Auditor Automation Usage and Professional Skepticism

Christian Peters

Project Number: 2020B03

FAR Working Papers

FOUNDATION FOR AUDITING RESEARCH

2022/12 - 11

Auditor Automation Usage and Professional Skepticism

CHRISTIAN PETERS
Tilburg School of Economics and Management
Tilburg University
c.p.h.peters@tilburguniversity.edu

This version: December 9, 2022

I am indebted to my supervisors Bart Dierynck and Kathryn Kadous for their wisdom and guidance. I thank Jeremy Bentley, Eddy Cardinaels, Anna Gold, Jesse van der Geest, Martin Jacob, Maximilian Müller, Mark Nelson, Victor van Pelt, Cathy Shakespeare, David Veenman, Elien Voermans, and seminar participants at IE Business School, KU Leuven, Radboud University Nijmegen, Tilburg University, Universiteit van Amsterdam, Vrije Universiteit Amsterdam, and WHU – Otto Beisheim School of Management for helpful comments and suggestions. Furthermore, I thank Jeannette Dekkers-Akkermans for her help with the experimental sessions. Finally, I gratefully acknowledge the financial support of the Foundation for Auditing Research.

AUDITOR AUTOMATION USAGE AND PROFESSIONAL SKEPTICISM

ABSTRACT

Audit firms increasingly employ automated tools and techniques in auditing procedures. The premise of using automation is that it increases audit effectiveness and audit efficiency. For these effectiveness and efficiency gains to materialize, auditors need to use automation in an adequate manner. Regulators, however, have raised concerns that auditors may over- or under-rely on automation. I predict that auditors are subject to an automation bias and use cues from automated tools and techniques as a replacement for vigilant information seeking, thereby reducing professional skepticism when relying on automation. My findings are in line with my predictions. When auditors rely on work conducted by automated tools and techniques, they are less skeptical than when relying on the same work conducted by an audit team member. Based on psychology theory, I employ a counterarguing mindset intervention that alleviates the negative effects of automation on professionally skepticism. Finally, I also test whether a reduction in vigilance caused by automation usage spills over to subsequent tasks. I do not find evidence indicating a spillover effect. Implications for practice and theory are discussed.

Keywords: Auditing, Automation, Behavioral Mindsets, Professional Skepticism

JEL Classifications: M42, C91, O33

Data and materials availability: Available upon request.

I. INTRODUCTION

Audit firms invest unprecedented amounts in automated tools and techniques (ATT, hereafter automation) with the aim to increase both audit effectiveness and audit efficiency (e.g., EY 2018a, KPMG 2019, PwC 2019, Bloomberg Tax 2020).[1,2] For the intended benefits of automation in audit engagements to materialize, auditors should not over- or under-rely on automated tools and techniques (e.g., KPMG 2016, PwC 2017, Zhang, Thomas, and Vasarhelyi 2022). Most auditing studies focus on auditors' potential under-reliance on automation and show that auditors are sometimes reluctant to rely on new technologies (Emett, Kaplan, Mauldin, and Pickerd 2021, Cao, Duh, Tan, and Xu 2022, Commerford, Dennis, Joe, and Ulla 2022). Theory and policymakers suggest that 'over-reliance' is also a concern (e.g., Parasuraman and Riley 1997, Harris 2017, IAASB 2021a). For instance, the IAASB (2021a, p.2) suggests that overreliance on automation may result in a lack of professional skepticism. Yet, little is known about the potential consequences of auditors' over-reliance on automation for professional skepticism (e.g., IAASB 2021a, PCAOB 2022).

The first aim of this study is to contribute to fill this gap and to examine how auditors' automation usage affects their professional skepticism. Professional skepticism is a foundational construct in auditing and can be viewed as the force that drives auditors to recognize potential errors and irregularities (Nolder and Kadous 2018). My predictions are rooted in automation bias

1 KPMG (2019), PwC (2019), and EY (2018a, 2022) announced to invest US$5 billion, US$3 billion, and US$2 billion in digital transformation, respectively. Most of the investments focus on upskilling digital skills of employees, developments of technologies, and engagement in strategic alliances with tech companies such as Microsoft, IBM, and Google.

2 The IAASB (2021a, p. 1) uses the term 'automated tools and techniques' to describe all of the emerging technologies that are being used when designing and performing audit procedures today, such as artificial intelligence (AI) applications, robotics automation processes, and data analytics. Throughout this manuscript I use the terms 'automated tools and techniques' and 'automation' interchangeably. Importantly, I focus on AI-enabled automation which can use feedback and learn to adjust instead of traditional automation, where the latter is more deterministic (the technical differences are beyond the scope of this study but are discussed by Raj and Seamans (2019).

and behavioral mindset theory.[3] I predict that auditors are less skeptical towards automation compared to the same information provided by an audit team member. That is because auditors have greater trust in automation than in humans. As a result of the greater trust, auditors may engage in a premature cognitive commitment when relying on automation (cf. Langer 1989, Parasuraman and Riley 1997). This premature cognitive commitment is likely to result in vigilance reductions and hamper an auditor's cognitive processing and readiness to respond to certain issues, thereby reducing their professional skepticism.

A second aim of the study is to test a theory-based intervention that reduces the negative effect of automation usage on professional skepticism. I propose a counterarguing mindset intervention to mitigate the negative effects of automation usage on professional skepticism. Counterarguing is defined as "*the generation of arguments against the validity of information's implications*" and requires auditors to generate reasons why a proposition is not true or a state of affairs could not occur (Wyer and Xu 2010, p. 110, Xu and Wyer 2012). Counterarguing can be particularly effective to prompt professional skepticism as regulators refer to professional skepticism as an attitude that includes a questioning mind and a critical assessment of evidence (AICPA 1997 AU §316.02, PCAOB 2010a ¶7, IAASB 2021b).[4] Theory on counterarguing mindset predicts that the effect of the counterarguing mindsets is more impactful when the cognitive behavior activated by the mindset is different from the behavior that would occur in the absence of the mindset (Xu and Wyer 2012). The cognitive behavior activated by prompting a counterarguing mindset is likely more different in the case of automation than for an audit team

---

3 Automation bias is defined as "*the tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing*" (Mosier and Skitka 1996).

4 A counterarguing mindset prompt should trigger a different causal reasoning process in which auditors should be more likely to generate arguments that refute the validity of information's implications (Xu and Wyer 2012). Based on behavioral mindset theory, I propose that a counterarguing mindset prompt before the audit tasks causes the level of professional skepticism to vary.

member as a result of the greater trust in automation than humans. Hence, I predict that the difference in professional skepticism between automation and an audit team member is smaller when a counterarguing mindset is prompted than in the absence of such a prompt.

A third aim of the study is to investigate whether using automation has negative externalities on subsequent, arguably unrelated, tasks. Specifically, I investigate whether a reduction in auditor's vigilance and professional skepticism caused by automation usage, spills over to subsequent tasks conducted by that auditor, even if there is no automation involved in those subsequent tasks. Prior literature shows that mindsets tend to be sticky as mindset switching is costly (Hamilton, Vohs, Sellier, and Meyvis 2011). Auditors that have vigilance reductions as a result of automation usage may therefore face difficulties acting professionally skeptical when working on subsequent tasks. As a result, the adverse behavioral ramifications of automation usage may not only lead to performance reductions when using automation, but also in subsequent audit tasks.

I conduct an experiment with 119 professional auditors recruited at a large public university in Western Europe. The auditors first conducted a case in which I asked them to review the workpapers of an inventory counting procedure. I employ a 2×2 between-subjects design to test my predictions. The first manipulation varies whether the workpaper of the inventory count is prepared by automation or by an audit team member. Inventory counting procedures are relatively structured tasks. I choose for such a structured task for two reasons.[5] First, Abdolmohammadi (1999) documents that two-third of substantive tests are structured audit tasks, whereas only one percent of audit tasks is classified unstructured. Second, structured audit tasks are the first-order

[5] Task structure is defined as the level of specification of what is to be done in a task (Simon 1973).

candidate for being automated (Zhang *et al.* 2022). The second manipulation varies whether a *counterarguing* mindset or *no* mindset is prompted to auditors. After the inventory counting task, auditors had to audit the client's step-one analysis of a goodwill impairment test, adapted from (Kadous and Zhou 2019). In this task, there were no differences between conditions. Instead, the task was used to test a potential spillover effect arising from the manipulations. In the audit of the goodwill impairment test, auditors had to judge the reasonableness of the fair value, state the (skeptical) action they would take, and list reasons for being skeptical or the additional evidence they would require.

I find that, absent a counterarguing mindset intervention, auditors are less skeptical when they rely on work conducted by the audit firm's automated tools and techniques than when relying on the same work conducted by an audit team member. Next, I find that a counterarguing mindset intervention weakens the negative effects of automation on professional skepticism. Finally, I investigate whether reductions in professional skepticism that are caused by automation usage also spill over to subsequent unrelated tasks. I do not find any evidence indicating a spillover effect.

This study extends two streams of literature. First, this paper contributes to a nascent but growing stream of literature that focuses on the adoption of technology in the auditing profession (e.g., Munoko, Brown-Liburd, and Vasarhelyi 2020, Christ, Emett, Summers, and Wood 2021, Commerford *et al.* 2022). Most papers in this area focus on the technical capabilities of technology (e.g., Yoon, Hoogduin, and Zhang 2015, No, Lee, Huang, and Li 2019). However, in comparison to the technical capabilities of technology, much less is written about the behavioral ramifications of technology. This study seeks to fill this void. One notable exception is Commerford *et al.* (2022), who find that auditors tend to under-rely on algorithmic advice versus human advice when auditing complex estimates, especially when management uses objective inputs. My study differs in several ways from Commerford *et al.* (2022), with the focus of my study on over-reliance being the most remarkable difference.[6] The warnings of auditing regulators and standard setters against potential overreliance on automated tools and techniques highlight the importance of investigating over-reliance (e.g., Harris 2017, IAASB 2021a).

