law. *Annual Review of Law and Social Science*, *3*(1), 21–43.
Feldman, L., Adolphs, R., Marsella, S., Martines, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68.
Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public Interest*, *13*(1), 3–66.
Fisher, A. (2019). An Algorithm May Decide Your Next Pay Raise, Fortune. Accessed on June 2021. https://fortune.com/2019/07/14/artificial-intelligence-workplace-ibm-annual-review/
Galloway, A. R. (2006). *Gaming: Essays on algorithmic culture*. University of Minnesota Press.
Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. Boczkowski, & K. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society*. MIT Press.
Guo, A., Kamar, E., Wortman Vaughan, J., Wallach, H., & Ringel Morris, M. (2019). Toward fairness in AI for people with disabilities: A research roadmap. *ACM SIGACCESS*, *125*.
Haack, S. (2011). *Defending science-within reason: Between scientism and cynicism*. Prometheus Books.
Hacking, I. (1995). The looping effects of human kinds. In D. Sperber & J. Premack (Eds.), *Causal cognition*. Oxford University Press.
Hart, M. (2005). Subjective decision making and unconscious discrimination. *Alabama Law Review*, *56*, 741–791.
Healy, G., Kirton, G., Ozbilgin, M., & Oikelome, F. (2010). Competing rationalities in the diversity project of the UK judiciary: The politics of assessment centres. *Human Relations*, *63*(6), 1–28.

Hmoud, B., & Laszlo, V. (2019). Will artificial intelligence take over human resources recruitment and selection? *Network Intelligence Studies*, *13*, 21–30.

Hodgkinson, G. P., & Ford, J. K. (2014). Narrative, meta- analytic, and systematic reviews: What are the differences and why do they matter? *Journal of Organisational Behavior*, *35*, 1–5.

Hovorka, D.S., Peter, S. (2018). Thinking with Monsters. In: Schultze, U., Aanestad, M., Mähring, M., Østerlund, C., Riemer, K. (eds) *Living with Monsters? Social Implications of Algorithmic Phenomena, Hybrid Agency, and the Performativity of Technology*. IS&O 2018. IFIP Advances in Information and Communication Technology, vol 543. Springer, Cham. https://doi.org/10.1007/978-3-030-04091-8_12

Introna, L. D. (2016). Algorithms, governance, and governmentality: On governing academic writing. *Science, Technology, & Human Values*, *41*(1), 17–49.

Jelf, G. (1999). A narrative review of post-1989 interview research. *Journal of Business and Psychology*, *14*, 25–58.

Jones, O., & Gatrell, C. (2014). Editorial: The future of writing and reviewing for IJMR. *International Journal of Management Reviews*, *16*, 249–264.

Jonsen, K., & Özbilgin, M. (2014). Models of global diversity management. In B. M. Ferdman & B. R. Deane (Eds.), *Diversity at work: The practice of inclusion* (pp. 364–390). Jossey-Bass/Wiley.

Kalberg, S. (1980). Max Weber's types of rationality: Cornerstones for the analysis of the rationalization processes in history. *American Journal of Sociology*, *85*(5), 1145–1179.

Kaufman, B. E. (2020). The real problem: The deadly combination of psychologization, scientism, and normative promotionalism takes strategic human resource management down a 30-year dead end. *Human Resource Management Journal*, *30*(1), 49–72.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. The *Academy of Management Annals*, *14*(1), 366–410.

Keyes, O. (2018). The misgendering machines: Trans/HCI implications of automatic gender recognition. *Proceedings of the ACM on Human-Computer Interaction*, *2*(CSCW), Article 88.

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, *20*(1), 14–29.

Klein, H. K., & Lyytinen, K. (1985). The poverty of scientism in information systems. In E. Mumford, R. Hirscheim, G. Fitzgerald, & A. T. Wood-Harper (Eds.), *Research methods in information systems*. North Holland.

Köchling, A., Riazy, S., Wehner, M. C., & Simbeck, K. (2021). Highly accurate, but still discriminatory. *Business & Information Systems Engineering*, *63*(1), 39–54.

Kramer, E. S. S., & Ward, R. (2010). Internal facial features are signals of personality and health. *Quarterly Journal of Experimental Psychology*, *63*(11), 2273–2287.

Lai, P. (2016). Natural language processing. 1 Georgetown Law Technology Review, 98.

Langer, M., Konig, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment*, *27*, 217–234.

Latour, B. (2005). *Reassembling the Social – An introduction to actor-network-theory*. Oxford University Press.

Latour, B. (2011). Networks, societies, spheres: Reflections of an actor-network. *Theorist International Journal of Communication*, *5*, 796–810.

Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics*, *160*, 377–392.

Lepri, B., Oliver, N., Letouze, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy and Technology*, *31*, 611–627.

