

# Auditor judgment in the fourth industrial revolution

Rita Samiolo<sup>1</sup> | Crawford Spence<sup>1</sup> | Dorothy Toh<sup>2</sup>

<sup>1</sup>King's College London, London, UK

<sup>2</sup>University College London, London, UK

## Correspondence

Crawford Spence, King's College London,  
London, UK.

Email: [crawford.spence@kcl.ac.uk](mailto:crawford.spence@kcl.ac.uk)

## Abstract

Discourse proclaiming the advent of a fourth industrial revolution predicts significant disruption to various work domains in the near future. Auditing is one of the domains where bold claims about the potential of technology are being made, with technology expected to augment auditors' judgments and, in time, possibly automate them. Drawing on 44 in-depth interviews with auditors, regulators, and emergent artificial intelligence software providers, we question the prevailing narrative around technological change in auditing which suggests that ostensibly simple, low-level technical tasks are areas where little judgment is at play and thus are ripe for automation. We show that significant elements of deliberation, sense-making, and reflexivity, arguably critical for the socialization of early career auditors into the profession, may be lost when automating areas of work perceived as low value, leading us to question what it means to apply judgment in auditing. Conversely, higher-level aspects of the audit process may be assisted by technology and augmented in different ways, yet new technological structures generate new areas of indeterminacy that pose new and yet unresolved demands on auditors' judgment. Overall, the paper shows how auditor habits are changing and highlights the risks posed by new technologies to the acquisition of practical knowledge by auditors.

## KEYWORDS

AI, auditing, augmentation, automation, judgment, structure

Accepted by Yves Gendron.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Contemporary Accounting Research* published by Wiley Periodicals LLC on behalf of Canadian Academic Accounting Association.

## Le jugement de l'auditeur à l'heure de la quatrième révolution industrielle

### Résumé

Le discours annonçant l'avènement d'une quatrième révolution industrielle prédit d'importantes perturbations dans divers domaines de travail dans un avenir proche. L'audit est l'un des domaines où l'on fait des affirmations audacieuses sur le potentiel de la technologie, laquelle est censée fortifier le jugement des auditeurs et, à terme, peut-être même l'automatiser. Sur la base de 44 entretiens approfondis avec des auditeurs, des régulateurs et des fournisseurs de logiciels d'intelligence artificielle émergente, les auteurs remettent en question le discours dominant sur le changement technologique en audit qui suggère que les tâches techniques apparemment simples et peu exigeantes représentent des secteurs où peu de jugement intervient, et qu'elles sont donc propices à l'automatisation. Les auteurs démontrent que des éléments importants de délibération, de construction de sens et de réflexivité, sans doute essentiels à la socialisation des auditeurs en début de carrière dans la profession, peuvent être perdus lors de l'automatisation de secteurs de travail perçus comme étant de faible valeur, ce qui les amène à s'interroger sur le sens de l'exercice du jugement en audit. De façon inverse, des aspects de niveau supérieur du processus d'audit peuvent être assistés par la technologie et renforcés de différentes manières, mais les nouvelles structures technologiques génèrent de nouvelles zones d'indétermination posant de nouvelles exigences, encore non satisfaites, en matière de jugement des auditeurs. Dans l'ensemble, cet article montre comment les habitudes des auditeurs évoluent et expose les risques liés aux nouvelles technologies pour l'acquisition de connaissances pratiques par les auditeurs.

### MOTS-CLÉS

audit, automatisation, IA, jugement, renforcement, structure

## 1 | INTRODUCTION

Discourse of a fourth industrial revolution singles out the domain of financial auditing as being at particularly high risk of imminent disruption (Frey & Osborne, 2017; Schwab, 2017; Susskind & Susskind, 2015; World Economic Forum, 2020). Perceptions of audit work as routine, calculative, prescribed, and process-driven, along with the quantitative nature of the financial data being handled, have contributed to this view of audit's susceptibility to automation and implied impending demise. Yet technological innovation within auditing has a long history (AICPA, 2015), showing that technology has often failed to deliver on its bolder promises in the past (Manson et al., 1997; Salijeni et al., 2019). Furthermore, the diffusion of novel technologies such as big data analytics can be more difficult and slower than implied by many accounts of technological change, due to both tensions internal to audit firms and lack of guidance

offered by regulators (Austin et al., 2021). This is even more likely to be the case when it comes to the deployment of more advanced technologies such as machine learning (ML) and artificial intelligence (AI), which promise to mimic auditor judgment. Whereas previous studies into the impact of the fourth industrial revolution have tended to focus on the replicable or repetitive elements of audit work (Frey & Osbourne, 2017; Susskind & Susskind, 2015) or on abstract perceptions of the future (World Economic Forum, 2020), we posit that a greater understanding of the interplay between individual judgment and the adoption and use of what we term new “audit technologies” (referred to as the digital tools and software programs deployed in the course of conducting and delivering financial audits, to include robotic process automation, data analytics, ML, and AI), is especially necessary to forming an overall view on the future of audit work. This dynamic also speaks closely to the tension, or complementarity, between human and machine, which lies at the heart of both future of work discourse (Autor et al., 2003) and the perennial structure-judgment debate in auditing (Dirsmith & McAllister, 1982; Power, 2003).

The structure-judgment debate in auditing has persisted for decades and is often given new life when audit embraces new technology. However, it is curious that few studies have sought to empirically unpack the dynamics of this interrelationship (Kohler et al., 2021). Building upon conceptual work on habit (Camic, 1986; Crossley, 2013; Turner & Cacciatori, 2016), we explore the ways in which audit judgment is both shaping and being reshaped by the advent of audit technologies. By drawing on 44 in-depth interviews and follow-up discussions with primarily UK-based practitioners, regulators, AI providers, the professional institute, and others, we focus on how professional judgment emerges as a key stake in the deployment of new technologies. Our findings make us question the prevailing narrative about technological change in auditing, which suggests that ostensibly simple, low-level technical tasks are areas where little judgment is at play and thus are ripe for automation, whereas other more complex judgments will be “augmented” by new technologies. We show that significant elements of deliberation, sensemaking, and reflexivity, arguably critical for the socialization of early career auditors into the profession, may be lost when automating areas of work perceived as low value, leading us to question what it means to apply judgment in auditing. Complex decisions may be assisted by technology and somewhat “augmented,” yet new technological structures aimed at supporting judgment generate new areas of indeterminacy which pose new demands on auditors’ judgment. It is only in certain conditions that augmentation actually takes place, with auditors seen to be overriding outputs from technology when these cannot be reconciled with extant knowledge of the client. Here increasing structure goes hand in hand with the need to apply judgment to machine outputs, but the ability to exercise such judgment might depend to a significant extent on knowledge of the client developed in a pre-augmentation world.

Thus, the recent wave of technological change poses a challenge to the classic distinction between audit structure and judgment and offers the opportunity to examine the terms and conditions of their interdependence. Questions remain as to whether such interdependence might over time tilt toward technological determinism, which authors like Francis (1994) saw as eroding the hermeneutical nature of audit.

## 2 | PRIOR LITERATURE

### 2.1 | Technological innovation and the structure-judgment tension within audit

Technological change within audit is not new, with a long history of attempts by auditors to legitimize their practices in the face of demands for improved quality and efficiency (Carpenter & Dirsmith, 1993; Curtis et al., 2016; Fischer, 1996; Power, 1997, 2003). Previous research has explored different dimensions of technological change, broadly defined, including

social constructivist perspectives on the emergence of specific techniques such as statistical sampling in the 1960s (Carpenter & Dirsmith, 1993; Power, 1992) or the development of business risk methodologies alongside shifting market forces and changes in the audit field (Robson et al., 2007). A significant body of literature has examined, from a more normative perspective, the impact, or the potential, of specific technologies on the audit process itself. This work covers, *inter alia*, the advent of the paperless audit (Bierstaker et al., 2001), the possibilities of continuous auditing (Alles et al., 2008; Rezaee et al., 2001), or the use of computer-assisted audit techniques (CAATs) (Bierstaker et al., 2014; Braun & Davis, 2003) in the audit process. In each case, this normative research addresses how specific technologies can be optimally used to either “improve” audit practice or explores the opportunities for better auditing implied within.

More recently, the rise of robotic process automation (RPA) (Eulerich et al., 2022) and big data and data analytics (collectively, BDA) in auditing has garnered significant practitioner, regulatory (Financial Reporting Council, 2020; IAASB, 2016, 2018; PCAOB, 2018), and academic interest (Appelbaum et al., 2017, 2018; Brown-Liburd et al., 2015; Salijeni et al., 2019, 2021). Building on earlier literature exploring the possibilities of such tools (Cao et al., 2015) and related challenges (Alles, 2015) or setting out urgent research needs (Appelbaum et al., 2017), emergent studies have started to explore empirically the actual use of BDA in auditing (Austin et al., 2021; Eilifsen et al., 2020; Salijeni et al., 2019, 2021).

For example, Salijeni et al. (2021) find that BDA has reconfigured aspects of the audit process and changed relational dynamics between different groups within audit firms, with some of the wider relational dynamics and tensions within the field explored by Austin et al. (2021). Adopting a sociomaterial perspective, and specifically focusing on the technical affordances of BDA, Salijeni et al. (2021) conclude that further research is needed for understanding how judgments are derived at the sites of where BDA technologies are being deployed. This need for greater scholarly attention on professional judgment with respect to audit technologies is also consistent with earlier papers that highlight the significant challenges BDA poses to auditor decision-making when confronted with information overload, multiple ways to analyze data, and ambiguity (Brown-Liburd et al., 2015).

More recently, research into the emergence of new technology in auditing has started to investigate AI’s possibilities for the audit approach and process (Kokina & Davenport, 2017; Raschke et al., 2018), the research agenda it invites (Issa et al., 2016), its implications for professional identity (Goto, 2021), and the ethical considerations it engenders (Munoko et al., 2020), including the potential for unintended consequences such as individual deskilling effects or, crucially, the risk of the auditor abdicating its responsibility for judgment (Munoko et al., 2020).

Questions on the interplay between technology and professional judgment are related to the long-standing tension between the trend for increasingly formalized and structured approaches on the one hand and the desire to uphold audit as a domain requiring individual discretion and subjective judgment on the other (Cushing & Loebbecke, 1986; Dirsmith & McAllister, 1982; Kohler et al., 2021; Power, 1992, 2003). The history of audit is marked by the growing “importance of procedural efficiency in the light of the difficulty of observing the output of audit” (Power, 2003, p. 388) and by the ongoing attempt to “balance . . . expert discretion and impersonal rules” (Power, 1996, p. 8) in the search for standardization, so as to make professional judgments replicable and thus defensible in court. New digital and automated tools can be seen as the next section in this long-term pursuit of efficiency and standardization via structured approaches.

Building on Dirsmith and McAllister’s (1982) notion of “mechanistic audit,” Cushing and Loebbecke (1986) provide one of the earlier definitions of a structured audit approach: “A systematic approach to auditing characterized by a prescribed, logical sequence of procedures, decisions, and documentation steps, and by a comprehensive and integrated set of audit policies and tools designed to assist the auditor in conducting the audit” (p. 32). Technology has often been

conceived within this framework as contributing to greater structure, although, as we discuss, this relationship needs qualifying. Structure, in turn, has predominantly been conceptualized within this early literature in opposition to judgment. For example, Abdolmohammadi and Wright (1987, p. 4) define structured tasks as those involving “routine, well-defined problems . . . requiring little judgement,” whereas unstructured tasks are those with “unique, undefined problems, few or no available guidelines, many undefined alternatives” and needing “judgement and insight” (Abdolmohammadi & Wright, 1987, p. 4).