Second, I contribute to the literature on professional skepticism in auditing (e.g., Nolder and Kadous 2018). With the emergence of automated tools and techniques such as data analytics, artificial intelligence, and robotic process automation, auditors increasingly have to exhibit professional skepticism to information prepared by those objects (Olsen and Gold 2018). Olsen and Gold (2018, 132) mention that the research question whether professional skepticism may be exercised differently toward a person versus technology is an important one and has hitherto not been investigated. Despite these claims, to the best of my knowledge, no research has yet investigated the effects of automation usage on auditors' professional skepticism. My study shows that these effects are negative, but these negative effects can be mitigated when auditors are prompted with a counterarguing mindset.

My findings are potentially relevant to regulators (such as the PCAOB), policymakers (including the IAASB), and audit firms. That is, for the potential advantages of automation to materialize, it is important that auditors' automation usage is based on thorough analysis of auditors' cognitive and motivational decision-making processes. Despite that regulators worry that auditors may rely too much on automation (e.g., Harris 2017, IAASB 2021a), the behavioral ramifications of auditors' automation usage have only recently started to draw attention from researchers. As a result, we know little about potential negative effects of automation usage on

6 Next to that, I investigate the effect of a counterarguing mindset intervention and a potential spillover effect. Furthermore, Commerford *et al.* (2022) focus on complex accounting estimates, whereas I focus on structured audit tasks. Implications of differences are discussed in Section 2.3.

professional skepticism. My study shows that automation usage may have a negative effect on professional skepticism. However, I also evaluate a theory-based intervention that addresses this concern. Regulators and audit firms can use this intervention to alleviate professional skepticism reductions when auditors use automated tools and techniques. An important caveat is that it is not an aim of this study to run a horserace between auditors and automation. My experiment does not lend itself to draw valid conclusions with respect to such questions. Rather, I examine the effect of automation usage on professional skepticism and build from literature drawn from auditing, psychology, and the management sciences. Moreover, I provide tools that audit firms can use to overcome the identified problems that auditors potentially face when working with automation.

II. BACKGROUND LITERATURE AND HYPOTHESIS DEVELOPMENT

2.1. Adoption of Automated Tools and Techniques in the Auditing Profession

Audit firms adopt automated tools and techniques to increase both audit effectiveness and efficiency (EY 2018a, KPMG 2019, PwC 2019). Key benefits of automation in the audit environment are that automation allows auditors to process an entire population of transactions instead of a sample (No *et al.* 2019), incorporate Big Data from social media websites with audit evidence (Yoon *et al.* 2015), mine large amounts of unstructured and structured data (Harris 2017), and share valuable insights with clients (Austin, Carpenter, Christ, and Nielson 2021). Research shows that in certain aspects of the audit engagement, the usage of automation leads to performance gains. For instance, Christ *et al.* (2021) find that the use of drones and automated counting software improves audit efficiency, audit effectiveness, and documentation quality in inventory counting procedures.

Although automation may lead to performance gains in the audit, auditor expertise may not easily be fully replicated by automated tools and techniques (e.g., KPMG 2016, Zhang *et al.*

2022). Therefore, audit firms emphasize that automation will not replace auditors but enhance their efficiency and effectiveness. That is, auditors, in the end, make the critical decisions and offer key analysis and insights (KPMG 2016, PwC 2017, Zhang *et al.* 2022). For instance, Christ *et al.* (2021) demonstrate that even in relatively objective tasks such as automated inventory counting using drone technology, auditors are involved to (i) ensure that the images taken by the drones are collectively comprehensive (to ensure *completeness*) and mutually exclusive (to ensure *existence*), (ii) verify whether the counting algorithm functioned well, and (iii) follow up with the client on discrepancies.[7,8]

As auditors' judgment is still needed even though tasks are automated (e.g., KPMG 2016, Zhang *et al.* 2022), it is important that the use of automation by auditors is based on thorough analysis of auditor cognition and decision-making processes. When adopting automation, many audit firms, audit regulators, and academics focus on gains in audit efficiency and audit effectiveness that can be achieved through adopting automation (e.g., IAASB 2017, EY 2018b, Christ *et al.* 2021, Austin *et al.* 2021). However, the potential benefits of automation may not (fully) materialize if there are unintended behavioral ramifications as a result of the adoption. Audit regulators have already expressed concern that auditors may 'over-rely' on automation in audits (Harris 2017, IAASB 2021a). For instance, in a speech to the PCAOB/AAA Annual Meeting, PCAOB board member Harris (2017) stated that "*[a]uditors should take care that they are not*

[7] Also in other fields, the combination of humans and AI outperforms either one of them alone, even in very objective tasks such as chess. For instance, a typical example of man versus machine is the 1997 chess match between IBM's "Deep Blue" and then world champion Garry Kasparov, ultimately won by "Deep Blue". However, it is not that widely known that a team of both a person and a computer outperformed either another computer or an expert chess player alone (Cassidy 2014).

[8] Related research in financial accounting indicates that human judgment is still essential to augment machine-based models. Specifically, Costello, Down, and Mehta (2020) find that lenders who rely on machine-generated credit scoring models, perform better when they have discretion to adjust the machine-based model when assessing the creditworthiness of opaque borrowers. Also in auditing, a main consideration is that auditors can bring their intuition, judgment, creativity, and experience to interpreting the data, leading to deeper insights than those of AI alone (KPMG 2016, PwC 2017).

over relying on data analytics. As powerful as these tools are, or are expected to become, they nonetheless are not substitutes for the auditor's knowledge, judgment, and exercise of professional skepticism."

Despite worries that auditors may 'over-rely' on automation, prior literature also finds that decision-makers may under-rely on technologies, even if they outperform human decision-makers (Dzindolet, Pierce, Beck, and Dawe 2002, Dietvorst, Simmons, and Massey 2015, 2018). Some studies have specifically investigated auditors' reliance on technologies and find that auditors tend to under-rely on technologies. First, Commerford *et al.* (2022) show that, when auditing complex estimates, auditors rely less on artificial intelligence when client management uses structured estimation processes. Second, Emett *et al.* (2021) show that engagement reviewers judge audit procedures conducted with data & analytics tools to be of lower audit quality as they entail less effort by the auditor. Third, Cao *et al.* (2022) find that the negative effects of inspection risk on reliance on data and analytics are alleviated by prompting auditors with a growth instead of a fixed mindset.

2.2. Automation Bias

Literature in human factors and organizational behavior has examined conditions for decision-makers to effectively use automation and suggests that there may be detrimental performance effects as a result of automation usage.[9] Parasuraman and Riley (1997) posit that decision-makers

[9] Merriam-Webster defines automation as an "automatically controlled operation of an apparatus, process, or system by mechanical or electronic devices that take the place of human labor." Automation focuses on streamlining repetitive, instructive tasks. Examples of traditional automation include the autopilot in an airplane, auto-generation of marketing e-mails, and automated production lines. Whereas automation is manually configured – meaning that automation works based on preprogrammed workflows, scenarios and the like – artificial intelligence goes beyond automation by mimicking and eventually superseding human intelligence and actions. Although there are differences between the two concepts, such as the usage of data, audit firms mainly use artificial intelligence effectively to automate audit procedures such that auditors can focus on higher-level tasks. This type of AI-enhanced automation is used by audit firms (e.g., KPMG 2016).

may either use, misuse, disuse, or abuse automation.[10] A well-documented bias that may particularly arise when using automation is the *automation bias* (Mosier and Skitka 1996, Parasuraman and Manzey 2010). Automation bias is defined as "*the tendency to use automated cues as a heuristic replacement for vigilant information seeking and processing*" (Mosier and Skitka 1996). More specifically, decision-makers have a tendency to over-rely on automation, resulting in *errors of omission* (i.e., failure to notice problems) and *errors of commission* (i.e., act on incorrect advice given by automation). Two main factors reinforce the occurrence of automation bias (Mosier and Skitka 1996, Parasuraman and Manzey 2010). First, decision-makers, including auditors, tend to conserve their cognitive resources (e.g., Hobfoll 1989, 2001, Dierynck and Peters 2021). Second, decision-makers tend to rely more on automation than on another person under some conditions (Dijkstra 1998, Dijkstra, Liebrand, and Timminga 1999, Logg, Minson, and Moore 2019). When decision-makers rely more on automation than on another person, decision-makers tend to develop a premature cognitive commitment when using automation, which affects their subsequent attitude towards the automation (Langer 1989, Parasuraman and Riley 1997) That is, when decision-makers over-trust automation and aim to conserve cognitive resources, this causes them to engage in mindless behavior and an inappropriate allocation of attentional resources leading to a loss of situational awareness and reductions in vigilance (e.g., Parasuraman and Manzey 2010).

2.3. Automated Tools and Techniques and Auditors' Professional Skepticism

With the adoption of automation, auditors increasingly have to apply professional skepticism to information prepared by automation (Olsen and Gold 2018). I investigate whether auditors tend to

[10] *Use* of automation refers to the voluntary activation or disengagement of automation, *misuse* refers to the overreliance on automation, which may result in behavioral biases, *disuse* refers to the neglect or underutilization of automation, and *abuse* refers to the implementation of automation without due regard for the consequences for human performance (Parasuraman and Riley 1997).

over-rely on automation, and whether this results in a reduction of professional skepticism when tasks are conducted by automated tools and techniques. Regulators, researchers, and audit methodologies emphasize the importance of exercising an appropriate level of professional skepticism (e.g., Nelson 2009, PCAOB 2010a ¶7, Quadackers, Groot, and Wright 2014). Yet, audit regulators identify a lack of professional skepticism as a root cause of audit deficiencies (e.g., IFIAR 2018). Professional skepticism is often described as a requirement of due professional care (PCAOB 2010a ¶7) and consists of the need to maintain a questioning mind and critically assess audit evidence throughout the planning and performance of an audit (IAASB 2012 ¶13, PCAOB 2003 ¶13, PCAOB 2010b ¶7). Appropriate exercise of professional skepticism is essential for identifying and responding to conditions that indicate material misstatement and reduces the risk of (i) overlooking unusual circumstances, (ii) overgeneralizing when drawing conclusions from audit observations, and (iii) using inappropriate assumptions in determining the nature, timing, and extent of the audit procedures and evaluating the results thereof (IAASB 2012 ¶15).