Levy, K., & Barocas, S. (2016). Designing against discrimination in online markets. *Berkeley Technology Law Journal*, *32*(3), 1183–1237.

Li, T. M. (2007). Practices of assemblage and community forest management. *Economy and Society*, *36*(2), 263–293.

Lindebaum, D., Vesa, M., & den Hond, F. (2020). Insights from "The Machine Stops" to better understand rational assumptions in algorithmic decision making and its implications for organisations. *Academy of Management Review*, *45*(1), 247–263.

MacCormick, J. P. (2012). *Stochastic algorithms for visual tracking*. Springer.

Martin, K. (2018). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*, 835–850.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.

McDonald, K., Fisher, S., & Connelly, C. E. (2017). E-HRM systems in support of "smart" workforce management: An exploratory case study of system success. *Electronic HRM in the smart era* (pp. 87–108). https://doi.org/10.1108/978-1-78714-315-920161004

McKinsey. (2018). The Fairness Factor in Performance Management, McKinsey Quarterly. Accessed on May 2021. https://www.mckinsey.com/business-functions/organisation/our-insights/the-fairness-factor-in-performance-management#

Mercer. (2019). Performance Transformation in the Future of Work: Insights from Mercer's 2019 Global Performance Management Study. Accessed on May 2021. https://www.mercer.com/our-thinking/career/performance-transformation-in-the-future-of-work.html

Mergen, A., & Ozbilgin, M. (2021). Understanding the followers of toxic leaders: Toxic illusio and personal uncertainty. *International Journal of Management Reviews*, *23*, 45–63.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1–21.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*(1), 33–43.

Ong, J. H. (2019). *Ethics of artificial intelligence recruitment systems at the United States secret service*.

Orlikowski, W. (2007). Sociomaterial practices: Exploring technology at work. *Organisation Studies*, *28*(9), 1435–1448.

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organisations: Research approaches and assumptions. *Information Systems Research*, *2*(1), 1–28.

Parry, K. W., Cohen, M., & Bhattacharya, S. (2016). Rise of the machines: A critical consideration of automated leadership decision making in organisations. *Group & Organisation Management*, *41*(5), 571–594.

Pentland, B. T., & Feldman, M. S. (2008). Designing routines: On the folly of designing artifacts while hoping for patterns of action. *Information and Organisation*, *18*, 235–250.

Porter, S. R., & Jones, M. R. (2018). *Where you grow up can affect your future*. United States Census Bureau. https://www.census.gov/library/stories/2018/10/opportunity-atlas.html

Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, pp. 469–481.

Rhodes, C., & Pullen, A. (2018). Critical business ethics: From corporate self-interest to the glorification of the sovereign pater. *International Journal of Management Reviews*, *20*, 483–499.

Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *American Sociological Review*, *77*(6), 999–1022.

Rivera, L. A. (2015). Go with your gut: Emotion and evaluation in job interviews. *American Journal of Sociology*, *120*(5), 1339–1389.

Roscoe, P., & Chillas, S. (2014). The state of affairs: Critical performativity and the online dating industry. *Organisation*, *21*(6), 797–820.

Savage, G. C., & Lewis, S. (2018). The phantom national? Assembling national teaching standards in Australia's federal system. *Journal of Education Policy*, *33*(1), 118–142.

Schildt, H. (2020). *The data imperative: How digitalization is reshaping management, organizing, and work*. Oxford University Press.

Simpson, R., Kumra, S., Lewis, P., & Rumens, N. (2019). Towards a performative understanding of deservingness: Merit, gender and the BBC pay dispute. *Gender, Work and Organization*, *27*(2), 181–197.

Spicer, A., Alvesson, M., & Karreman, D. (2009). Critical performativity: The unfinished business of critical management studies. *Human Relations*, *62*(4), 537–560.

Sutton, A., Lansdall-Welfare, T., & Cristianini, N. (2018). Biased embedding from wild data: Measuring, understanding and removing. https://arxiv.org/pdf/1806.06301.pdf

Trewin, S. (2018). *AI fairness for people with disabilities: Point of view*. Cornell University.

Uhlmann, E. L., & Cohen, G. L. (2007). "I think it, therefore it's true": Effects of self-perceived objectivity on hiring discrimination. *Organizational Behavior and Human Decision Processes*, *104*, 207–223.

Winograd, T. (1980). What does it mean to understand language? *Cognitive Science*, *4*(3), 209–241.

How to cite this article: Vassilopoulou, J., Kyriakidou, O., Özbilgin, M. F., & Groutsis, D. (2022). Scientism as illusio in HR algorithms: Towards a framework for algorithmic hygiene for bias proofing. *Human Resource Management Journal*, 1–15. https://doi.org/10.1111/1748-8583.12430