Since these early conceptions, a substantial body of literature has offered critical insights into the growth of “structure,” which, as Power (2003, p. 381) put it, may be seen as being “about legitimacy and control” and “not necessarily consistent with better or more efficient auditing” (see also Bowrin, 1998) and which may remain only loosely coupled to audit practice (Humphrey & Moizer, 1990). For their part, Carpenter et al. (1994, p. 375) see audit structure as the abstract system of knowledge, the sampling tools, and the expert systems that “commodify and deprofessionalize the profession by providing the wherewithal to encode expertise in the formal structure of the organization” and as promoting a “one best way mode” or “mechanistic culture.” Similarly, Carpenter and Dirsmith observed, with specific reference to statistical sampling as part of a wider set of structured approaches (Carpenter & Dirsmith, 1993, pp. 55–56):

By standardizing and normalizing judgment and encoding it in such forms of organizational structure as firm audit manuals, and by specifying thresholds of such judgment errors as incorrect acceptance or rejection of a client’s financial statement assertions, statistical sampling transfers power from the practitioner to the administrative component of public accounting firms. . . . Statistical sampling thus joins a complex of techniques transforming collections of individual practitioners into modern professional bureaucracies, wherein judgment becomes encoded in the very structure of the organization.

Carpenter et al. (1994) see pressures toward greater structure as possibly countervailed by the resistance proffered by “more seasoned practice partners” (p. 375), casting the tension in terms of organizational structure versus resistance of experienced individual practitioners.

Francis (1994, p. 253) echoes these critical stances against structure, defined as “the use of formalized, standardized and predetermined representations of the audit process and evidence gathering/interpreting procedures.” He associates structure with greater standardization, such that the application of structure goes hand in hand with the assumption that audits are “homogeneous across contexts” and thus able to “benefit economically and technologically from . . . standardization and rule-bound rationality,” where the latter is seen to efface judgment (p. 253):

Judgment is given over to a kind of simplistic and unreflective rule-bound rationality that defines both what is appropriate evidence and the algorithms for “objectively” evaluating the evidence. Thus the “end” of the audit is already given and structure simply provides the technocratic means to achieve that end instrumentally. There is no reflexive mediation by the auditor on either the ends or the means as they’re both given!

At the other end of the spectrum, Knechel (2007) describes increasing structures in audit firms in relation to the need to minimize errors of judgment (p. 386), reflecting a more optimistic and normative view of structure as reducing “bad” judgment. As Francis summarizes, those who favor structure tend to believe “that the systematic character of structured audits enhances the accounting firm’s quality control over audits, thus reducing the risk of audit failures and at

the same time providing documentation that the accounting firm has conformed to generally accepted audit standards and hence is not negligent” (Francis, 1994, p. 252).

Judgment tends to be used in this literature to refer to the nonstructured, often tacit, situational responses of the individual auditor, which are shaped through experience and socialization but not coded in manuals, protocols, and methods and thus are potentially nonreplicable (Carpenter et al., 1994). Francis sees judgment as a hermeneutic practice entailing subjectivity and a “capacity for practical reasoning over the ends and means of the audit” (Francis, 1994, p. 253), which structuration puts at risk in that the decision aids which standardize audit methodologies might become an end in themselves rather than be at the service of auditors’ “genuine understanding.”

While preoccupation with human error or bad judgment on the one hand and with the erosion of judgment by structure on the other, suggest a sort of zero-sum game, as Power put it, “Both structured and unstructured audit approaches are problematic ideals or programs (Rose & Miller, 1992) which can never be fully or perfectly realized” (Power, 2003, p. 381). Francis (1994) concurs by observing that (p. 251) “the terms ‘structure’ and ‘judgement’ could be argued to be ideological rather than technical categories.” Francis further notes that structure and judgment are better understood in terms of an ongoing tension—a duality rather than a dichotomy (p. 251):

The binary opposition of structure-judgement can be easily deconstructed. Structure does not “end” judgement, rather it relocates judgement and directs it in certain ways. In this respect any audit methodology necessarily contains both structure and judgement and a tension operates between them.

Francis argued that the programmatic ambition to increase audit structure amounts to a belief in objectivity that erodes the ability to perceive that judgment is always at play and audits are always, to some extent, subjective. Belief in structure, according to Francis, “deforms both the hermeneutic character of auditing and the potentiality for practical reasoning” (p. 251). In this respect, despite noting the interdependence between structure and judgment, Francis focused his discussion on the negative implications of structure for audit as an interpretive practice and did not fully explore what the terms of such interdependence might be.

More recently, IFRS adoption has sparked new interest in the structure-judgment tension. Kohler et al. (2021) examined such tension in the context of IFRS interpretation within the audit process, illustrating how newly established professional practice functions (PPFs) play a key role in addressing the increased demand for both structure and judgment that complex accounting issues pose in global accounting firms. Kohler et al. show that as global firms strive to control local practice, PPFs enable greater global consistency by intensifying structure while also promoting the relocation of judgment from the local to the global level.

The advent of new audit technologies, too, appears to have reignited interest in the conditions and implications of the structure-judgment relationship. For example, Boland et al. (2019) examined the relationship between structured audit technologies (defined as checklists, decision aids, standardized forms, and processes—electronic and manual) and PCAOB inspection outcomes. While improving inspection outcomes is a clear motivator for the adoption of structured technologies, Boland et al. (2019) do not find their use to be effective in achieving such outcomes, citing auditors’ concerns about the negative impact structured technologies can have on critical thinking and skepticism. Indeed, Eulerich et al.’s (2022) recent study of RPA rollout in audit firms finishes with a call for more research that explores whether overreliance on technology “erodes auditors’ understanding of the audit process” (p. 712) and whether auditors become more or less professionally skeptical as a result, echoing the worries expressed in much of the earlier literature on this topic.

Preoccupations with the undermining of judgment and its “bureaucratization” within audit firm structures, whether coming from auditing scholars or practitioners, testify to how the notion of judgment represents a key symbolic resource around which the authority and legitimacy of the auditor is maintained (Power, 1992). It is thus not surprising that the discourse surrounding new technologies needs to promote their potential for supporting auditors’ judgment. Just as Power observed 30 years ago, “The progressive investment of scientific rationality in the audit process is paralleled by an intensification of the discourse of expert judgement” (Power, 1992, pp. 37–38). Such discourse today predicates that new technologies should replace human auditors in low (deterministic) judgment areas—a process broadly corresponding to what is termed “automation” in the more practitioner-oriented AI literature (see review in Raisch & Krakowski, 2021)—and focus auditors’ attention on areas requiring more judgment, where AI is seen to have the potential to “augment” such judgment rather than automate it (Davenport & Kirby, 2016; Kokina & Davenport, 2017). In this respect, Moffitt et al. (2018, p. 1) suggest that parts of the audit “that are prone to the utilization of workflow and time and motion improvements” or “that have repeatable judgements that, by and large, are deterministic if the information is available” should be prime candidates for automation. Machine-assisted, “augmented” human judgment is seen as needed in complex areas where uncertainty is high, whereas automation is promoted in low uncertainty, ostensibly highly standardized parts of the audit, where little or no judgment is at play. The rationale for automation appears the same that has characterized prior standardization efforts and the further structuration of the audit process: efficiency and replicability. Conversely, the rhetoric of augmentation is premised on the idea that new technology (which will entail elements of structuration) can support complex judgments in less replicable parts of the audit process—implying a new cooperation of sorts between structure and judgment.

Nascent research on augmentation in other professional domains, however, suggests that augmentation (and, by extension, the new structure-judgment cooperation that is implied by the innovation discourse) is rarer than conveyed by prevailing narratives, with the discarding of machine output a not-infrequent outcome of the use of new tools in professional work. For example, Lebovitz et al. (2022) show that the use of opaque AI tools for critical judgments in the domain of medicine increases the uncertainty experienced by decision-makers, who often struggle to reconcile AI outputs with their prior judgments. The result is that doctors often ignore such outputs or else passively accept them without reflection. The authors found that only if users have the ability and resources to interrogate the output of AI tools and reconcile it to their own prior judgment does some form of “augmentation” take place (which the authors term “engaged augmentation”). Such augmentation is defined as a form of learning stemming from the collaboration human-machine, in which knowledge claims by humans and machines are compared and made compatible through interrogation practices, ultimately reducing the uncertainty experienced. According to the authors, when AI interrogation is not possible—for either lack of time or absence of complementary technologies and approaches that can help corroborate or discard the output of AI tools—no real augmentation takes place. Faulconbridge et al. (2023) also highlight several instances where legal and accounting professionals dismiss new technology rather than integrate it into their work. Faulconbridge et al. (2023) argue that such dismissal reinforces practitioners’ sense of professional self, a process they refer to as defensive boundary work.

Thus, the automation/augmentation agenda in auditing invites two key questions concerning (1) the implications of reducing human input in the areas where decision-making is seen as more deterministic and standardizable (areas of audit work that are ostensibly highly structured) and (2) the nature of the machine-assisted, “augmented” judgments auditors are expected to embrace in the “higher level,” less structured, and standardizable parts of the audit process.

All in all, the automation/augmentation seemingly at play in the recent wave of technological innovation point to the need to revisit the structure-judgment duality, seen not so much as

a zero-sum game, or as a one-way process, but as an ongoing interplay where the erosion of the capacity for practical reasoning noted by Francis (1994) may be only one (albeit crucial) vector in a more complex dynamic shaping the evolution of audit. In order to examine this interdependence further, we draw on a multidimensional notion of habit, to which we now turn.

## 2.2 | Audit practice and auditor habits

Organizational research has long contended with the tension between more structured, repetitive, “routine” work on the one hand (Pentland & Feldman, 2005) and more reflexive, deliberative, or “mindful” work on the other (Weick & Sutcliffe, 2006), concluding that “the enactment of neither mindful nor routinized behavior is possible without the other” (Feldman & Pentland, 2003; Levinthal & Rerup, 2006, p. 503). Levinthal and Rerup (2006) identified several forms of interdependence between “mindful” and “routine” work. Drawing on Weick et al. (1999), the authors note that in order to confront new problems requiring innovative action, organizations tend to draw on repertoires of existing routines and recombine them in novel ways. That is, work experience that is externalized and encoded in organizational structure via routines can be a reservoir on which to draw to confront nonroutine situations requiring novel deliberations and judgment—a process well documented in the context of professional service firms (Brivot, 2011, cited in Kohler et al., 2021, p. 19). Levinthal and Rerup (2006) also note that elements of mindfulness underlie most routine actions, which might be repetitive but still require interpretation in order to encode the type of response needed. Furthermore, routines tend to evolve only once their outcomes—often ambiguous and multiple—are encoded via novel deliberations. Finally, Levinthal and Rerup (2006) observe that most organizations have routine monitoring systems in place that are precisely meant to sustain attention to important signals, supporting organizational reflexivity and deliberation when such signals emerge.

Audit represents one of the quintessential examples of such routinized monitoring of exceptions and signals, in which repetition and mechanical behavior are intertwined with more deliberative, reflexive, or indeed “skeptical” attitudes. As the following quote by Weick and Sutcliffe (2001, pp. 87–88) illustrates, the tension mindfulness-mindlessness in organizational research is cast in terms that very closely resonate with the structure-judgment tension in auditing:

Mindful moments are important if the contexts in which you operate are dynamic, ill-structured, ambiguous, unpredictable. In less dynamic contexts, mindfulness is less necessary and the economies of mindlessness are more appropriate. Mindfulness takes effort and cost; mindlessness in the form of routine can be cost-efficient.

Auditing is thus an ideal practice to explore concerns with the “routinization of mindfulness” which Levinthal and Rerup (2006, p. 506) point to as an apparent oxymoron deserving more attention.