Professional skepticism comprises both of a skeptical attitude and a skeptical mindset (Nolder and Kadous 2018). Whereas a skeptical attitude is typically defined as a stable individual trait (e.g., Hurtt 2010, PCAOB 2010a ¶7, Quadackers *et al.* 2014), a skeptical mindset is typified as a state which can be aroused by situational factors (e.g., Hurtt, Brown-Liburd, Earley, and Krishnamoorthy 2013, Bauer 2015, Robinson, Curtis, and Robertson 2018, Kadous and Zhou 2019). A salient situational factor is whether the work is conducted by a person or by automation (Olsen and Gold 2018).[11] If auditors tend to 'over-trust' imperfect automated tools and techniques

[11] This is also consistent with the Elaboration Likelihood Model of Persuasion (Petty and Cacioppo 1986). In this model, the persuasiveness of the source (i.e., automated tools and techniques vs. auditor) is a peripheral cue that may affect the auditor's attitudes toward the work (cf. Dijkstra *et al.* 1998). Especially when auditors are less motivated or unable to judge work on its contents, auditors may base their decision on a peripheral cue such as the persuasiveness of the source.

(cf. Harris 2017), this likely causes auditors to make a premature cognitive commitment, resulting in an attentional bias in which they engage in less cognitive processing. Cognitive processing is an important determinant of an auditor's ability to exercise appropriate skeptical judgment, especially when tasks require deeper processing (Griffith, Hammersley, Kadous, and Young 2015, Griffith, Kadous, and Young 2016, Nolder and Kadous 2018). If an auditor's skeptical judgment is inhibited by the use of automation, a deterioration of an auditor's intentions and skeptical actions is likely to follow (Nelson 2009). This leads to the following hypothesis.

HYPOTHESIS 1: Auditors exhibit less professional skepticism when they rely on work conducted by automated tools and techniques compared to work conducted by another auditor.

Given the richness of decision-making environments, it is not surprising that prior literature has arrived at different predictions than mine. I highlight three reasons why my study does not undermine other predictions, but instead complements them. First, many prior studies compare an individual's reliance on automation to reliance on one's own judgment (e.g., Dzindolet *et al.* 2002, Dietvorst *et al.* 2015, 2018). An important result from the decision-making literature is *egocentric discounting*: individuals underweight the advice of others compared to their own judgments when making decisions as a result of egocentrism (Yaniv and Kleinberger 2000; Logg *et al.* 2019). As a result, automation reliance is related to decision-makers' estimates of the trustworthiness of automation relative to estimates of *their own* ability, which is potentially subject to egocentric discounting and overconfidence (Logg *et al.* 2019).

Second, individuals tend to have a "*perfect automation*" schema (Dzindolet *et al.* 2002, Madhavan and Wiegmann 2007). A perfect automation schema is conceptualized as cognitive beliefs that automation will perform with near-perfect reliability and individuals that have such a schema are less-forgiving when automation errs (Merritt, Unnerstall, Lee, and Huber 2015). This *all-or none* thinking with respect to automation performance may cause individuals to under-rely on automation when making judgments or forecasts about the future, as the future is inherently probabilistic. Many prior studies that document under-reliance on automated tools and techniques involves probabilistic forecasts about the future (Eastwood, Snook, and Luther 2012, Dietvorst *et al.* 2015, 2018, Commerford *et al.* 2022).

Third, most studies investigate reliance on automated tools and techniques by asking individuals to report the degree to which they wish to rely on automated tools and techniques (e.g., Dietvorst *et al.* 2015, 2018). However, this should not be confused by Mosier and Skitka's (1996) automation bias, where individuals tend to heuristically rely on automation. There is a difference between being consciously asked to what extent one wishes to rely on automation (i.e., a conscious decision) and using heuristics when one is actually relying on automation (i.e., partially an unconscious process). All in all, prior literature shows that these conditions are important in determining reliance on automation.

2.4. Joint Effect of a Counterarguing Mindset and Automation on Auditor Professional Skepticism

To mitigate the negative consequences of automation on auditor professional skepticism, I propose prompting a *counterarguing mindset*.[12] Counterarguing is defined as "*the generation of arguments against the validity of information's implications*" (Wyer and Xu 2010, p. 110). Counterarguing requires auditors to generate reasons why a proposition is not true or a state of affairs could not occur (Wyer and Xu 2010, Xu and Wyer 2012). Xu and Wyer (2012) find that these mindsets can

[12] Sets of cognitive processes that produce a disposition or readiness to respond to a particular matter can be characterized as mindsets (Gollwitzer 1990).

be situationally induced and reflect the activation and use of cognitive procedures in subsequent unrelated situations. Specifically, counterarguing mindsets activate cognitive behavior that leads to a tendency to refute the validity of assertions. Such a mindset persists in subsequent tasks, even if they serve a different purpose. For instance, Xu and Wyer (2012) find that individuals that watch a political speech by a politician they opposed are less likely to consider a product in subsequent commercial breaks.[13]

The effect of the counterarguing mindsets is more impactful when the cognitive behavior activated by these mindsets is different from the behavior that would occur in the absence of these mindsets (Xu and Wyer 2012). In the absence of a mindset prompt, I predict that auditors trust automation to a greater extent than an audit team member and as a result engage in less cognitive processing (cf. Hypothesis 1). Given that prompting a counterarguing mindset activates cognitive behavior that leads to a tendency to refute the validity of subsequent assertions, the difference between the cognitive behavior activated by the mindset and the cognitive behavior in the absence of the mindset is greater when auditors rely on work conducted by automation than by another person. This implies that a counterarguing mindset weakens the negative relationship between automated tools and techniques and professional skepticism. This leads to the following hypothesis.

[13] Closely related to a counterarguing mindset is *counter-explanation*. Counter-explanation is conceptualized as generating reasons why a certain assessment may not be true. Prior literature has studied the effect of counter-explanation in auditing and financial accounting settings. For instance, generating and reading counter-explanations reduced auditors' likelihood assessments for suggested causes in analytical review tasks. Kadous, Krische, and Sedor (2006) show that financial analysts that generate counter-explanations make less optimistic forecasts, and show that this effect is alleviated when generating counter-explanations is relatively difficult for an analyst, thereby providing an important boundary condition. A key difference between a counterarguing mindset and counter-explanation is that a counterarguing mindset is prompted by unrelated statements whereas counter-explanation refers to explaining why a focal assertion may not be true. Counter-explanation may thus directly impact a decision-maker's assessments of something, whereas a counterarguing mindset is prompted by unrelated statements and should only situationally affect a decision-maker's assessments of something. In my setting, as counter-explanation could cause auditors to form more pessimistic perceptions about automation that may lead to disuses, and hence be harmful in other domains.

HYPOTHESIS 2: The negative effect of automated tools and techniques usage on professional skepticism is weakened when auditors are prompted with a counterarguing mindset.

2.5. Mindset Spillover to Distinct Subsequent Tasks

As automation can lead to increased audit efficiency and increased audit effectiveness (e.g., Christ *et al*. 2021, Austin *et al*. 2021), it might be an optimal strategy for an auditor to devote less attention to tasks conducted by automation when it is free or nearly free from errors or when the cost of an error is sufficiently low such that it is acceptable. In that case, devoting less attention may result in a more efficient process but is unlikely to lead to significant reductions in audit quality. However, I propose that the reduced professional skepticism imposed by automation may spillover to subsequent distinct tasks. Prior studies in auditing show that judgments from unrelated tasks can spillover to other judgments (Phillips 1999, Piercey 2011, Van Rinsum, Maas, and Stolker 2018). For instance, Van Rinsum *et al.* (2018) find that using a disclosure checklist causes auditors to have higher levels of pro-client bias in domains distinct from those that the disclosure checklist is informative about. Hence, if automation is nearly flawless and the costs of an error are sufficiently low, a potential spillover to other tasks may still result in audit quality reductions.

The theoretical buildup to Hypothesis 1 highlighted that auditors engage in a premature cognitive commitment when working with automation. As a result, this premature cognitive commitment is likely to hamper an auditor's cognitive processing and readiness to respond to certain issues. Although auditors are not locked into a single mindset and optimal decision-making may require mindset switching, Hamilton *et al.* (2011) show that mindset switching is costly. That is, they argue that mindset switching diminishes self-regulation resources, which are limited for auditors, like other decision-makers (Baumeister 1998, Baumeister, Bratslavsky, Muraven, and

Tice 1998, Mullis and Hatfield 2018, Hurley 2019, Dierynck and Peters 2021). When switching mindsets, auditors need to override habitual, natural, or dominant responses and this taxes their self-regulatory resources. As a result, spillover effects from automation usage may be induced in two different ways. First, mindsets induced by automation may be "sticky" and a mindset imparted in automated tasks may carry over to audit tasks where no automation is involved (cf. Wyer and Xu 2010, Hamilton *et al.* 2011). Second, auditors may switch mindsets and lose self-regulatory resources that are needed to maintain cognitive focus, complete complex tasks, and make decisions (cf. Mullis and Hatfield 2018). This leads us to hypothesize that when auditors have conducted tasks using automation before conducting a subsequent task, this causes them to exercise less professional skepticism in that task. In other words, I predict that the professional skepticism reduction from a task relying on automation spills over to subsequent tasks that are not conducted by automated tools and techniques.

HYPOTHESIS 3: Auditors exhibit less professional skepticism in a subsequent unrelated task when they relied on automated tools and techniques in a previous task.

III. METHOD

3.1. Participants

119 professional auditors were recruited during sessions of a part-time professional accounting education program at a large public university in Western Europe.[14,15] Auditors were provided with a web-based experiment that was developed using Qualtrics software. Auditors were

[14] More specifically, I recruited participants during lectures of the Post-Master Accountancy program. A Post-Master Accountancy program is a program that auditors follow part-time (usually on Fridays) to obtain a public accounting license equivalent to CPA.

[15] The Institutional Review Board (IRB) at the author's institution approved the experimental study in this paper.

informed that the experiment would approximately take between 30 and 45 minutes. Most auditors were male (n = 79, 66.39 percent), had an average work experience of 1.80 years (st. dev. = 0.72 years), and were on average 24.84 years old (st. dev. = 1.84 years). The sample consists of 102 staff auditors, eleven senior staff auditors, and six auditors that classify themselves as 'Intern/Trainee.'

3.2. Experimental Case and Procedures

I presented auditors with a scenario in which they assume the role of an auditor at a year-end audit of a client operating in the agriculture industry. Specifically, I told auditors that they were responsible for auditing the inventory audit procedures and the client's step-one analysis of a goodwill impairment test.[16] These inventory audit procedures consist of the counting procedures of the client firm's livestock to provide assurance over the existence (i.e., all inventory exists and is real) and completeness (i.e., all inventory owned is reported) of the inventory.[17] The scenario adopts a four-step process as put forward by Christ *et al.* (2021). That is, the client conducts physical counts of the inventory, while the auditor has observed these physical counts. As the livestock is dispersed across wide areas, nonstationary, and large, this makes counting manually difficult. As in Christ *et al.* (2021), the audit team captures images by flying drones over the agricultural assets (cf. PwC 2016) and processes the images to ensure only relevant assets are captured. Next, in Christ *et al.* (2021), they apply automated tools and techniques to count the livestock using the *Countthings algorithm*.[18] Auditors in the *Automation* conditions are informed that the livestock is counted by the algorithm. In the *Human* conditions, an audit team member counts the livestock (which will be discussed in more detail in Section 3.3.). Finally, the auditors

16 PwC (2016) estimates that companies spent $127.3 billion on inventory management in 2015.

17 Abdolmohammadi (1999) documents that auditors rate substantive audit procedures related to the inventory counting procedures as relatively structured tasks.

18 For a demonstration of the *Countthings* algorithm, see https://countthings.com/ (last accessed on August 17, 2022).

were tasked to review each image after the source (either *Automation* or *Human*) had reported the initial count.

Auditors were tasked to conduct the final step in this four-step process. That is, they were presented the count by the preparer, and had to manually identify whether assets had been missed or items had been incorrectly included. To do this, they would have to perform recounts. Auditors received four aerial photos to review. Auditors were informed that the inventory of the livestock is material to the financial statements, both quantitatively and qualitatively due to the moderate likelihood of management fraud in inventories. As they were also informed that misstatements are likely systematically biased into one direction, it was important that *any* deviation from client's reported numbers was detected and discussed with the audit team. Next, they received background information about the client and the audit procedures. Directly after receiving the background information, auditors were subject to the counterarguing mindset manipulation (discussed in Section 3.3.). After being subject to the counterarguing mindset manipulation, auditors were provided an example of an inventory count and were subject to the source manipulation (also discussed in Section 3.3.). Next, auditors continued to the main task.

In the main task, auditors had to review the inventory counting procedures that were already prepared by either automated tools and techniques or by an audit team member, depending on the condition they were in. Specifically, they had to review four photos of livestock that was captured by a drone (see also Appendix A). In each of the photographs, auditors were provided with an initial number and had to verify whether this was correct (*yes/no*). Only if auditors selected "*no*", they were asked what the correct number should be. Whereas in the first three photos the correct number was provided by the workpaper preparer (*Automation* or *Human*), there were six seeded errors in the fourth photograph. The review of the inventory counting procedures can be characterized as a relatively structured task (Abdolmohammadi 1999). I use a structured task for two reasons. First, Abdolmohammadi (1999) shows that most substantive audit tasks are structured (i.e., 67% structured, 32% semi-structured, and 1% unstructured). Second, Zhang *et al.* (2022) argue that structured tasks are potentially automated, while automating unstructured tasks is less likely. Hence, when auditors use automation, this is more likely to occur in a structured task.

After reviewing the inventory audit procedures, auditors were tasked to audit the client's step-one analysis of a goodwill impairment case, which was adapted from Kadous and Zhou (2019).[19] I use the goodwill impairment case to measure a potential spillover effect (i.e., Hypothesis 3). In the case, the client uses a discounted cash flow (DCF) model to estimate the fair value of a business unit. Auditors' task was to evaluate the projections for future revenues and form a preliminary conclusion about the reasonableness of the fair value of goodwill (cf. Kadous and Zhou 2019). Auditors were informed that the firm's internal valuation specialist had already determined that the DCF model was appropriate from the client and the team had tested the mathematical accuracy of the model and found no exceptions. The only parts that the audit team still needed to evaluate were the five-year projections of revenues and the discount rate used in the DCF model. Auditors were tasked with evaluating the revenue projections, whereas an internal specialist would audit the discount rate. In the case, there were five seeded issues indicating that the revenue projections may have been too rosy and fair value is overstated, and some of these issues were in the discount rate section. Even though auditors were not explicitly asked to audit the discount rate assumption, auditing standards require them to do so (Kadous and Zhou 2019).[20]

[19] Out of the four important assumptions that underlie the client's discounted cash flow model in Kadous and Zhou (2019), I use only two: the projections for future revenue and the estimated discount rate. This was to shorten the case to provide auditors with enough time to finish the experiment.

[20] Kadous and Zhou (2019) report that auditing standards AU sec. 336 *Using the Work of a Specialist* and International Standards on Auditing (ISA) 620, *Using the Work of an Auditor's Expert*, among others require auditors to obtain an understanding of the methods and assumptions used by the specialist.

There were both surface- and deep-level cues, depending on the amount of cognitive processing needed to find the issues. At the end of the case, auditors were asked to judge the reasonableness of the case, state what (skeptical) action they would take following the case, and identify the reasons for doing so, or the additional evidence they wanted to request from the client.

After the tasks, auditors were provided a post-experimental questionnaire, in which they were asked about the process, personality, and demographics. The process variables include manipulation checks, attention checks, and questions about how auditors felt during the experiment. The other variables include questions about feelings about the auditing profession and the Hurtt professional skepticism scale (Hurtt 2010). Demographic variables include age, gender, work experience, position in the firm, and certifications. Figure 1 shows the instrument flow.

3.3. Independent Variables

I conduct an experiment with a 2×2 between-subjects design. I manipulate whether the preparer of the working paper (i.e., the source) is either a human colleague (*Human*) or the working paper is prepared by automated tools and techniques (*Automation*) using a vignette description.[21,22] Using a vignette description is in line with most of the studies manipulating automation vis-à-vis human (e.g., Dzindolet *et al.* 2002, 2003, Dietvorst *et al.* 2015, 2018, Castelo, Bos, and Lehmann 2019, Logg *et al.* 2019, Commerford *et al.* 2022). Specifically, both *Human* and *Automation* are described in identical terms, except for them being named as *the audit team member* and *the counting algorithm*. Panel A of Table 1 provides an overview of the source manipulations.

---

21 An important design choice is that I compare the initial count conducted by automation with an initial count conducted by a human colleague instead of one's own initial count. In this case, I address the concern that individuals underweight others' advice due to egocentrism, a robust result from utilization of human advice (Yaniv and Kleinberger 2000). My results in Hypothesis 1 are thus unable to speak to comparisons between an auditor's choice to conduct a task by oneself or by automation (see also Section 2.3.).

22 Although I do not specify the rank of the audit team member in the experimental case, Abdolmohammadi (1999) reports that most substantive procedures related to inventory counts are conducted from the staff level on.

The second manipulation varies whether auditors are prompted with a counterarguing mindset intervention (i.e., the generation of arguments against the validity of information's implications) or are not prompted with a mindset intervention as control group (*Mindset Present* and *Mindset Absent* conditions, respectively). I operationalize the mindset intervention by asking auditors to list their thoughts about propositions that they are likely to disagree with, thereby triggering a counterarguing mindset (e.g., Xu and Wyer 2012). In the counterarguing mindset, the propositions were worded in such a way that led auditors to disagree with them (e.g., "Human activity has no major impact on the environment"). Although I expect auditors in both conditions to have similar thoughts about these propositions due to randomization, the thoughts trigger a counterarguing mindset in the *Mindset Present* conditions, because auditors are prompted with these thoughts. As a result, they are likely to induce a counterarguing mindset. In the *Mindset Absent* conditions, auditors are tasked to write their thoughts about arguably neutral things: the pyramids of Egypt, the solar system, and the First World War. Panel B of Table 1 provides an overview of the mindset manipulations.

3.4. Dependent Variables

In the main task (i.e., the audit of the inventory counting procedures), I use three dependent variables to proxy for professional skepticism: *Time Spent*, *Agree with Preparer*, and *Seeded Errors Identified*. In the spillover task (i.e., the audit of a client's step-one analysis of a goodwill impairment test), I use five dependent variables to proxy for professional skepticism: *Reasonableness*, *Surface Issues*, *Deep Issues*, *Total Issues*, and *Contact Directly*. These variables are described below.

I start by outlining the dependent variables for the inventory counting procedures. The first dependent variable used is *Time Spent*; the amount of time spent on reviewing the four inventory counting tasks, measured in seconds. Given that the audit of the inventories consists of recounts, the time spent on a task is arguably a valid proxy for how much effort auditors apply (i.e., more extensive recounting takes more time). If auditors are less skeptical, they are more likely to choose a less effortful strategy and spend less time (Nolder and Kadous 2018). Second, I use *Agree with Preparer*, an indicator variable equal to "1" if an auditor judged the initial count by the workpaper preparer to be correct for a given photograph, and equal to "0" if not. Third, I use *Seeded Errors Identified*, defined as the number of seeded errors identified by auditors in the inventory counting procedures. Given that there are six seeded errors in the fourth (and none in the other photographs), the variable is bounded by zero (lower-bound) and six (upper-bound). The number of seeded errors identified is also a proxy for professional skepticism (Nolder and Kadous 2018).