This tension has recently been reexamined through the concept of habit, with the aim to reflect on the conditions that promote more structuration and mechanical routines, or else greater deliberation and reflexivity (Turner & Cacciatori, 2016). The notion of habit has a long history in social science and philosophy. It has been somewhat contested, in that different disciplines have promoted conflicting understandings of it. As observed by Crossley (2013; see also Camic, 1986), “habit” came to be associated with conditioned or mechanical behavioral due to “the successful colonization of the concept by the early behaviorists,” and as a result the notion was neglected by sociologists and/or replaced with habitus, most notably in the work of Marcel Mauss and Pierre Bourdieu. However, thinkers such as Merleau-Ponty and Dewey revisited and rehabilitated the notion of habit by freeing it of its behaviorist baggage and reconnecting it

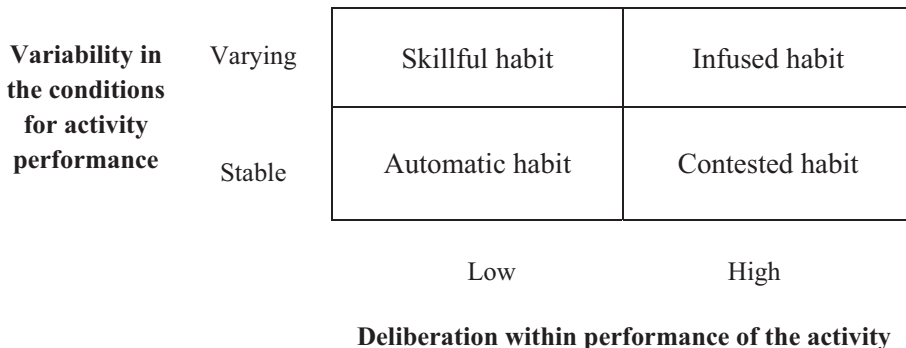


both with Bourdieu's work and with the philosophical traditions in which the notion originated. In Crossley's reading, both Merleau-Ponty's and Dewey's notion of habit are dynamic and entail reflexivity and change (2013, p. 152): "A portion of our habits is always being short-circuited, forcing reflective intervention and reworking." Thus, the notion of habit cuts across the distinction structure-judgment.

Turner and Cacciatori (2016, p. 74) build on such analysis and take their cue from Camic's (1986) definition of habit as "a more or less self-actuating disposition or tendency to engage in a previously adopted or acquired form of action." They unpack this common definition and offer a typology of habit that encompasses two dimensions: the degree of automaticity of the activity that forms the habit and the variability of the conditions under which such activity is performed (see Figure 1). Automaticity is here understood as lack of deliberation, defined by the authors, following Dewey, as "a dramatic rehearsal (in imagination) of various competing possible lines of action" (Dewey, 2002 [1922], p. 190, cited in Turner & Cacciatori, 2016, p. 81). Thus, the notion of habit as a mechanical activity, which was repudiated by Bourdieu and Mauss, is located by Turner and Cacciatori at "one extreme of the habit continuum" (p. 81) and corresponds to situations in which activities are performed under stable conditions entailing little or no deliberation, denoted by the authors as "automatic habit."

At the other extreme, when conditions are less stable and a high level of deliberation is at play, one finds "infused habit," a more flexible and adaptive habit where people display "unexpected potentialities" (p. 82). Drawing on Dewey, the word "infused" refers to the injection of thought into habit, "which solves surmountable problems around the edges of existing skills and routines, thereby enhancing the scope and adaptability of existing habits" (Winter, 2013, p. 134, cited in Turner & Cacciatori, 2016, p. 82). An example of infused habit might involve the deliberations of a North American football coach or quarterback around what "play" to make in a given situation (e.g., the team finds itself five points behind with 30 seconds to go in the fourth quarter against opponents with a particularly aggressive defense).

Between these two extremes, Turner and Cacciatori (2016) identify two additional types of habit. The first is "contested habit," which, like automatic habit, tends to be at play when conditions are stable, but entails the ability to inhibit repetition and suspend the otherwise automatic activity through a certain degree of deliberation. The second is "skillful habit," in which conditions are unstable and require adaptability, but with low levels of reflexivity and deliberation, yet "still purposive and displaying intelligence and understanding of the situation" (p. 82), as in the example, borrowed from Bourdieu, of a skilled soccer player who has to adapt quickly in the maelstrom of a flowing soccer match. Note that soccer, unlike North American football, does not have the same frequency of interruption and so affords less scope for deliberation of the kind described above. Like Crossley (2013), Turner and Cacciatori note that deliberation,



**FIGURE 1** Typology of habit. *Source:* Turner and Cacciatori (2016).

too, can become habitual—“the ‘learnt’ way of professionals reasoning about their task” (2016, p. 86) and is at play in different degrees in all the typologies of habit described above except for the more extreme case of automatic habit.

When seen in the context of these debates, audit structure and judgment appear as idealized poles in a continuum that sees audit practice entail different degrees of reflexivity and deliberation. In its more extreme version, structure, as discussed by its critics in the audit literature, can become akin to “automatic habit” (Turner & Cacciatori, 2016), which is not necessarily without intelligence as such, but represents the accumulation and externalization of experience and its encoding in organizational structure (Carpenter & Dirsmith, 1993; Levinthal & Rerup, 2006). What defines automatic habit is a lack of individual reflexivity and deliberation, the “mindless” state of being on “automatic pilot,” which Francis (1994, discussed above) saw as hindering the capacity for practical reasoning and distorting the hermeneutic nature of auditing.

While more idealized notions of audit judgment may resonate with Turner and Cacciatori’s definition of “infused habit,” in which high degrees of reflexivity and deliberation are at play to deal with nonrepetitive and complex decision problems, in practice one can think of judgments in terms of varying degrees of reflexivity and deliberation. The notion of professional skepticism, for example, which is often evoked alongside auditor judgment, can also resonate with the definition of “contested habit,” in that it entails the ability to question, rather than unreflexively accept, claims by the client and the process of their validation, in what is often—though not always—fairly repetitive work. “Skillful habit” might be encountered in situations such as RPA selecting samples for testing in a black-box fashion and thus presenting the auditor with variable tasks, necessitating the auditor to adapt to such variability and carry out and evaluate the results of that testing with sufficient purposive intelligence and understanding. On the other hand, the integration of opaque AI and professional judgment (Lebovitz et al., 2022), to the extent that it generates new uncertainty and requires to critically interrogate machine output and adapt professional judgment in the light of such output, might take us closer to the model of infused habit. That is, interactions human-machine fostered by new technologies may, under certain conditions, require or promote more complex and deliberation-rich forms of habit.

Understanding the type of habit that is cultivated within specific organizational practices can help make sense of the various degrees of reflexivity and deliberation, as well as automaticity and repetitiveness, which intertwine in the work of practitioners. That is, such a typology can help identify and nuance the particular balance of structure and judgment at play, as well as understand whether such balance is likely to be tilted in one or the other direction over time by virtue of the particular forms of habit emerging. As Turner and Cacciatori put it (Turner & Cacciatori, 2016): “Whether or not routinization at an organizational level sustains mindfulness can depend on the type of habit that predominates for participants” (p. 83).

We show that the ostensive intensification of structure that comes with the new digital tools that increasingly mediate the audit “ritual” (Pentland, 1993) has various implications for judgment, expanding and eroding judgment in different ways and under varying parameters. This, in turn, brings to the surface more nuanced dynamics of the structure-judgment interdependency. In areas of automation, automatic forms of habit appear to be promoted, so that the cultivation of the judgment required for auditors of future is at risk. Conversely, in areas of augmentation, a shift toward higher orders of habit is now invited—though not necessarily produced—in the exercise of professional judgment. Not only does augmentation intensify both structure and judgment, but the nature of the judgment being engendered is also shifting. The implications of these findings will be further developed in the discussion.

### 3 | METHODS AND DATA

The study comprised multiple methods. Documentary analysis of key practitioner and regulatory publications was undertaken alongside both ethnographic encounters and interviews with practitioners and other key actors in the field. Ethnographic encounters were undertaken throughout 2019 and included participation in a roundtable discussion on the “ethics of AI and ML” in London, held jointly by the profession’s regulator and one of the professional bodies; a networking event held by one of the fastest-growing, third-party AI software providers to audit firms; a live webinar hosted by this same software provider demonstrating their AI platform; and attendance at Europe’s largest accountancy and finance trade fair, where both global brands and start-ups showcased their latest “state of the art” technologies. Field notes were taken during and immediately after these events, which were used as a basis for discussion with the research team. These ethnographic encounters and subsequent discussions provided the context and starting point from where more focused interviews were conducted.

#### 3.1 | Interviews

Our primary source of data was 42 semistructured interviews (plus two follow-up discussions), mostly with audit practitioners, heads of audit methodology, and digital auditing or data analytic specialists at a range of audit firms. The experience levels of the practitioners interviewed ranged from a first-year junior to senior partners on the verge of retirement. Thirteen participants were female and the rest were male. While the majority of the interviewees were based in the United Kingdom, included within the sample was a senior partner from a European office of a Big 4 firm, a UK partner currently on secondment in Switzerland, and another partner who had only just returned to the United Kingdom after a prolonged period in the Hong Kong offices of his firm. Two of the firms provided live demonstrations of their latest audit technologies (in data analytics and AI), with one of these firms walking the research team, at their offices, through their use of live client data and explaining in detail their interpretations of the output. The second demonstration took place on Zoom. Audio recording of both these demonstrations was allowed as they did not capture any record of sensitive visuals. These demonstrations permitted the research team to visualize the software currently in use in performing data-driven audits and to be walked through examples of how risky transactions might be identified.

Interviews were also drawn from the regulatory body, a professional institute, and a start-up looking to penetrate the mid-tier audit market with AI technologies. These interviews afforded understanding of wider field-level dynamics and differing perspectives. One interview was held with a senior recruitment consultant who specialized in recruitment for the Big 4 firms, to gain a sense of current hiring trends and any skill requirements. A list of interviews and their affiliations can be found in the [Appendix](#).

As our focus was on understanding technological change in a broad sense, a greater proportion of interviewees was concentrated in more experienced audit professionals (senior manager, director, or partner) who could offer greater insight into their firm’s technology strategy and, importantly, into how audit practices may have changed over time. Interviews with senior associates and managers (usually with 3–5 years work experience) were held to understand their lived experiences working with the tools. With the Big 4 auditing firms clearly taking a global lead on both investing in and developing audit technologies, participants from these firms were more actively targeted. Within the Big 4, those who were at the forefront of new technology deployment were targeted where possible. Some of the senior professionals interviewed could be defined as “elite informants”—that is, “key decision makers who have extensive and exclusive information and the ability to influence important firm outcomes” (Solarino & Aguinis, 2020, p. 650).

In particular, these were audit partners who were leading the firm's deployment for the new technology or who had a firm-wide remit for technological innovation (BF1, BF20). For example, in addition to championing the use of technology within the firm, BF1 had a dedicated team within his own audits to deploy the technologies. Being the responsible individual signing off on the audit, he was keen to optimize the technology's potential on the audit while also having a close interest into how the technology was deployed, such that interpretation of the output could be made. BF14–BF19 were all engaged on audits on which their firms' latest technologies were being used. BF23 was fully engaged in rolling out the technology across the firm to audit teams, a role similarly taken by NBF5, who was the local "champion" of data analytics in his firm. PS1 led a team that wrote the code and algorithms for deployment within her organizational audits.

Consistent with guidance on how to solicit the richest possible information, a highly flexible and adaptable approach was adopted when conducting these interviews. Some of the key challenges noted by those authors around power dynamics and access to information were mediated by the lead interviewer who, having had extensive experience herself in working in the auditing profession, was able to approach sensitive topics in ways that did not put the data collection process at risk and also to read between the lines of interviewee discourse (Solarino & Aguinis, 2020). This previous experience as an auditor, mentioned at the start of each interview or, indeed, at the outset in the request for interviews, helped significantly in building a rapport with the interviewees such that they appeared less self-conscious in their responses. The lead interviewer was able to relate responses to her own experiences, building trust and familiarity with the interviewees. There was also no need for interviewees to explain acronyms or the realities and pressures of the way audits were conducted, which further helped the interviewees to respond more reflectively as opposed to descriptively.