Next, I outline the dependent variables for the audit of the client's step-one analysis of the goodwill impairment test. I base these dependent variables on Kadous and Zhou (2019). First, *Reasonableness* is defined as auditors' assessment of the overall reasonableness of the fair value, measured on an 11-point Likert scale, ranging from 0 ("not at all likely to be reasonable") to 10 ("extremely likely to be reasonable"). Second, *Surface Issues* are the issues in the goodwill impairment case that require relatively little cognitive processing (two in total). Third, *Deep Issues* are the issues that require relatively a lot of cognitive processing (three in total). Both *Surface Issues* and *Deep Issues* were coded by two independent raters.[23] Fourth, *Total Issues* is the sum of *Surface Issues* and *Deep Issues*. Fifth, *Contact Directly* is an indicator variable that equals "1" when auditors decide to call their manager immediately regarding issues that may indicate the fair value is not reasonable, and equals "0" otherwise (cf. Kadous and Zhou 2019). The overreliance on management's process, failure to gather sufficient evidence, and failure to identify seeded

[23] Two doctoral students coded the number of issues identified by the auditors. Both were blind to experimental conditions. Cohen's Kappa was 0.85 (0.89) for *Surface Issues* (*Deep Issues*), indicating good interrater agreement.

issues are typically seen as resulting from a lack of professional skepticism (PCAOB 2011, Hurtt *et al.* 2013, Griffith *et al.* 2015).

IV. RESULTS

4.1. Manipulation Checks

4.1.1. Source Manipulation

In the post-experimental questionnaire, auditors were asked to evaluate several statements. First, auditors were asked to evaluate who or what conducted the initial count.[24] In the *Human* conditions, auditors were significantly more likely to indicate that the initial count was conducted by a colleague than in the *Automation* conditions ($t = 8.34, p < 0.01$, two-tailed). Similarly, in the *Automation* conditions, auditors were significantly more likely to indicate that the initial count was conducted by an algorithm than in the *Human* conditions ($t = -8.50, p < 0.01$, two-tailed). In addition to directly asking auditors who or what conducted the initial count, auditors were also asked to evaluate the statements about their perceptions during the counting task.[25] Auditors in the *Human* conditions agreed significantly more to the statement about the initial count being conducted by an algorithm than auditors in the *Automation* conditions ($t = 8.51, p < 0.01$, two-tailed). Also, auditors in the *Automation* conditions agreed significantly more to the statement about the initial count being conducted by a person ($t = -9.10, p < 0.01$, two-tailed). Overall, the results of this manipulation check indicate that the source manipulation was successful.

4.1.2. Mindset Manipulation

[24] Specifically, auditors were asked to evaluate the following statement: "[w]ho or what conducted the initial count of the livestock?", where the options were (i) "a colleague", (ii) "the Countthings algorithm", (iii) "both a colleague and the Countthings algorithm", (iv) "neither a colleague nor the Countthings algorithm, and (v) "I don't know."

[25] Specifically, auditors were asked to evaluate the following statements: "While working on the inventory counting task, I thought about the initial count being conducted by a person" and "While working on the inventory counting task, I thought about the initial count being conducted by an algorithm." They evaluated these statements on a seven-point Likert scale ranging from *strongly disagree* to *strongly agree*.

I elicited auditors' attitude toward each of the three propositions used in the counterarguing mindset conditions in the post-experimental questionnaire. These were coded as agreement or disagreement.[26] 88 out of 119 auditors (73.95 percent) disagree with each of the three propositions, potentially generating arguments against the validity of the propositions.[27] Only one participant (0.84 percent) agreed with each of the three propositions. On average, auditors disagreed with the propositions that were prompted in the counterarguing mindset condition ($M = 1.93$ on a seven-point Likert scale ranging from "1 – Strongly Disagree" to "7 – Strongly Agree"). The mean evaluation was significantly lower than the midpoint of the scale ($p < 0.01$, two-tailed, for each proposition). Overall, the results of this manipulation check indicate that auditors indeed tend to disagree with the propositions that were prompted in the counterarguing mindset conditions, allowing them to generate arguments against their validity and thus activating a counterarguing mindset.

4.2. Tests of Hypotheses

4.2.1. Does Automation Usage Reduce Auditors' Professional Skepticism?

The first hypothesis examines whether auditors exhibit less professional skepticism when they rely on work conducted by automated tools and techniques compared to work conducted by another auditor. To test this, I compare the amount of time spent of tasks, the propensity to agree with the workpaper preparer, and the number of seeded errors identified in the *Mindset Absent* conditions, where no counterarguing mindset is prompted (i.e., a simple effect). First, I analyze the amount of time auditors spent on reviewing the tasks. Figure 1 (top figure) shows the amount of time spent

[26] Auditors were asked to evaluate the extent to which they agreed or disagreed with the following statements on a seven-point Likert scale ranging from *strongly disagree* to *strongly agree*: (i) "For a business, it is acceptable to do anything to make a profit", (ii) "Higher education should not be available to all, but only to a small minority of selected students", and (iii) "Human activity has no major impact on the environment."

[27] This number is similar to the counterarguing mindset manipulation in Wyer and Xu (2012), where 76 percent of auditors disagreed with each of three statements.

by auditors on each photograph. Table 2 reports the time spent by condition. Results show that auditors in the *Automation*/*Mindset Absent* condition spent significantly less time reviewing the initial count ($M = 497.46$ seconds) than auditors in the *Human*/*Mindset Absent* condition ($M = 705.72$ seconds), and this difference is statistically significant at the one-percent level ($t = 3.46, p < 0.01$, two-tailed).[28]

Next, I analyze auditors' propensity to agree with the workpaper preparer. That is, after each of the four photos, auditors had to indicate whether they judged the initial count by the *Human*/*Automation* to be correct. Here, I again analyze a simple effect and compare differences in judgments that result from the source manipulation (*Automation* vs. *Human*), while only examining auditors in the conditions where a mindset intervention was absent. Figure 1 (bottom figure) provides graphical evidence that auditors in the *Automation* conditions are more likely to agree with the workpaper preparer, even in cases when the workpaper preparer is wrong. Despite having seen the same photos and the same counts, auditors in the *Automation*/*Mindset Absent* report that significantly more counts are correct than auditors in the *Human*/*Mindset Absent* ($t = -3.26, p < 0.01$, two-tailed).[29] This suggests that auditors' automation usage reduces professional skepticism when a counterarguing mindset intervention is absent. That is, when using automation auditors are significantly more likely to judge it to be correct. This suggests that auditors' automation usage can indeed result in professional skepticism reductions, which is in line with Hypothesis 1.

---

[28] A nonparametric Mann-Whitney test, that does not rely on any distributional assumptions, also shows a significant effect in line with the prediction ($z = 3.29, p < 0.01$, two-tailed). In addition, I find that the results are robust to winsorizing at the 1st and 99th percentile, the 5th and 95th percentile, and at the 10th and 90th percentile.

[29] When a separate *t*-test is conducted for each photo, I find that our inferences remain the same. That is, for photo 1 ($t = -2.56, p = 0.012$, two-tailed), photo 2 ($t = -1.98, p = 0.052$, two-tailed), photo 3 ($t = -1.86, p = 0.068$, two-tailed), and photo 4 ($t = -1.87, p = 0.066$, two-tailed) a higher proportion of auditors in the *Automation/Mindset Absent* condition report that the initial count is correct than in the *Human/Control* condition.

Next, I examine whether auditors failed to find seeded errors. In the experimental case there were four photos where auditors had to review the inventory counting procedures. Whereas the number reported by the initial counter was correct in the first three photos, there were errors seeded in the fourth photo. Specifically, six false positives were seeded into the case. The correct count was 131 and the initial counter reported 137 (see Appendix 1). First, I find that in the *Automation*/*Mindset Absent* condition 12 out of 27 auditors (44.4 percent) incorrectly judge the initial count to be correct, whereas in the *Human*/*Mindset Absent* condition only 7 out of 32 auditors (21.9 percent) incorrectly judge the initial count to be correct. The difference in proportion between conditions is in line with Hypothesis 1 (t = -1.87, $p = 0.066$, two-tailed). If auditors indicated that the number of the initial counter was incorrect, they were asked to provide their own count. Based on their own count, I examine how many seeded errors auditors identify. I find that absent any mindset intervention, auditors in the *Automation* condition identify less seeded errors than those in the *Human* condition (1.59 vs. 2.75 out of 6, $t = 1.91$, $p = 0.061$, two-tailed).

Collectively, these results demonstrate that when auditors rely on work conducted by automated tools and techniques, they will be less professionally skeptical, as indicated by lower time spent (*Time Spent*), a higher likelihood to (incorrectly) judge a task to be correct (*Agree with Preparer*), and a lower number of seeded errors to be identified (*Seeded Errors Identified*). This is consistent with my predictions.

4.2.2. Does a Counterarguing Mindset Help to Alleviate the Professional Skepticism Reductions?

The second hypothesis predicts that the negative effect of automated tools and techniques usage on professional skepticism is weakened when auditors are prompted with a counterarguing mindset. First, I examine how automation usage (*Automation*) and a counterarguing mindset intervention (*Counterarguing*) affect auditors time spent on reviewing the inventory counting tasks (*Time

Spent). Figure 3 graphically demonstrates the observed interaction plot for auditors' time spent reviewing the inventory counting procedures. A visual inspection reveals that the negative effect of automation is visible in the *Mindset Absent* conditions, but not in the *Mindset Present* conditions. In fact, the effect in the *Mindset Present* conditions is slightly positive. I also provide formal tests of significance. Panel C of Table 2 reports an ANOVA model. I find the difference in slopes is significant at the one percent level ($p < 0.01$, two-tailed). Together with the visual fit, this indicates that the slope of the effect of *Automation* on *Time Spent* is significantly less negative in the *Mindset Present* conditions than in the *Mindset Absent* conditions, implying that a counterarguing mindset alleviates the negative effect of automation on professional skepticism. The two main effects of *Automation* and *Counterarguing* are insignificant ($p > 0.10$, two-tailed).