The interviewees were recruited through personal contacts and purposive snowballing techniques. Interviews took place between March 2019 and November 2020, with two follow-up discussions held in March 2022. The majority of interviews were held face-to-face, whereas the five interviews following 2020 COVID-19 lockdowns were conducted remotely via video calls. Apart from two follow-up interviews and one phone interview undertaken, all interviews, lasting 35–100 minutes and averaging 1 hour, were recorded and transcribed. For these three interviews, detailed notes were taken and added to the data files immediately afterwards. One other interview was undertaken by questionnaire at the request of the interviewee. One author conducted and led every interview; 16 interviews were undertaken together with at least one other author. To ensure completeness and consistency of coverage, a broad interview guide was adopted for the interviews that contained the high-level themes to be explored and to allow space for interviewees to articulate themselves within their own interpretative schemes (Power & Gendron, 2015). These questions were adapted based on the role, area of expertise, and organizational affiliation of the interviewees. The interview guide covered how technology was changing audit practices (or not), the drivers behind and barriers to any such change, how technology was influencing areas requiring professional judgment (after this concept was identified repeatedly during early interviews and therefore established as a key theme), whether the skills required of audit professionals were evolving, organizational shifts taking place on account of technology, and any wider field-level dynamics relating to technological change, such as regulatory pressure or constraints. The interviews were not constrained to a specific technology in particular. Instead, the focus of the interviews was technology in the broad sense, with open-ended reflections sought on how and which technologies had engendered the greatest shifts in working practices. After the first 30+ interviews, themes started to reappear and repeat themselves and, hence, interviews were stopped at 42, with follow-up conversations subsequently organized with 2 interviewees.

## 3.2 | Data analysis

After the first 20 interviews, NVivo, the qualitative data analysis software, was utilized to code the transcripts along inductive lines, identifying key trends and changes to audit practices. Two of the authors coded several of the initial interviews “blind,” comparing notes and coding nodes to ensure a consistent understanding of the approach taken. Once this approach and initial nodes were preliminarily agreed upon, one member of the team then proceeded to code the remaining interviews. In performing the coding, guidance was taken from Gioia et al. (2013), whereby particular care was taken to “give voice to the informants” (p. 17) as much as possible such that opportunities at a later date for discovery of new concepts would be enabled. This exercise resulted in 83 first-order nodes and subnodes being generated across the 20 interviews.

Each member of the research team then read the 83 first-order nodes in detail and sought individually to group these into second-order themes, still largely emerging from the data itself (Strauss & Corbin, 1998). Meetings were held to review the potential second-order themes and to establish consensus on what these should be going forward. From the 83 first-order nodes, a number of high-level themes were identified and agreed upon as salient and worthy of further exploration (e.g., technological impacts, data, audit quality, judgment, field dynamics, epistemic factors). From this point onwards, a more specific and detailed list of interview questions was jointly drafted to act as a guide for the next set of interviews, although the broad interview guide initially developed was still deemed appropriate. The remaining 22 interviews were then conducted and subsequently coded following the same process as described above. At the end of the coding process, there were 105 nodes in total, from which 10 final second-order themes were developed. These were reviewed and agreed upon by the research team through ongoing discussions, and a final third wave of axial coding (Strauss & Corbin, 1998) was undertaken, which involved more explicit iteration between research literature and data. This final coding—which was informed more by Turner and Cacciatori’s (2016) dimensions of variability and deliberation rather than strictly by their four ideal habit types—facilitated consideration of where and how judgments were being formulated or shifted and the identification of specific modes of interplay between judgment and structure that were afforded by new tools. Hence, the focus on judgment emerged from the interviews as a key theme, with discussions and further questioning raised to clarify and probe interviewee statements and reflections. The concept of judgment was raised by most interviewees themselves alongside our questioning on technology use in practice.

While coding of text constituted a formal way of documenting and crystallizing empirical findings, it was recognized that data collection and analysis are not discrete, but rather are overlapping processes (Miles & Huberman, 1994). Care was thus taken not to overly “fetishize the transcript” and privilege the written text over the musical and emotive quality of the spoken word (Gabriel, 2019) or of the visual representations that the researchers were exposed to during the data collection phase. In this regard, care was taken to identify *how* interviewees expressed certain views or placed emphasis on certain themes rather than more superficial inferences based on word counts. Many of the interviews undertaken jointly or the demonstrations of software packages that the research team participated in served as important sources of collective anchoring and were returned to repeatedly during the various discussions held around data analysis.

In accordance with Lincoln and Guba (1985), member checking whereby “data, analytic categories, interpretations and conclusions are tested with members . . . from whom data were originally collected” is the most crucial for establishing credibility. Hence, follow-up discussions were held with two key interviewees (BF5 and PS1) in March 2022 to explore the themes and findings of the project. The two interviewees selected for this had provided some of the rich reflections in their interviews, had been repeatedly interviewed over some time (for clarification, elaboration, and saturation purposes), and also represented “opposite” ends of the spectrum:

one was responsible for writing the code to deploy more audit technology within her organization, and the other was a “recipient” of technology, offering deep reflections on its problematic use in practice and possible unintended consequences. These discussions offered further points of detail but not any substantive changes in our analyses.

## 4 | EMPIRICAL ANALYSIS AND FINDINGS

Audit technologies are currently held up, particularly by the Big 4 global audit firms, which have invested significantly in their development, as having the potential to “transform” and “revolutionize” (Ernst & Young [EY], 2020) audit work, enhance audit quality, and deliver greater value to clients while improving both efficiency and the nature of audit work, particularly at the more junior levels (Sharma, 2020). Investments are being made to support most areas of the audit process, including project management, access of electronic client data, automation of working papers and repetitive tasks, sample selection, and substantive testing. Data analytics, ML, and AI are being used predominantly in the analysis and identification of anomalous transactions within large volumes of data, which also assist in audit risk assessments.

Audit firms appear to be moving toward technologies that will “augment,” or possibly even replace, the judgment and expertise of auditors. For PricewaterhouseCoopers (PwC), claims are that its “revolutionary bot” will use AI and ML “to replicate the thinking and decision-making of expert auditors” (PwC, 2023). KPMG predicts, “The development and maturity of cognitive technologies and the ability to mimic human judgment over the next three to five years will be a game changer” (Forbes, 2018). Or, as the Audit Innovation Lead at Deloitte summarizes, “The current reality is a world in which examples of cognitive capability combine with the application of human judgement,” where the balance of human versus artificial judgment is seen as likely to shift in favor of the latter (Canell, cited in Holmes, 2018).

In contrast to this media-finessed public relations narrative, our interviews elicited a more subdued picture of technological adoption (see also Austin et al., 2021), providing an opportunity to study the phenomenon at its early stages. While data analytics are now widespread within the industry (and have been for some time, with precursors such as CAATs (see Salijeni et al., 2019 in place since the 1980s), the actual use of AI we observed was limited to experiments with simple forms of ML whose deployment was confined to select clients with information systems most amenable to such an approach. Technologies currently being deployed tend to focus on automating the more repetitive and administrative tasks (i.e., comparing bank balances to source documentation) or certain testing procedures (i.e., matching revenue transactions with their corresponding cash entries), as well as on capturing and analyzing complete sets of ledger entries for anomalous entries, the criteria of which are either predetermined by the auditor or, increasingly, machine learned.

### 4.1 | The automation of “low-level” work and the cultivation of judgment

One of the most significant areas being invested in is the automation of existing work processes, which attempts to substitute human directly with machine activity. While standardization within the audit has been a long-standing trend (Francis, 1994; Power, 1996), technological advances have enabled automation to follow on from standardization, elevating it to a key strategic priority for many firms. In the firms’ efforts to drive efficiency, due to both staff shortages and cost pressures, automation allegedly provides one means whereby a significant amount of time can be saved, cutting out numerous trivial steps (such as junior auditors filling in all of the square bracket functions in Excel) or the time-consuming task of writing up working papers:

We've now got automation, so bots that will write the working papers for our auditors. Hopefully, they [auditors] just have to do a bit of review, a bit of checking. . . . All the little things that used to take ages at the end manually are now done through technology. (BF16)

Working papers are the virtual artifacts that constitute an audit: the files that auditors use to document their work. They are a fundamental part of the audit “ritual” (Francis, 1994; Pentland, 1993) and component of the audit trail (Power, 2021). They can crystallize auditor habits through, among others, the repeated assessment of what to document and the cumulative sign-off that is core to the audit ritual. The interviewee above describes a trend toward the automatic generation of these papers, freeing up auditors from documentation time, which can in principle be redeployed to focus on perceived higher value activities such as review and evaluation. Tasks that might have entailed a fair amount of “automatic habit” are now shifted onto machines, potentially opening up the space for tasks requiring more reflection and deliberation—or so the innovation narrative goes.

However, some interviewees challenged this conclusion. The automation of working papers constitutes, we were told, a significant shift in terms of not just how auditors spend their time, but in the process of junior auditors learning the craft of the audit—the process of which is increasingly being challenged by technology (Westermann et al., 2015). One audit senior referred to this shift in terms of a loss of learning:

For each work paper, for example, you would write your procedures down, you would write your objectives and you would complete a conclusion. On the standardized work papers, the procedures are already stated as soon as you've launched a paper. The objectives are already stated, once you've launched the paper. The conclusion is a tick box exercise. You are picking your conclusion out of a list. You're not physically writing it. So . . . for that reason I definitely agree, that the learning aspect may have been lost in the standardization process. (BF22)

While the innovation narrative suggests that the automation currently at play pertains to replicable and time-consuming procedures in which little or no judgment is required, the above quote indicates that automation goes hand in hand with a further standardization of practice that limits reflective and sensemaking tasks such as setting objectives and writing conclusions, with open-ended writing being replaced by box-ticking.

Reinforcing this point, the same interviewee notes that while the automated work papers contain reminders of all the possible procedures that could be performed, encouraging thoroughness, consistency, and hence, in some ways, quality, they also have the effect of causing him, personally, to “think less” (BF22):

You go to so many training exercises, where they teach you how to test revenue, how to test trade receivables and what the correct procedures are. You're learning this throughout your whole qualification, but you don't apply it as you get into your career, because the standardized work paper just tells you what to do.

Further instances of automation included “low-level compliance work,” such as tests of detail. Interviewees expressed some concern about this shift, lamenting that trainees would no longer have to go through the same basic processes of transaction analysis that seniors had gone through earlier in their career; as a result, trainees might be at risk of not developing a “holistic” perspective of their client or of the audit:

Yes, there is concern as more and more low-level compliance work is automated. That experience by “doing” goes away so there is a concern there as to how we can continue to have audit juniors who have . . . that holistic view of the business. . . . I remember when I qualified as an auditor and there was a lot of just pulling out samples of invoices and ticking them. Okay I’ve ticked one and I don’t necessarily need to tick loads more to understand that. But it is a concern that that experience for the whole system, the whole job could be lost. (I1)

Or else, trainees may be less attuned or sensitive to what an unusual or suspect transaction might actually look like:

We still think it’s really important to have that experience and background because, first of all, you’re always on the lookout for something outside of the norm. And second of all, even if you do have a sample of 50 and you have to go to the client and ask for backup for those 50, you need that experience to go, “Yeah, that backup works.” It makes sense. I can see it. . . . So, you still need some of the auditing background. Our concern there is, are we giving people enough of that in the early years? (BF16)

This concern was echoed by another interviewee:

The other challenge, over time, is if auditors don’t do some of the nitty gritty work, how do they learn the judgmental stuff? My personal view is, going back to when I was a new auditor, there was a heck of a lot of stuff that didn’t use my intelligence that I did as a junior auditor. So, I think there’s a lot of stuff that’s quite mundane that junior auditors had to do back in the day that actually. . . . That can still help them to learn how to do an audit. (R2)

More directly, tasks that have historically been seen to require judgment,<sup>1</sup> such as the selection of testing samples, are also increasingly being automated, transferring the need for human judgment within the procedure to that as determined by machine. In some cases, the mechanism by which this occurs is not clear to the auditors themselves, raising questions regarding how the auditor will adequately *interpret* the output. For example, the following interviewee refers to the use of automation bots in somewhat “black box” terms:

The bot comes back and says, “Here are your material transactions that you need to test.” I’m not sure exactly how it works but the way that we use it, is there’s a mailbox, we send it to a mailbox and then it comes back to us saying, “Your bot has run and this is the output from the bot.” (BF14)

When questioned about how auditors, especially those at more junior levels, can be trained to question machines, one interviewee highlighted the need to work both with technology while developing judgmental capacities simultaneously:

It’s an important point. . . . We need them [auditors] to be experienced and be able to say does this feel right. But that doesn’t run contrary to technology, that’s . . . are they gaining the right experience of using the technology? (BF16)

<sup>1</sup>Whereas previous attempts to automate and standardize sampling approaches resulted in the advent of statistical sampling (Power, 1992), alternate nonstatistical approaches such as judgmental and haphazard sampling were still prevalent within practice, according to our interviewees.