Second, I examine the effects of *Automation* and *Counterarguing* on the number tasks that auditors judge to be correctly prepared by the workpaper preparer (*Agree with Preparer*). Figure 4 graphically depicts the observed interaction plot for *Agree with Preparer*. The visual fit shows that auditors are more likely to agree with the workpaper preparer when the workpaper is prepared by automation than by a human colleague, suggesting that they are less skeptical. Also, the increase from *Human* to *Automation* is greater in the *Mindset Absent* conditions than in the *Mindset Present* conditions. This suggests, at least visually, that a counterarguing mindset alleviates the effects of *Automation* on professional skepticism. I also conduct formal tests of significance in Panel C of Table 3. The interaction effect is marginally significant based on a one-tailed test in line with my prediction (($p = 0.07$, one-tailed; $p = 0.14$, two-tailed). Next to that, I find a main effect for *Automation* on *Agree with Preparer* ($p < 0.01$, two-tailed), while finding no significant main effect for *Counterarguing* ($p > 0.10$, two-tailed).

Third, I also test the effects of *Automation* and *Counterarguing* on the number of seeded errors identified (*Seeded Errors Identified*). Figure 5 shows the observed interaction plot for *Seeded Errors Identified*. The interaction plot shows a disordinal interaction, where the slope of the effect of *Automation* on *Seeded Errors Identified* is positive (negative) when a counterarguing mindset is present (absent). Table 4 reports formal tests of significance. I find that the interaction effect is statistically significant at the five percent level ($p = 0.05$, two-tailed), which is in line with earlier findings and Hypothesis 2. None of the main effects is significant ($p > 0.10$, two-tailed).

Collectively, the results suggest that the negative effect of automated tools and techniques usage on professional skepticism is weakened when auditors are prompted with a counterarguing mindset. This is in line with my prediction and indicates that a counterarguing mindset could be a helpful tool for audit firms to use when auditors work with automation. However, the results also indicate that audit firms need to be careful. That is, an unexpected finding is that *Counterarguing* may also reduce *Time Spent* when the workpaper is prepared by a human colleague.[30] One potential explanation for this unexpected finding may be that auditors that are prompted with a counterarguing mindset may perceive other humans to have a counterarguing mindset as well and are therefore less skeptical in reviewing the work of other humans.[31]

4.2.3. Do Professional Skepticism Reductions Caused by Automation Spill Over to Subsequent Audit Tasks?

The third hypothesis predicts that auditors exhibit less professional skepticism in a subsequent task when they relied on automated tools and techniques in a previous task. I test this using the five dependent variables elicited from the goodwill impairment case that followed the inventory

---

30 Tests of simple effects of *Counterarguing* on the three dependent variables in the *Human* conditions show a significant effect when *Time Spent* is the dependent variable ($p < 0.01$, two-tailed), while showing an insignificant effect when *Agree with Preparer* and *Seeded Errors Identified* are the dependent variable ($p > 0.10$, two-tailed).

31 This is consistent with social projection theory and this social (Robbins and Krueger 2005).

counting task: *Reasonableness*, *Surface Issues*, *Deep Issues*, *Total Issues*, and *Contact Directly* (see Appendix B for variable definitions). Panel A of Table 5 reports the descriptive statistics of these variables. The mean of the variables is similar to mean of those variables in Kadous and Zhou (2019). I start by investigating whether *Automation* affects the dependent variables in the *Mindset Absent* conditions. For all five dependent variables, I do not find a significant effect ($p >$ 0.10, two-tailed). In Panel B of Table 5, I use a negative binomial regression model to test Hypothesis 3, given the nature of the dependent variables (cf. Kadous and Zhou 2019). In the spillover case, I do not find evidence for statistically significant main effects, nor for a significant interaction effect. The only exception is the positive coefficient of a main effect *Counterarguing* on *Reasonableness* ($p = 0.02$, two-tailed). Hence, overall I do not find evidence that professional skepticism reductions caused by automation usage in a first task spill over to a second task.

4.3. Supplemental Analyses

In this section, I perform supplemental analyses to provide further process evidence about the role of attention and the role of effortful analysis in explaining the findings. The literature on automation bias identifies reductions in attention and effortful processing following automation usage as drivers of the automation bias.

4.3.1. Process Evidence: Attention During Inventory Counting Procedures

The cognitive processes underlying the automation bias involve reductions in vigilance and attention spent to the task. I investigate whether auditors' attention during the review of the inventory counting procedures differs between conditions. During the inventory counting procedures, auditors had to review four photos. Whereas three of the four photos included cattle, the other photo they had to count included sheep. To test the attention spent by auditors, I asked auditors in the post-experimental questionnaire to recall what other animals than cattle were shown in the photos.[32] 93 out of 119 auditors were able to correctly recall that the other animals in the inventory counting procedures were sheep. First, not surprisingly, I find that auditors that were able to correctly recall the animals spent significantly more time on the inventory counting task than auditors that did not recall the animals ($t = -2.43$, $p = 0.02$, two-tailed). Second, I find a marginally significant interaction effect of *Automation* and *Counterarguing* on a dummy variable that captures whether auditors were able to correctly recall the animals ($F = 3.54$, $p = 0.06$, two-tailed).

4.3.2. Process Evidence: Path Analysis

Next, I conduct path analyses to test whether *Time Spent* mediates the relationship between *Automation* and *Seeded Errors Identified*. That is, I test whether a reduction in effortful processing causes the negative effect of automation on auditors' propensity to identify seeded errors.[33] Thereby, I further examine to what extent the negative effects of automation on professional skepticism are driven by less effortful processing and attention spent to the evidence provided by the automation (cf. Parasuraman and Manzey 2010). Figure 6 displays the mediation models for both the *Mindset Absent* (top figure) and *Mindset Present* (bottom figure) conditions. In the *Mindset Absent* conditions, I find that *Automation* significantly reduces *Time Spent* ($\beta = -208.26$, $z = -3.48$, $p < 0.01$, two-tailed) and *Time Spent* is significantly positively related to *Seeded Errors Identified* ($\beta = 0.002$, $z = 1.99$, $p = 0.047$, two-tailed). Although the total effect (i.e., the c-path) of *Automation* on *Seeded Errors Identified* is marginally significant ($\beta = -1.16$, $z = -1.95$, $p = 0.051$, two-tailed), the direct effect (i.e., the c'-path) in the mediated

[32] Specifically, the item included the question: "*In the inventory counts, there were four aerial views of animals. Three of them contained cattle (i.e., cows and bulls). Which animals did the other one contain?*"

[33] Nolder and Kadous (2018, p. 7) identify both the time spent on task and the seeded errors identified as measures of a skeptical mindset. Yet, arguably the degree to what effortful analysis is conducted (i.e., time spent) can affect the number of seeded errors identified.

model is insignificant ($p > 0.10$, two-tailed). The indirect effect is insignificant at conventional two-tailed significant levels, but marginally significant at one-tailed significance levels ($p = 0.12$, two-tailed; $p = 0.06$, one-tailed).

In the *Mindset Present* conditions, I do not find evidence for a significant total effect of *Automation* on *Seeded Errors Identified* ($p > 0.10$, two-tailed). Nor do I find a significant effect of *Automation* on *Time Spent* ($p > 0.10$, two-tailed). The only relationship in the mediation model that shows marginal significance is the relationship between *Time Spent* and *Seeded Errors Identified* (i.e., the b-path, $\beta = 0.002$, $z = 1.66$, $p = 0.097$, two-tailed). This supplementary analysis shows that when a counterarguing mindset intervention is absent, auditors engage in less effortful processing and this leads them to identify less seeded errors. When auditors are prompted with a counterarguing mindset, the automation does not cause them to engage in less effortful processing, and therefore there are no adverse effects of automation on the number of seeded errors identified.

V. CONCLUSION

In this study, I investigate the effect of auditors' automation usage on their professional skepticism. I find that absent a counterarguing mindset intervention, auditors are less skeptical when they rely on work conducted by the audit firm's automated tools and techniques than when relying on the same work by an audit team member. This indicates potential drawbacks of using automated tools and techniques. To alleviate these drawbacks, I employ a counterarguing mindset intervention that successfully alleviates the negative effects of automation usage on professional skepticism. Finally, I investigate whether reductions in professional skepticism that are caused automation usage also spill over to subsequent unrelated tasks. I do not find any evidence indicating a spillover effect.

These findings are relevant for audit practice and theory. That is, for the potential advantages of automation to materialize, it is important that auditors' reliance on automated tools and techniques is based on thorough analysis of auditors' cognitive and motivational decision-making processes. Audit firms could, for instance, employ counterarguing mindset interventions to mitigate the negative effects of automated tools and techniques on professional skepticism. Also, my findings could help regulators and standard setters, such as the IAASB and PCAOB, to better make decisions in inherently difficult trade-offs regarding the use of automation in an audit. Next to that, this study contributes to the literature on professional skepticism in auditing (e.g., Nolder and Kadous 2018). With the emerging of automated tools and techniques such as data analytics, artificial intelligence, and robotic process automation, auditors increasingly have to exhibit professional skepticism to information prepared by those objects (Olsen and Gold 2018). Yet, to the best of my knowledge, no study has hitherto investigated how professional skepticism may be different towards automation.

My study is also subject to limitations. In the auditing setting, there are numerous possible automated tools and techniques, audit team members, audit tasks, and auditors. In my study, partially due to the nature of experiments, I was constrained in testing various alternatives. Therefore, readers need to be cautious in generalizing findings to other tasks. Future research can further explore different variations, and potentially explore boundary conditions. Overall, the relationship between automated tools and techniques and professional skepticism appears to be a fruitful area for future research.

REFERENCES

ABDOLMOHAMMADI, M. J. (1999). A comprehensive taxonomy of audit task structure, professional rank and decision aids for behavioral research. *Behavioral Research in Accounting*, *11*, 51-92.

AMERICAN INSTITUTE OF CERTIFIED PUBLIC ACCOUNTANTS (AICPA). (1997) *Consideration of fraud in a financial statement audit*. New York, NY: AICPA.

AUSTIN, A. A., CARPENTER, T. D., CHRIST, M. H., & NIELSON, C. S. (2021). The Data Analytics Journey: Interactions Among Auditors, Managers, Regulation, and Technology. *Contemporary Accounting Research*, *38*(3), 1888-1924.

BAUER, T. D. (2015). The Effects of Client Identity Strength and Professional Identity Salience on Auditor Judgments. *The Accounting Review*, *90*(1), 95-114.