Yet when pressed on this point of how new auditors would gain sufficient training and experience with machine-produced output, this interviewee—while conceding that this could be an issue—deferred, in a circular fashion, to the importance of tacit interpretations:

It's really hard to teach auditing without some experience in auditing. It's just, when I was first starting, you look at something like a bank record and you say, I've got to pick a sample to look at. How do you know what to pick? And then people will go, oh, you just know. You know? The ones that look interesting. How? And then after a few years *you just know*. (BF16; emphasis in original)

How such tacit knowledge can be acquired in the new automated environment remains very much in question.

Reinforcing this point about the ways in which technology may constrain the sensemaking of auditors, one Head of Global Methodology explained, in relation to automated tests, "Tests are limited by data sets available—they only test what is there!" He explains further that auditing should also be about thinking of and looking for what is *not* there, beyond the data sets available (BF5c). In this respect, automation risks concentrating the auditor's attention to the available data and once again, away from the bigger picture, possibly limiting the cultivation of higher forms of habit.

As another interviewee put it in the following exchange, technology is reducing the number of instances in which sampling is necessary (when a whole population of transactions can be analyzed) and decreasing auditor discretion in the choice of sampling techniques. The interviewee perceived these changes as narrowing the scope of judgment:

**BF21:** I think [proprietary software's name] is narrowing judgment and the standardization process is narrowing your judgment. You don't have to pick your sample based on judgment. They're going to test 100% of it. So, it's taken that out. Or, it will be more mandated in terms of what sampling technique you can use. Because there's a few you pick based on what you think works for your client, what you think works for your data set. And a lot of that choice is taken away through that standardization process. So I think, in that sense, it definitely narrows it.

**Interviewer:** Do you think that's a good thing?

**BF21:** It's probably a good thing for the quality of audit because, obviously, if you're using your judgment to pick a sample, you'll pick the smallest sample you can. Whereas this will mandate proper coverage testing, even if we may not like it. So I think it's better for the quality of the audit.

**Interviewer:** Do you think there are any areas where it's not good for the quality of audit?

**BF21:** Actually, maybe I take that back because sometimes you're mandated to pick a sample of revenue which is 70 out of 100 items of what you know is not a risky revenue balance. Because it will just be tick to invoice, or something like that. Whereas before you might have been able to test only 20 and get your coverage that way. Now you have to test 70, you're wasting your time doing that and not spending your time where you know the risks are. So, actually, maybe it's not such a great thing because . . . the opportunity to use your judgment is to put it where the risks are and you don't always have that choice anymore.

Thus, while limiting the scope of individual choice may help enforce thorough procedures ("proper coverage testing") when auditors might have been tempted to choose smaller samples,

automated test sample selections based on predetermined and predefined approaches might contain irrelevant detail or excessive requirements—particularly for smaller-sized clients. Thus, while automation is in general seen as an improvement for consistency and efficiency purposes, it can also lead, counterintuitively, to inefficiencies. The same interviewee spoke of instances where even more documentation time would be necessitated to justify the inapplicability of mandated procedures, after the effect. This pertained in particular to automated work papers that set out the full spectrum of potential procedures to be performed on a particular account balance, which by definition would not necessarily be client specific. She also observed how the use of new tools in some respects amounted to more box-ticking (and thus forms of habit that trend toward the automatic end of the spectrum), which reduced the “fun” of audit—running counter to the firms’ narratives of trying to improve the quality of work for auditors:

But then the fact is that now it’s spit out 20 things and I’ll have to sit down and explain 20 things when, really, I know half of them aren’t really a problem. It’s just that extra documentation that kind of takes the fun out of audit, because you’re not focusing on the actual risky . . . or you can see the risky parts, you can do those, but you also have to do all this other documentation which is just kind of tick-boxing. But that’s what it’s become now. (BF21)

In summary, the technologies underpinning automation are having a disruptive impact on the constellation of auditor habits. On the one hand, the intensification of standardization accompanying the automation of working papers, or the use of RPA for sampling, seems to lead auditors to think less and box-tick more, which is indicative of a move toward more “automatic” forms of habit. These findings run counter to the public pronouncements of, in particular, Big 4 firms, who cite automation as a way of freeing auditors to focus more on high-level work or high-risk transactions. On the other hand, in certain instances technology also prompts auditors to apply more judgment and invites habits on the high deliberation side of Turner and Cacciatori’s (2016) matrix, by demanding an interrogation of how software is coded or by prompting auditors to question machine outputs when they are not sufficiently client-specific (as in the example of mandated sample sizes discussed by BF21).

## 4.2 | Data analytics, ML/AI, and the “augmentation” of judgment

Current developments in data analytics incorporate ML/AI technologies which seek to analyze the full population of accounting journals, determine what transaction parameters are “normal” and from there, identify “anomalies” deemed to be at higher risk of error or fraud. The premise of these technologies is to perform this risk assessment on the population of entries, enabling human efforts to be focused on those transactions that do not follow a typical flow or pattern, or else violate authorization levels or other predefined parameters (e.g., weekend entries would not be expected to be posted for an organization with normal weekday office hours).

Our interviewees described the impact of this new technologically enabled risk analysis largely in terms of an expansion or enhancement of the auditor’s capabilities in comparison with extant sampling approaches, with the technology being able to scan all transactions and enabling an auditor’s attention to be directed to those transactions that might matter more:

When you think about it, there is only so much time. And I think we only have so many judgments that we can make, decisions we can make. I think it allows you to focus where you need to focus, where actually the judgments need to be made. (BF4)

Another interviewee explained that this newfound ability to focus on “relevant” areas is underpinned by a “superintelligence” that overcomes the limits of sampling:

Today the way you do it is manually. So you just look at the one out of a million and the likelihood for you to hit something is very low. So the fact that you have a superintelligence that goes and scans everything based on certain patterns that you are looking at is actually probably going to give you more meaningful outcomes. And those more meaningful outcome will help you assess your understanding of the processes and the controls. (BF20)

Such enhancement is often described in terms of “augmentation.” Indeed, as one interviewee put it, “It’s all about augmenting human judgment” (R1). The following interviewee further emphasizes the enhancement of capabilities in terms of volume and sheer capacity, describing the augmentation as a “scale up”:

*This is just a scale up.* So what I’m saying here is, it’s not that you’re trusting the [black] box. It’s actually [that] the boxes [are] augmenting your capability. So the box is helping you do what you cannot do as a human being . . . and tells you, “These are the implications.” So that is where I see the element of AI and other technologies breaking the complexity of the audit rather than replacing the auditors. (BF20; emphasis added)

Another interviewee, referring to data analytics, largely concurred:

Data analytics, in its simplest form, probably gives us the opportunity to tick a lot more invoices, test a lot more transactions, but someone still needs to be able to interpret that, as far as I can see. (NB5)

“Scaling up” infers doing the same, but much more of it. Yet, interviewees also describe the technology as performing complex tasks that an auditor cannot otherwise perform, identifying new patterns or relations that cannot be seen by the human eye or through basic software such as SQL or Excel (at least not at once and in the same timescale). The following interviewee explains the process in similar terms:

My argument is: the machine is not telling me something I don’t know, but the machine is bringing together two or three filters—more than that, more complex filters, than what I can apply humanly at the same time. And it also takes away the possibility of human error because if I were writing something in SQL or trying to do the same thing, I could go wrong. But [here, this is something] which is proven, the logic is proven. (BF23)

Some interviewees argued that technology will facilitate more skeptical mindsets and thus better judgment, by virtue of bringing more elements to the attention of the auditor—another facet of augmentation:

I can see it will only augment judgment areas. . . . When we first started getting [regulator quality] review reports, [we would get]: This is the facts . . . this is the facts, and there was a lack of professional skepticism. And the lack of professional skepticism was that you didn’t consider alternative potential outcomes. So . . . biases didn’t allow you to perceive that an alternative outcome could have been possible, which manifested itself. So I believe that technology . . . should enable us to come to

better judgments, will facilitate better judgments, because it will make more information available based on which we can feasibly suggest alternative outcomes. (BF5)

For these interviewees, AI tools augment capacity and capabilities, which then complement or support judgment. Crucially, judgment is seen by these interviewees as *external* to the tools and as “something that could *never* be automated” (BF25; emphasis added). Such a viewpoint is reflective of a more widespread affirmation of human-machine separation, where the auditor is still in control (Faulconbridge et al., 2023) and can question machine outputs and “join the dots”:

So professional judgment is outside the tool. The tool is an enabler for me to exercise the judgment I have to. (BF23)

We are at the heart of it, which means having an auditor with a skeptical mindset questioning what is in front of him is at the heart of what we do. Being helped with super capacity because it processes things while he’s sleeping away. (BF20)

So for me to have a perspective that these guys are not seeing means that I have done something not supernatural, but I’ve done something super-augmented from my capabilities to be able to give that perspective. And this is where . . . joining the dots comes in. (BF20)

Interviewees also spoke of situations whereby the tools may generate too many “false positives” (BF20) requiring new deliberations. These interviewees seemed adamant that they broadly understood the output of new tools and that they could apply a skeptical mindset to them. For these interviewees, at least, new technologies appear to increase the propensity to identify transactions needing attention from the auditor and to govern the complexity of the audit by cutting through layers of data and identifying patterns otherwise not visible, which require judgment to evaluate and decipher.

Another set of issues identified by one interviewee, which have paradoxical effects on judgments such as the application of materiality, relates to the greater accuracy yielded by new tools:

The client system pulls exchange rates from an acceptable source. We pull exchange rates from an acceptable source. Okay? There will be differences. Inevitably there are differences. We’re in a world where under the current standards these differences are difficult, [whereas] in a sample world . . . they’re difficult to deal with because we’ve never been at this level of accuracy before. (BF5)

Here, the interviewee cites differences arising from the audit firm’s technologically enabled lease valuations, with integrated foreign currency calculations, with those produced by the client’s own methods, struggling to establish how newly arisen material differences would be handled within the extant regulatory framework. How to determine such a tolerance, with the appropriate judgment in doing so, is a *new* problem posed by technology:

**BF5:** What I’m trying to say is when we introduce machines into this process which are inevitably valuable because it enables us to test many more transactions . . . We don’t have this ability to . . .

**Interviewer:** Build in the tolerance?

**BF5:** Yes. Or a professional framework for tolerance. . . . So I can see us being in a world . . . of us doing audit work which actually is probably many times better in terms of lease valuation, going back to lease details, applying a technology to validate, but still having audit issues. . . . In the absence of provision in the standards, we have to *create some sort of our own judgmental aspect* that goes in a matching process, there is inevitably some tolerance. (Emphasis added)

**Interviewer:** Well, presumably over time that will be built into the machine.

**BF5:** Yeah, okay, but now that's a very good point. Just because it's built into the machine doesn't make it right. Doesn't make it equal to professional standards. It just means it's hidden in the code. All right?