BAUMEISTER, R. (1998). The Self. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4th ed., Vol. 4, pp. 680–740). New York: McGraw-Hill.

BAUMEISTER, R., BRATSLAVSKY, E., MURAVEN, M., & TICE, D. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265.

BLOOMBERG TAX. (2020). *Big Four Invest Billions in Tech, Reshaping Their Identities.* Available at https://news.bloombergtax.com/financial-accounting/big-four-invest-billions-in-tech-reshaping-their-identities.

CAO, T., DUH, R. R., TAN, H. T., & XU, T. (2022). Enhancing Auditors' Reliance on Data Analytics under Inspection Risk Using Fixed and Growth Mindsets. *The Accounting Review.*

CASSIDY, M. (2014). *Centaur Chess Shows Power of Teaming Human and Machine*. Huffington Post. Available at: https://www.huffpost.com/entry/centaur-chess-shows-power_b_63836

CASTELO, N., BOS, M. W., & LEHMANN, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, *56*(5), 809-825.

CHRIST, M. H., EMETT, S. A., SUMMERS, S. L., & WOOD, D. A. (2021). Prepare for takeoff: Improving asset measurement and audit quality with drone-enabled inventory audit procedures. *Review of Accounting Studies*, 1-21.

COMMERFORD, B. P., DENNIS, S. A., JOE, J. R., & ULLA, J. W. (2022). Man versus machine: Complex estimates and auditor reliance on artificial intelligence. *Journal of Accounting Research*, *60*(1), 171-201.

COSTELLO, A. M., DOWN, A. K., & MEHTA, M. N. (2020). Machine+ man: A field experiment on the role of discretion in augmenting AI-based lending models. *Journal of Accounting and Economics*, *70*(2-3), 101360.

DIERYNCK, B., & PETERS, C. P. H. (2021). Auditor Task Prioritization: The Effects of Time Pressure and Psychological Ownership. *Available at SSRN 3450363*.

DIETVORST, B. J., SIMMONS, J. P., & MASSEY, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114-126.

DIETVORST, B. J., SIMMONS, J. P., & MASSEY, C. (2018). Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science*, *64*(3), 1155-1170.

DIJKSTRA, J. J. (1999). User agreement with incorrect expert system advice. *Behaviour & Information Technology*, *18*(6), 399-411.

DIJKSTRA, J. J., LIEBRAND, W. B., & TIMMINGA, E. (1998). Persuasiveness of expert systems. *Behaviour & Information Technology*, *17*(3), 155-163.

DZINDOLET, M. T., PIERCE, L. G., BECK, H. P., & DAWE, L. A. (2002). The perceived utility of human and automated aids in a visual detection task. *Human Factors*, *44*(1), 79-94.

DZINDOLET, M. T., PETERSON, S. A., POMRANKY, R. A., PIERCE, L. G., & BECK, H. P. (2003). The Role of Trust in Automation Reliance. *International Journal of Human-Computer Studies*, *58*(6), 697-718.

ERNST & YOUNG (EY). (2018a). *EY to spend US$1 billion as part of its innovation drive.* (August 13). Available at: https://www.ey.com/en_nl/news/2018/08/ey-to-spend-us-1-billion-as-part-of-its-innovation-drive.

ERNST & YOUNG (EY). (2018b). *How Artificial Intelligence Will Transform the Audit.* (July 20). Available at: https://www.ey.com/en_us/assurance/how-artificial-intelligence-will-transform-the-audit.

ERNST & YOUNG (EY). (2022). *EY announces US$1b investment in a next generation technology platform to facilitate trust, transparency and transformation through assurance services.* (June 16). Available at: https://www.ey.com/en_gl/news/2022/06/ey-announces-us-1b-investment-in-a-next-generation-technology-platform-to-facilitate-trust-transparency-and-transformation-through-assurance-services

GOLLWITZER, P. (1990). Action Phases and Mind-Sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition* (Vol. 2, pp. 53–92). New York: Guilford Press.

GRIFFITH, E. E., HAMMERSLEY, J. S., KADOUS, K., & YOUNG, D. (2015). Auditor Mindsets and Audits of Complex Estimates. *Journal of Accounting Research*, *53*(1), 49-77.

GRIFFITH, E. E., KADOUS, K., & YOUNG, D. (2016). How insights from the "new" JDM research can improve auditor judgment: Fundamental research questions and methodological advice. *Auditing: A Journal of Practice & Theory*, *35*(2), 1-22.

HAMILTON, R., VOHS, K. D., SELLIER, A. L., & MEYVIS, T. (2011). Being of two minds: Switching mindsets exhausts self-regulatory resources. *Organizational Behavior and Human Decision Processes*, *115*(1), 13-24.

HARRIS, S. B. (2017). Technology and the Audit of Today and Tomorrow. Speech delivered at PCAOB/AAA Annual Meeting: Washington DC, April 20, 2017.

HOBFOLL, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513-524.

HOBFOLL, S. E. (2001). The influence of culture, community, and the nested self in the stress process: Advancing Conservation of Resources theory. *Applied Psychology*, *50*(3), 337-421.

HURLEY, P. J. (2019). Ego depletion and auditors' JDM quality. *Accounting, Organizations and Society*, *77*, 101048.

HURTT, R. K. (2010). Development of a scale to measure professional skepticism. *Auditing: A Journal of Practice & Theory*, *29*(1), 149-171.

HURTT, R. K., BROWN-LIBURD, H., EARLEY, C. E., & KRISHNAMOORTHY, G. (2013). Research on Auditor Professional Skepticism: Literature Synthesis and Opportunities for Future Research. *Auditing: A Journal of Practice & Theory*, *32*(sp1), 45-97.

INTERNATIONAL AUDITING AND ASSURANCE STANDARDS BOARD (IAASB). (2012). Overall objectives of the independent auditor and the conduct of an audit in accordance with international standards on auditing. ISA 200. New York, NY: IFAC.

INTERNATIONAL AUDITING AND ASSURANCE STANDARDS BOARD (IAASB). (2017). *Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics.* New York, NY: IFAC.

INTERNATIONAL AUDITING AND ASSURANCE STANDARDS BOARD (IAASB). (2021a). FAQ: Addressing the Risk of Overreliance on Technology: Use of ATT and Use of Information Produced by Entity's Systems. New York, NY: IFAC.

INTERNATIONAL AUDITING AND ASSURANCE STANDARDS BOARD (IAASB). (2021b). 2020 Handbook of international quality control, auditing, review, other assurance, and related service pronouncements. New York, NY: IFAC.

International Forum of Independent Audit Regulators (IFIAR). 2018. *Survey of Inspection Findings 2017*.

KADOUS, K., & ZHOU, Y. (2019). How Does Intrinsic Motivation Improve Auditor Judgment in Complex Audit Tasks?. *Contemporary Accounting Research*, *36*(1), 108-131.

KPMG, LLP. (2016). Harnessing the Power of Cognitive Technology to Transform the Audit. Wilmington, DE: KPMG

KPMG, LLP. (2019). KPMG Expects to Invest US $5 Billion on Digital Leadership in Professional Services. Available at: https://home.kpmg/xx/en/home/media/press-releases/2019/12/kpmg-expects-to-invest-5-billion-on-digital-leadership-in-professional-services.html.

LANGER, E. J. (1989). *Mindfulness.* Reading, MA: Addison-Wesley.

LOGG, J. M., MINSON, J. A., & MOORE, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90-103.

MADHAVAN, P., & WIEGMANN, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277-301.

MERRITT, S. M., UNNERSTALL, J. L., LEE, D., & HUBER, K. (2015). Measuring individual differences in the perfect automation schema. *Human Factors*, *57*(5), 740-753.

MOSIER K. L., SKITKA, L. J. (1996) Human decision makers and automated decision aids: made for each other. In: R Parasuraman, M Mouloua, eds. *Automation and Human Performance:*

Theory and Applications. Hillsdale, NJ, England: Lawrence Erlbaum Associates; 1996: 201–220.

MULLIS, C. E., & HATFIELD, R. C. (2018). The Effects of Multitasking on Auditors' Judgment Quality. *Contemporary Accounting Research*, *35*(1), 314-333.

MUNOKO, I., BROWN-LIBURD, H. L., & VASARHELYI, M. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics*, *167*(2), 209-234.

NELSON, M. W. (2009). A Model and Literature Review of Professional Skepticism in Auditing. *Auditing: A Journal of Practice & Theory*, *28*(2), 1-34.

NO, W. G., LEE, K., HUANG, F., & LI, Q. (2019). Multidimensional audit data selection (MADS): A framework for using data analytics in the audit data selection process. *Accounting Horizons*, *33*(3), 127-140.

NOLDER, C. J., & KADOUS, K. (2018). Grounding the professional skepticism construct in mindset and attitude theory: A way forward. *Accounting, Organizations and Society*, *67*, 1-14.

OLSEN, C., & GOLD, A. (2018). Future research directions at the intersection between cognitive neuroscience research and auditors' professional skepticism. *Journal of Accounting Literature*, *41*, 127-141.

O'NEIL, C. (2016). Weapons of Math Destruction: *How Big Data Increases Inequality and Threatens Democracy*. Penguin Books.

PARASURAMAN, R., & MANZEY, D. H. (2010). Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors*, *52*(3), 381-410.

PARASURAMAN, R., & RILEY, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, *39*(2), 230-253.

PETTY, R. E., & CACIOPPO, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In *Communication and persuasion* (pp. 1-24). Springer, New York, NY.

PHILLIPS, F. (1999). Auditor Attention to and Judgments of Aggressive Financial Reporting. *Journal of Accounting Research*, *37*(1), 167-189.

PIERCEY, M. D. (2011). Documentation requirements and quantified versus qualitative audit risk assessments. *Auditing: A Journal of Practice & Theory*, *30*(4), 223-248.

PRICEWATERHOUSECOOPERS (PWC). (2016). Clarity from Above: PwC Global Report on the Commercial Application of Drone Technology. *Available at: https://pwc.blogs.com/files/clarity-from-above-pwc.pdf*.