The higher accuracy afforded by new technologies and the technology gap between auditor and client require the development of new professional standards to decide when and how discrepancies that would not have arisen in a traditional audit amount to material misstatements in the client's financial statements. Thus, attempts to adopt technology to standardize analyses, arguably increasing structure, open up new spaces where human judgment is required—at least until it can be further encoded into software, potentially generating new gray areas for human deliberation.

While nowhere in our interviews did anyone speak of an instance where a program or algorithm could take over the ultimate decision-making or the drawing of conclusions from the human auditor, with AI technologies, judgments of *what* to test (taken from the AI-identified risk profiles), and equally as important, what *not* to test (the remaining population discarded as “normal”) are gradually being displaced to machines. In this respect, interviewees were not always so certain about the intelligibility of machine outputs. The transparency, and I (intelligence), of these algorithms was in question and was noted to be a crucial matter:

Explainable AI is going to be really important in auditing because we need to be really clear about if we've picked a particular thing as a test because it is risky, [we need to know] why is it risky? And making sure an auditor has applied some judgment to that. (PS1)

Perhaps unsurprisingly, this opacity has raised concerns with regulators, where extant frameworks call for the exercise of individual judgment in these very matters. As one interviewee from the UK regulator clearly set out:

Personally, I've noticed pilot usage . . . so if you're looking for odd items you might say, “I'm going to look for odd items in line with what's set out [in the standards] for general entry testing, but let's have another look and use this tool that we've got to see if this tool thinks, based on what it's seen elsewhere, if there's anything else that looks a bit odd.” But, lots of it could look a bit odd. We're laughing, but that's almost the crux of the problem; what is a bit odd? By what criteria? This technique says, based on what it's seen before. . . . Do I agree with your definition toward what a bit odd is? Is it relevant to what I'm actually doing? Can we have lots of things that look a bit odd but might be far more relevant for an internal audit? (R2)

Crucially, when confronted with uncertainty concerning machine output, interviewees tended to default to established habits (Turner & Cacciatori, 2016) and their extant knowledge of the client. Many interviewees referred to such knowledge of the client as a means by which

their evaluation and assessment of machine generated outputs were being performed, which helped them navigate the uncertainty experienced when those outputs were opaque. Indeed, “knowledge of the client” often took priority over machine-based outputs regarding judgments of what is normal or abnormal, as the following interviewee indicated:

Well, that’s the thing with the AI bits and pieces. [Firm’s proprietary software] has its own AI tool, as they call it, where it can give you some sort of suggestions for things to look at, but at the end of the day, it’s still down to the audit team, the audit manager, the audit partner to say, “*We actually know how the business operates. I know what’s weird and what isn’t.*” (NB6; emphasis added)

The following interviewee further emphasized this prioritization by indicating how the knowledge of the client can be used to set the parameters of the analyses being run, thereby tailoring the testing requirements to that knowledge:

Now, what data analytics has done in my time in the firm, in about 10 years, is that by being able to extract that population directly from the ledger, you’re able to run analytics on it to identify, *based on your knowledge of the client*, what specific risks we’d want to think about in terms of those transactions. (BF12; emphasis added)

Bespoke, contextual knowledge of the client is used by auditors to tackle the possible uncertainty associated with the use of new tools. In addition to providing auditors with the basis for manipulating the tools, thereby overriding the potential output from the technology, interviewees also spoke of how some outputs from data analysis and visualizations can assist with “getting to understand the client better” (BF10; BF14). Interviewee BF20 also referred to this earlier as helping to inform the auditor’s understanding of the client’s process and controls. Here, machine output is retained but only insofar as it builds on *existing* knowledge of the client (for a comparison, see Selten et al., 2023). The following interviewee reinforces this point by explaining situations when output is used in a purely ancillary sense, to bolster documentation:

But just having that data available, it makes it a lot easier because it’s just something that we can easily then include in our documentation and say, “This is the trend,” and [that it] makes sense with our understanding of what’s going on in the client’s environment, the environment they operate in. (BF10)

To this end, in some cases, interviewees indicated that output from technology, be these from automated bots, standardized working papers, or risk assessments, would at times simply be overridden by the auditor’s knowledge of the client, rendering the output from the technology somewhat redundant. For example:

You make a judgment *based on what you know of the client* and things you might have discussed that you don’t, necessarily, always write down. So it always comes to the end of the audit where you’re like, “Oh, I have to do this work paper,” and then you have to, kind of . . . find ways to document around it, which is just kind of fudging it away. (BF21; emphasis added)

At every point, it is very much a tool to try to make your life easier, but you still have to apply that judgment and that professional skepticism to say, “I want to test on the four highest categories of risk.” Even within then, there might be some transactions that *you just know* are fine. (NB6; emphasis added)

These quotes allude to how auditors will fall back on either existing practices or on their previous knowledge of the client in their decision-making, where output from the tools can be accommodated within the existing ways of working and knowing without too much disruption.

Overall, this section suggests that augmentation in auditing, as currently experienced, is subject to certain parameters. Technology has enabled an augmentation of sorts, by increasing capacity and capabilities, such as scanning entire populations of transactions via multiple criteria at once in order to risk analyze those populations. Yet, the implications of these enhanced capabilities for judgment manifested in more subtle ways. Uncertainty can be generated in both the means and the end—in how the technology has generated the output and also within the output itself, such as decisions concerning what to test and not to test. If machine output were passively accepted, we would be observing automation rather than augmentation (Lebovitz et al., 2022). However, our interviewees mostly indicated that extant knowledge of the client allows them to question and integrate machine output in their decisions. When uncertainty cannot be resolved, such that machine output is not reconcilable to that knowledge, we see evidence of such output simply being disregarded. In these situations, no real augmentation actually occurs (Lebovitz et al., 2022).

## 5 | DISCUSSION

Our data leads us to question the prevailing rhetoric about technological change in audit, which suggests that those parts of the audit that are highly standardized or where judgment is largely deterministic if information is available should be prime candidates for automation, whereas new tools will augment judgment in less structured areas (Kokina & Davenport, 2017; Moffitt et al., 2018), questioning what it means to apply—or not apply—judgment. It also allows us to unpack and revisit the interdependence between structure and judgment, which has long preoccupied the auditing literature from a variety of angles (Kohler et al., 2021; Power, 2003) and which the automation/augmentation agenda in auditing has brought to the fore once more.

As far as automation is concerned, our interview materials show that precious elements of judgment are at play in those repetitive tasks that are seen as tedious, low-value, and ripe for automation. Here, structure and judgment come across as more intertwined than suggested by the automation agenda. Blind to such interdependence, the implementation of automation in the name of efficiency and productivity appears to be crowding out judgment from the work of early career auditors, in several ways.

In the first place, as seen in the case of working papers, automation is taking over areas of work, which, despite being highly standardized, still required auditors to exercise some reflexivity (in the form of open-ended writing) and deliberation (in the choice of objectives and testing procedures). Even within such standard and repetitive work, elements of reflexivity and deliberation remain crucial, as exceptions inevitably arise demanding special cases to be encoded and sorted out, so that routines can be flexibly enacted through situated and context appropriate responses (Levinthal & Rerup, 2006). Removing pockets of reflexivity and deliberation in automated working papers is pushing auditors toward more box-ticking, arguably reducing mindfulness and promoting automatic habit (Turner & Cacciatori, 2016).

Second, we observed that deliberations made by bots regarding what to test are not always made sense of by auditors, while the more limited choice afforded in their sampling work was perceived by some interviewees to limit the scope for applying judgment, or to lead them to “think less.” As a result, we were told that trainees may become less attuned to what an unusual transaction might actually look like. This implies that the routine, highly structured sampling and testing practices of early career auditors are important to sustain their ability to “anomalize” (Weick & Sutcliffe, 2006) client claims and cultivate the complex forms of habit that are associated with a skeptical mindset. However, this particular interplay between

structure and judgment is now disrupted by the further standardization and automation of sampling and testing work.

Furthermore, automating low-level compliance work appeared to reduce the ability of junior auditors to make sense of the “whole system” or the “whole job,” which is arguably crucial for developing skills needed in more complex areas of work, including later in one’s career. As one interviewee noted, some early career “experience by doing” is being lost, and new tools appear instead to afford more fragmented insights, causing a loss of perspective. In other words, the “nitty gritty work” is essential to “learn the judgmental stuff,” as one interviewee from the UK regulatory body reminded us. Judgment can be and is built on more routinized work, via the “spontaneous recombination of wisdom accumulated from prior experimental learning” (Levinthal & Rerup, 2006, p. 505). As Sennett (2008) wryly points out, as clever as Mozart was, he would not have written many scores had he not spent thousands of hours at the keyboard engaged in more basic, repetitive actions in his very early years (pp. 37–38).

To sum up, important elements of judgment are variously interspersed in the audit “structure” that is currently being automated, and the related skills and habits might be lost for the next generation of auditors. This makes us question the implicit hierarchy that in the innovation discourse associates judgment with “high-level” work. Rather, the “low-level,” highly structured work of today is conducive to the high-level and more judgmental work of the future. That is, when conceived less as features of abstract and atemporal decision problems, as the innovation discourse tends to do and more as embedded in the experience of individual auditors, structure and judgment emerge not only as closely intertwined, but also as interdependent in a teleological sense: the structure of today, infused with elements of judgment, is to be understood and valued in its implications for cultivating the judgment required of auditors in the future.

When it comes to augmentation, the further structuration that accompanies the use of new technologies such as BDA and ML/AI is explicitly placed at the service of judgment, pointing to a new cooperation between structure and judgment. One hundred percent data ingestion and analysis vastly widen the “auditable” scope of the audit and, concomitantly, the range of judgments applied by auditors, who are now exposed to unseen levels of accuracy, greater ambiguity, or unworkable levels of exceptional items. New technologies are seen to expand auditor capabilities and support judgment in ways perceived as unprecedented, endowing them with a “superintelligence.” Yet this augmentation narrative, too, needs qualifying.

First, despite some real enthusiasm concerning the impact of data analytics and AI, the greater computation power that accompanies new audit tools is carving out new areas where auditor deliberations are both needed and uneasy. This was the case in the example of the lease calculator, where the superior accuracy of the tool implies substantive differences with the less sophisticated client valuations, even if the latter remain compliant with financial reporting standards, begging the question of whether and how valuation differences arising from a technological gap between auditor and client may count as material misstatements. Our interviewee here lamented the lack of a professional framework for tolerance in this kind of circumstance and affirmed the need to develop one, observing that building such tolerance into the tool would not be satisfactory as it would be equal to “hiding things into the code”—a form of erosion of judgment by structure, which our interviewee rejected. Building a sophisticated valuation tool creates new and yet unresolved problems of judgment that auditors are currently contending with. In such areas, structure and judgment are both visibly at stake, but as far as these new problems of judgment remain unresolved, the achievement of augmentation remains in question.

Technologically enabled risk analysis and the augmentation it is perceived to yield arguably represent an intensification of the “routinized mindfulness” (Levinthal & Rerup, 2006) that auditing epitomizes, with a parallel expansion of both structure and judgment. Structured practices based on data analytics and AI allow auditors to scan entire populations of transactions for abnormalities. Such new tools require an ostensibly high level of deliberation in order to



interpret and act on machine output. Interviewees tended to reject suggestions that judgment itself was being automated, claiming that “judgment is outside the tools” and that they were able to make sense of and even question machine output in new technologically assisted risk analysis. However, we also observed a degree of nervousness concerning the opacity of the ML/AI tools being pioneered. A strong theme emerging from our empirics is the anchoring of judgment on prior knowledge of the client, resorted to as the key resource to interrogate opaque AI output and reconcile it with auditors’ knowledge, integrating it into their judgment. That is, prior knowledge of the client emerges here as the key condition for true augmentation to occur (which Lebovitz et al., 2022, call “engaged” augmentation), and we found several instances in which machine output that could not be reconciled with such knowledge was simply discarded, with no real augmentation at play in such cases.

Resolving uncertainty by questioning and integrating machine output can engender new forms of deliberation and cultivate complex types of habit (Turner & Cacciatori, 2016). However, while increasing structure can engender expansive changes in the nature of the judgment exercised, today’s adopters of new auditing tools have learned their auditing skills in a pre-BDA and pre-AI world, and it is knowledge of the client developed in such a world that, by and large, allows them to question machine output today. The new cooperation between structure and judgment implied by the augmentation agenda may thus be largely dependent on skills developed in the pre-augmentation era.

Crucially, “knowledge of the client” also worked as a rhetorical resource to affirm the primacy of human auditors over machines, which our interviewees seemed to do again and again. Some of these quotes are more ideational than descriptive of practice, evoking the importance of tacit knowledge and intuition but not always able to articulate how these may be nurtured and protected in the new data-driven audit environment. This was particularly evident when interviewees sought to describe the process of augmentation AI is tasked with, noting that introducing AI simply meant scaling up information processing capabilities in order to offer the auditor more elements to subject to a skeptical mindset. This implies a quantitative rather than qualitative change in the way judgment is exercised, somewhat downplaying claims about the potential of AI for disrupting auditing as we know it. At the same time, this kind of reasoning presupposes an already skeptical mindset, but does not, to speak back to Eulerich et al.’s (2022) concerns, explain how such a skeptical mindset might be cultivated by the auditors of the future, other than affirming the need to keep human auditors in control through, for example, explainable AI. This represents something of a conundrum for audit firms: the more they automate easy tasks in order to free up space for high-level judgmental work, the less prepared auditors will be for that high-level judgmental work.

## 6 | CONCLUSION

This paper has explored the impact that new technologies are having in the domain of audit, highlighting areas of both automation and augmentation. Our case study was based in the United Kingdom and focused primarily on the Big 4 professional service firms, who are seen as being at the forefront of technological innovation in terms of audit. Our findings may not be transposable to other geographical contexts where embrace of technology might be more tentative (Goto, 2021), to other domains of financial services, to knowledge workers in general, or to mid-size or smaller audit practitioners. The empirical study was also conducted between 2019 and 2021, when data-driven audits were still in their developmental phase in some firms, so it should be recognized that the paper’s insights refer to a field in flux rather than one that has reached any kind of “settlement” (Fligstein & McAdam, 2011) in terms of technological innovation in audit. Yet, equally, technological change within audit could arguably be seen to be in a constant state of flux. In this sense, this time is not any different.

Beyond the implications for conceptualizations of audit, our findings also speak to debates on the future of knowledge work more broadly. As outlined at the outset of this paper, fourth industrial revolution discourse is gloriously bullish about the prospects for technology to reorganize the workforce, doing away with old, inefficient ways of doing things and identifying more and more opportunities for machines (Frey & Osborne, 2017; Susskind & Susskind, 2015; World Economic Forum, 2020). Some commentators even go so far as to encourage us to contemplate a world without work (Susskind, 2020) or to prepare ourselves for the singularity, a hypothetical point in time at which technological growth becomes uncontrollable and irreversible (Kurzweil, 2005). Based on survey data, macroeconomic trends, or simply blue-sky thinking, such discourse can generally be traced back to economists or technological determinists and reflects a profound aversion to any anthropological or sociological insights whatsoever. Detailed case studies such as that presented here, offering a less cognitivist or behaviorist notion of judgment, demonstrate that precious elements of the latter are still at play in areas relentlessly promoted as low value and in need of automation, while that ability to make sense of new automated tools in higher value parts of the audit process depends precisely on developing the kind of enriched forms of habit that automation risks crowding out in auditors' early careers.

Future research could usefully explore a number of different areas. More detailed case studies, with social scientific sensibilities, are needed in many work domains in order to assess the actual impacts that fourth industrial revolution technologies are having on work practices and workers. Such studies would provide a refreshing and much needed counterfoil to the often-hyperbolic discourses emanating from economists and technologists. In audit specifically, technology is advancing apace and there is an increasing role being played by hybrid teams comprised of both auditors and data scientists (see Bauer et al., 2019).<sup>2</sup> Future studies could also explore the dynamics between auditors and these new actors, particularly in terms of how expertise is increasingly distributed across time, space, and epistemic domains (Eyal, 2019). Outsourcing of work, something that also necessitates technological innovation, is important to explore in this respect. Although difficult to undertake for a number of reasons, ethnographies of auditors actually using new software and tools could yield more detailed, microlevel accounts of how professional vision (Goodwin, 1994, 1995) might be changing as a result of technological innovation in the audit space. In particular, as getting to know the client will be increasingly mediated by technology in the future, charting how this hampers or hinders the development of auditors' judgmental dispositions would be worthy of pursuit.

Furthermore, we privileged interviewees who were at the forefront of using new technologies, which was appropriate given our research questions. However, a different sampling strategy would elicit the views of smaller firms and those more distant from new technologies in order to give a fuller sense of how the audit field is impacted, or not impacted, by technological change. Finally, the extent to which new technologies actually improve audit quality or have the potential to make audit a more meaningful practice would be worth exploring (Financial Reporting Council, 2020). Many auditors have a subjective sense that audits are now better with the enhanced digital tools at their disposal, but relying on the affective dispositions of auditors themselves to assess audit quality is a risky business. Future research could elicit the views of chairs of audit committees or regulators in order to provide a more critical assessment of these claims—or investigate cases of contemporary audit failure.

<sup>2</sup>Our empirical study highlighted the increasingly important role played by data science teams at the heart of Big 4 firms that are populated by both auditors and data scientists or individuals who are competent in both. These teams increasingly support auditors in the field by ingesting client data, developing the software to facilitate testing or analytics. This suggests that audit is an increasingly distributed practice involving multiple forms of expertise and expert systems.

## ACKNOWLEDGMENTS

The authors thank participants at the 2023 Society for Advancement of Socio-Economics Annual Conference in Rio, the Global Management Accounting Research Symposium 2021, and seminars at Copenhagen Business School, Monash University, the Centre for Professional Service firms at Bayes Business School, Alliance Manchester Business School, and the London School of Economics, as well as Andrea Mennicken, John Millar, Mike Power, Alex Preda, Vaughan Radcliffe, and Mitch Stein, for comments on earlier versions of this paper. Crawford Spence is grateful for the time afforded by a Leverhulme Trust Fellowship RF-2020-0637 Audit in the Age of Machines for this project. The authors are also indebted to editor Yves Gendron and two anonymous reviewers for their insightful comments throughout the revision process.

## REFERENCES

- Abdolmohammadi, M., & Wright, A. (1987). An examination of the effects of experience and task complexity on audit judgements. *The Accounting Review*, 62(1), 1–13.
- AICPA. (2015). *Audit analytics and continuous audit: Looking toward the future*. [https://us.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/auditanalytics\\_lookingtowardfuture.pdf](https://us.aicpa.org/content/dam/aicpa/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/auditanalytics_lookingtowardfuture.pdf)
- Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons*, 29(2), 439–449.
- Alles, M. G., Kogan, A., & Vasarhelyi, M. A. (2008). Putting continuous auditing theory into practice: Lessons from two pilot implementations. *Journal of Information Systems*, 22(2), 195–214.
- Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big Data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory*, 36(4), 11–27.
- Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2018). Analytical procedures in external auditing: A comprehensive literature survey and framework for external audit analytics. *Journal of Accounting Literature*, 40, 83–101.
- Austin, A. A., Carpenter, T. D., Christ, M. H., & Nielson, C. (2021). The data analytics journey: Interactions among auditors, managers, regulation, and technology. *Contemporary Accounting Research*, 38(3), 1888–1924.
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4), 1279–1333.
- Bauer, T. D., Estep, C., & Malsch, B. (2019). One team or two? Investigating relationship quality between auditors and IT specialists: Implications for audit team identity and the audit process. *Contemporary Accounting Research*, 36(4), 2142–2177.
- Bierstaker, J., Janvrin, D., & Low, D. J. (2014). What factors influence auditors' use of computer-assisted audit techniques? *Advances in Accounting*, 30(1), 67–74.
- Bierstaker, J. L., Burnaby, P., & Thibodeau, J. (2001). The impact of information technology on the audit process: An assessment of the state of the art and implications for the future. *Managerial Auditing Journal*, 16(3), 159–164.
- Boland, C. M., Daugherty, B. E., & Dickins, D. (2019). Evidence of the relationship between PCAOB inspection outcomes and the use of structured audit technologies. *Auditing: A Journal of Practice & Theory*, 38(2), 57–77.
- Bowrin, A. R. (1998). Review and synthesis of audit structure literature. *Journal of Accounting Literature*, 17, 40–71.
- Braun, R. L., & Davis, H. E. (2003). Computer-assisted audit tools and techniques: Analysis and perspectives. *Managerial Auditing Journal*, 18(9), 725–731.
- Brivot, M. (2011). Controls of knowledge production, sharing and use in bureaucratized professional service firms. *Organization Studies*, 32(4), 489–508.
- Brown-Liburud, H., Issa, H., & Lombardi, D. (2015). Behavioral implication of Big Data's impact on audit judgement and decision making and future research directions. *Accounting Horizons*, 29(2), 451–468.
- Camic, C. (1986). The matter of habit. *American Journal of Sociology*, 91(5), 1039–1087.
- Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. *Accounting Horizons*, 29(2), 423–429.
- Carpenter, B., & Dirsmith, M. (1993). Sampling and the abstraction of knowledge in the auditing profession: An extended institutional theory perspective. *Accounting, Organizations and Society*, 18(1), 41–63.
- Carpenter, B. W., Dirsmith, M. W., & Gupta, P. P. (1994). Materiality judgements and audit firm culture: Social-behavioral and political perspectives. *Accounting, Organizations and Society*, 19(4–5), 355–380.
- Crossley, N. (2013). Habit and habitus. *Body & Society*, 19(2–3), 136–161.
- Curtis, E., Humphrey, C., & Turley, S. W. (2016). Standards of innovation in auditing. *Auditing: A Journal of Practice & Theory*, 35(3), 75–98.
- Cushing, B. E., & Loebbecke, J. K. (1986). *Comparison of audit methodologies of large accounting firms*. American Accounting Association.
- Davenport, T. H., & Kirby, J. (2016). *Only humans need apply: Winners and losers in the age of smart machines*. HarperCollins.

- Dewey, J. (2002). *Human nature and conduct*. Courier Corporation.
- Dirmsmith, M. W., & McAllister, J. P. (1982). The organic vs. the mechanistic audit. *Journal of Accounting, Auditing & Finance*, 5(3), 214–228.
- Eililfsen, A., Kinslerdal, F., Messier, W. F., Jr., & McKee, T. E. (2020). An exploratory study into the use of audit data analytics on audit engagements. *Accounting Horizons*, 34(4), 75–103.
- Eulerich, M., Pawlowski, J., Waddoups, N., & Wood, D. (2022). A framework for using RPA for audit tasks. *Contemporary Accounting Research*, 39(1), 691–720.
- Ernst & Young. (2020, January 13). *How digitalization impacts financial services companies and their audits*. [https://www.ey.com/en\\_gl/assurance/how-digital-transformation-impacts-financial-services-companies-and-their-audits](https://www.ey.com/en_gl/assurance/how-digital-transformation-impacts-financial-services-companies-and-their-audits)
- Eyal, G. (2019). *The crisis of expertise*. John Wiley & Sons.
- Faulconbridge, J., Sarwar, A., & Spring, M. (2023). How professionals adapt to artificial intelligence: The role of intertwined boundary work. *Journal of Management Studies*. Advance online publication. <https://doi.org/10.1111/joms.12936>
- Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly*, 48(1), 94–118.
- Financial Reporting Council. (2020). *AQR thematic review: The use of technology in the audit of financial statements*. [https://www.frc.org.uk/getattachment/1c1478e7-3b2e-45dc-9369-c3df8d3c3a16/AQT-Review\\_Technology\\_20.pdf](https://www.frc.org.uk/getattachment/1c1478e7-3b2e-45dc-9369-c3df8d3c3a16/AQT-Review_Technology_20.pdf)
- Fischer, M. J. (1996). “Real-izing” the benefits of new technologies as a source of audit evidence: An interpretive field study. *Accounting, Organizations and Society*, 21(2–3), 219–242.
- Fligstein, N., & McAdam, D. (2011). Toward a general theory of strategic action fields. *Sociological Theory*, 29(1), 1–26.
- Forbes. (2018, October 19). *Artificial intelligence, real breakthroughs: The practice and promise of AI in auditing*. <https://www.forbes.com/sites/insights-kpmg/2018/10/19/artificial-intelligence-real-breakthroughs-the-practice-and-promise-of-ai-in-auditing/?sh=7bbbaa7f6c10>
- Francis, J. (1994). Auditing, hermeneutics and subjectivity. *Accounting, Organizations and Society*, 19(3), 235–269.
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, 114, 254–280.
- Gabriel, Y. (2019). Placing emotion and fantasy at the heart of the researcher-researched relationship: The contribution of psychosocial studies. *Qualitative Research in Organizations and Management*, 15(2), 192–196.
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, 16(1), 15–31.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606–633.
- Goodwin, C. (1995). Seeing in depth. *Social Studies of Science*, 25(2), 237–274.
- Goto, M. (2021). Collective professional role identity in the age of artificial intelligence. *Journal of Professions and Organization*, 8(1), 86–107.
- Holmes, L. (2018, March 29). Big Four firms raising audit bar with AI. *The CFO*. <https://the-cfo.io/2018/03/29/big-four-firms-raising-audit-bar-ai/>
- Humphrey, C., & Moizer, P. (1990). From techniques to ideologies: An alternative perspective on the audit function. *Critical Perspectives on Accounting*, 1(3), 217–238.
- IAASB. (2016). *Exploring the growing use of technology in the audit, with a focus on data analytics. Request for Input*. <https://www.iaasb.org/publications/exploring-growing-use-technology-audit-focus-data-analytics>
- IAASB. (2018). *Exploring the growing use of technology in the audit, with a focus on data analytics. Feedback Statement*. <https://www.iaasb.org/publications/feedback-statement-exploring-growing-use-technology-audit-focus-data-analytics>
- Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting*, 13(2), 1–20.
- Knechel, R. W. (2007). The business risk audit: Origins, obstacles and opportunities. *Accounting, Organizations and Society*, 32(4–5), 383–408.
- Kohler, H., Pochet, C., & Gendron, Y. (2021). Networks of interpretation: An ethnography of the quest for IFRS consistency in a global accounting firm. *Accounting, Organizations and Society*, 95, 101277.
- Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies In Accounting*, 14(1), 115–122.
- Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin.
- Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgements: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126–148.
- Levinthal, D., & Rerup, C. (2006). Crossing an apparent chasm: Bridging mindful and less-mindful perspectives on organizational learning. *Organization Science*, 17(4), 502–513.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Manson, S., McCartney, S., & Sherer, M. (1997). Audit automation: Improving quality or keeping up appearances? In M. Sherer & S. Turley (Eds.), *Current issues in auditing* (3rd ed., pp. 254–272). Paul Chapman Publishing.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

- Moffitt, K. C., Rozario, A. M., & Vasarhelyi, M. A. (2018). Robotic process automation for auditing. *Journal of Emerging Technologies in Accounting*, 15(1), 1–10.
- Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. A. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics*, 167, 209–234.
- PCAOB. (2018). *Annual report*. [https://assets.pcaobus.org/pcaob-dev/docs/default-source/about/administration/documents/annual\\_reports/2018-pcaob-annual-report.pdf?sfvrsn=100daf3\\_2](https://assets.pcaobus.org/pcaob-dev/docs/default-source/about/administration/documents/annual_reports/2018-pcaob-annual-report.pdf?sfvrsn=100daf3_2)
- Pentland, B. T. (1993). Getting comfortable with the numbers: Audit and the micro-production of macro-order. *Accounting, Organizations and Society*, 8(7–8), 605–620.
- Pentland, B. T., & Feldman, M. S. (2005). Organizational routines as a unit of analysis. *Industrial and Corporate Change*, 14(5), 793–815.
- Power, M. (1992). From common sense to expertise: Reflections on the prehistory of audit sampling. *Accounting, Organizations and Society*, 17(1), 37–62.
- Power, M. (1996). From the science of accounts to financial accountability of science. In M. Power (Ed.), *Accounting and science: Natural inquiry and commercial reason* (pp. 1–35). Cambridge University Press.
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford University Press.
- Power, M. (2003). Auditing and the production of legitimacy. *Accounting, Organizations and Society*, 28(4), 379–394.
- Power, M. (2021). Modelling the microfoundations of the audit society: Organizations and the logic of the audit trail. *Academy of Management Review*, 46(1), 6–32.
- Power, M., & Gendron, Y. (2015). Qualitative research in auditing: A methodological roadmap. *Auditing: A Journal of Practice & Theory*, 34(2), 147–165.
- PwC. (2023). *Harnessing the power of AI to transform the detection of fraud and error*. <https://www.pwc.com/gx/en/about/stories-from-across-the-world/harnessing-the-power-of-ai-to-transform-the-detection-of-fraud-and-error.html>. Last accessed September 15, 2023.
- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation-augmentation paradox. *Academy of Management Review*, 46(1), 192–210.
- Raschke, R. L., Saiewitz, A., Kachroo, P., & Lennard, J. B. (2018). AI-enhanced audit inquiry: A research note. *Journal of Emerging Technologies in Accounting*, 15(2), 111–116.
- Rezaee, Z., Elam, R., & Sharbatoghlie, A. (2001). Continuous auditing: The audit of the future. *Managerial Auditing Journal*, 16(3), 150–158.
- Robson, K., Humphrey, C., Khalifa, R., & Jones, J. (2007). Transforming audit technologies: Business risk audit methodologies and the audit field. *Accounting, Organizations and Society*, 32(4–5), 409–438.
- Rose, N., & Miller, P. (1992). Political power beyond the state: Problematics of government. *The British Journal of Sociology*, 43(2), 173–205.
- Salijeni, G., Samsonova-Taddei, A., & Turley, S. (2019). Big Data and changes in audit technology: Contemplating a research agenda. *Accounting and Business Research*, 49(1), 95–119.
- Salijeni, G., Samsonova-Taddei, A., & Turley, S. (2021). Understanding how Big Data technologies reconfigure the nature and organization of financial statement audits: A sociomaterial analysis. *European Accounting Review*, 30(3), 531–555.
- Schwab, K. (2017). *The fourth industrial revolution*. Currency.
- Selten, F., Rober, M., & Grimmelikhuisen, S. (2023). ‘Just like I thought’: Street-level bureaucrats trust AI recommendations if they confirm their professional judgment. *Public Administration Review*, 83(2), 263–278.
- Sennett, R. (2008). *The craftsman*. Penguin.
- Sharma, T. (2020, July 27). How companies like EY and Deloitte are using AI to revolutionize accounting. *Global Tech Council*. <https://www.globaltechcouncil.org/artificial-intelligence/how-companies-like-ey-deloitte-are-using-ai-to-revolutionize-accounting/>
- Solarino, A. M., & Aguinis, H. (2020). Challenges and best-practice recommendations for designing and conducting interviews with elite informants. *Journal of Management Studies*, 58(3), 649–672.
- Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Sage.
- Susskind, D. (2020). *A world without work: Technology, automation and how we should respond*. Penguin.
- Susskind, R. E., & Susskind, D. (2015). *The future of the professions: How technology will transform the work of human experts*. Oxford University Press.
- Turner, S., & Cacciatori, E. (2016). The multiplicity of habit: Implications for routines research. In J. Howard-Grenville, C. Rerup, & H. Tsoukas (Eds.), *Organizational routines: How they are created, maintained, and changed* (pp. 71–95). Oxford University Press.
- Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected* (Vol. 9). Jossey-Bass.
- Weick, K. E., & Sutcliffe, K. M. (2006). Mindfulness and the quality of organizational attention. *Organization Science*, 17(4), 514–552.
- Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (1999). Organizing for high reliability: Processes of collective mindfulness. *Research in Organizational Behavior*, 21, 81–123.

- Westermann, K. D., Bedard, J. C., & Earley, C. E. (2015). Learning the “craft” of auditing: A dynamic view of auditors’ on-the-job learning. *Contemporary Accounting Research*, 32(3), 864–896.
- World Economic Forum. (2020). *The future of jobs report 2020*. [https://www3.weforum.org/docs/WEF\\_Future\\_of\\_Jobs\\_2020.pdf](https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf)

**How to cite this article:** Samiolo, R., Spence, C., & Toh, D. (2023). Auditor judgment in the fourth industrial revolution. *Contemporary Accounting Research*, 1–31. <https://doi.org/10.1111/1911-3846.12901>

## APPENDIX: INTERVIEWS HELD BETWEEN MARCH 2019 AND NOVEMBER 2020

No.	Code	Position	Organization	Specialization
1	NBF1	Senior manager	Mid-tier firm	Forensics
2	NBF2	Manager	Mid-tier firm	Technology specialist
3	NBF3	Manager	Mid-tier firm	Audit
4	NBF4	Partner	Mid-tier firm	Forensics
5	NBF5	Partner	Mid-tier firm	Audit
6	NBF6	Manager	Mid-tier firm	Audit, data analytics
7, 8	BF1a, b	Partner	Big 4	Head of audit
9	BF2	Partner	Big 4	Audit
10	BF3	Partner	Big 4	Innovation lead
11	BF4	Associate partner	Big 4	Audit and quality review
12, 13, 14	BF5a, b, c	Director	Big 4	Global methodology
15	BF6	Junior associate	Big 4	Audit
16	BF7	Partner	Big 4	Digital audit
17	BF8	Manager	Big 4	Technology risk assurance
18	BF9	Senior associate	Big 4	Audit
19	BF10	Assistant manager	Big 4	Audit
20	BF11	Senior associate	Big 4	Audit
21	BF12	Senior manager	Big 4	Audit and risk assurance
22	BF13	Manager	Big 4	Audit
23	BF14	Manager	Big 4	Audit
24	BF15	Senior manager	Big 4	Data analytics
25	BF16	Senior manager	Big 4	Transformation
26	BF17	Manager	Big 4	Data scientist
27	BF18	Senior manager	Big 4	Transformation
28	BF19	Senior associate	Big 4	Audit
29	BF20	Partner	Big 4	Digital assurance
30	BF21	Assistant manager	Big 4	Audit
31	BF22	Senior associate	Big 4	Audit
32	BF23	Manager	Big 4	Digital audit, analytics
33	BF24	Senior associate	Big 4	Audit
34	BF25	Manager	Big 4	Data analytics/science
35	BF26	Partner	Big 4	Audit

## APPENDIX (Continued)

No.	Code	Position	Organization	Specialization
36	BF27	Partner	Big 4	Technology audit
37	EP1	Director	External AI provider	Growth, sales
38, 39, 40	PS1a, b, c	Department head	Public sector	Data analytics
41	R1	Project director	Regulator	Project director
42	R2	Project manager	Regulator	IT and AQR inspector
43	I1	Director	Institute	IT faculty
44	RC1	Senior consultant	Recruitment agency	Big 4 audit recruitment

*Note:* Follow-up discussions (interviews 14 and 40) held with BF5 and PS1 took place in March 2022. At this time, interviewee PS1 had changed jobs and was now the Director of AI and Analytics at the external AI provider, where EP1 had been working. Interviews with NBF5 and NBF6 were held together. Responses for BF14 were collected via a questionnaire plus follow-up questions at the request of the interviewee, who was working in a different time zone at the time.