PRICEWATERHOUSECOOPERS (PWC). (2017). Confidence in the future. Human and machine collaboration in the audit. Available at: https://www.pwc.co.uk/services/audit-assurance/statutory-audit/insights/confidence-in-the-future-audit-innovation/human-and-machine-collaboration-in-the-audit.html

PRICEWATERHOUSECOOPERS (PWC). 2019. Investing in Technology and People: PwC's Global Annual Review 2019. *Available at: https://www.pwc.com/gx/en/about-pwc/global-annual-review-2019/downloads/ambition.pdf*.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). (2003). *Consideration of Fraud in a Financial Statement Audit*. Auditing Standard 2401. Washington, D.C., PCAOB.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). (2010a). *Due Professional Care in the Performance of Work*. Auditing Standard 1015. Washington, D.C.: PCAOB.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). (2010b). *The Auditor's Responses to the Risks of Material Misstatement*. Auditing Standard 2301. Washington, D.C.: PCAOB.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). (2011). *Concept Release on Auditor Independence and Audit Firm Rotation*. PCAOB Release No. 2011-006. Washington, D.C.: PCAOB.

PUBLIC COMPANY ACCOUNTING OVERSIGHT BOARD (PCAOB). (2022). *Data and Technology*. Available at: https://pcaobus.org/oversight/standards/standard-setting-research-projects/data-technology.

QUADACKERS, L., GROOT, T., & WRIGHT, A. (2014). Auditors' Professional Skepticism: Neutrality versus Presumptive Doubt. *Contemporary Accounting Research*, *31*(3), 639-657.

RAJ, M., & SEAMANS, R. (2019). Primer on artificial intelligence and robotics. *Journal of Organization Design*, *8*(1), 1-14.

ROBBINS, J. M., & KRUEGER, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review*, *9*(1), 32-47.

ROBINSON, S. N., CURTIS, M. B., & ROBERTSON, J. C. (2018). Disentangling the trait and state components of professional skepticism: Specifying a process for state scale development. *Auditing: A Journal of Practice & Theory*, *37*(1), 215-235.

SIMON, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence*, *4*(3-4), 181-201.

VAN RINSUM, M., MAAS, V. S., & STOLKER, D. (2018). Disclosure Checklists and Auditors' Judgments of Aggressive Accounting. *European Accounting Review*, *27*(2), 383-399.

WYER JR, R. S., & XU, A. J. (2010). The role of behavioral mind-sets in goal-directed activity: Conceptual underpinnings and empirical evidence. *Journal of Consumer Psychology*, *20*(2), 107-125.

XU, A. J., & WYER JR, R. S. (2012). The role of bolstering and counterarguing mind-sets in persuasion. *Journal of Consumer Research*, *38*(5), 920-932.

YANIV, I., & KLEINBERGER, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes*, *83*(2), 260-281.

YOON, K., HOOGDUIN, L., & ZHANG, L. (2015). Big Data as Complementary Audit Evidence. *Accounting Horizons*, *29*(2), 431-438.

ZHANG, C., THOMAS, C., & VASARHELYI, M. A. (2022). Attended process automation in audit: a framework and a demonstration. *Journal of Information Systems*, *36*(2), 101-124.

APPENDIX A
Examples of Photographs Used in the Inventory Counting Procedures

FIGURE A.1.

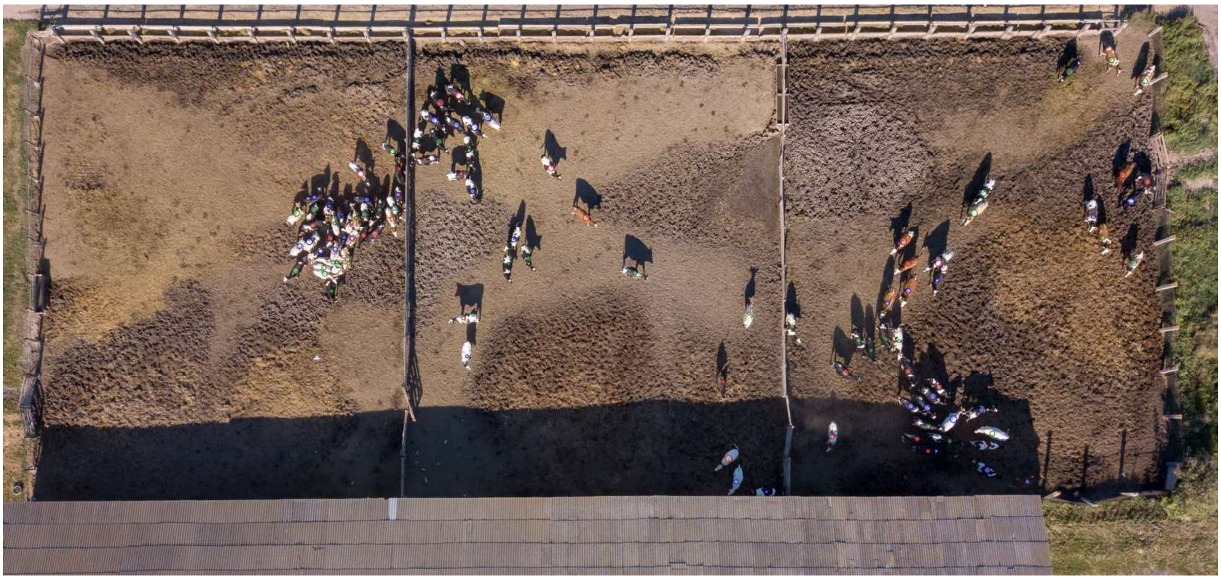


Figure A.1. presents a screenshot of the first aerial photograph of a pen that auditors needed to review. The count by the preparer is 98 and the actual number of cattle is also 98. Hence, there are no seeded errors in this subtask. The image in the experimental case was large, and an additional magnifier was provided, such that auditors were able to manually check whether the initial count was correct.

FIGURE A.2.

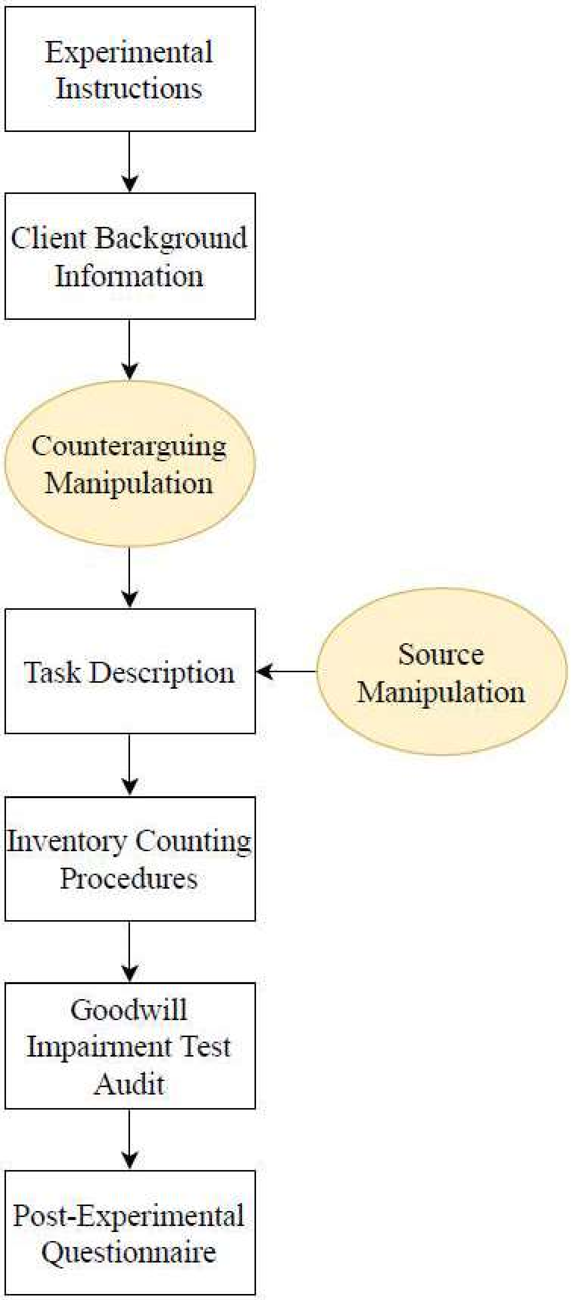
Figure A.2. presents a screenshot of the fourth and last aerial photograph of a pen that auditors needed to review. The count by the preparer is 137, while the actual number of cattle 131. There are 6 seeded errors in this subtask. The image in the experimental case was large, and an additional magnifier was provided, such that auditors were able to manually check whether the initial count was correct.

APPENDIX B
Variable Definitions

Variable	Definition
Dependent Variables - Main Task	
Time Spent	The amount of time spent on reviewing the four inventory counting tasks, measured in seconds.
Agree with Preparer	An indicator variable equal to "1" if auditors judged the initial count by the workpaper preparer to be correct, and equal to "0" if not.
Seeded Errors Identified	The number of seeded errors identified by auditors in the inventory counting procedures. Bounded by zero (lower-bound) and six (upper-bound).
Dependent variables - Spillover Task	
Reasonableness	Auditors' assessment of the overall reasonableness of the fair value, measured on an 11-point Likert scale, ranging from 0 ("not at all likely to be reasonable") to 10 ("extremely likely to be reasonable").
Surface Issues	The number of surface issues identified in the goodwill impairment case. (0 - 2)
Deep Issues	The number of deep issues identified in the goodwill impairment case. (0 - 3)
Total Issues	The total number of issues (surface + deep) identified in the goodwill impairment case. (0 - 5)
Contact Directly	An indicator variable that equals "1" when auditors decide to call their manager immediately regarding issues that may indicate the fair value is not reasonable, and equals "0" otherwise (cf. Kadous and Zhou 2019).
Independent Variables	
Automation	An indicator variable that equals "1" if auditors are in the *Automation* conditions and equals "0" if auditors are in the *Human* conditions.
Counterarguing	An indicator variable that equals "1" if auditors are in the *Mindset Present* conditions and equals "0" if auditors are in the *Mindset Absent* conditions.

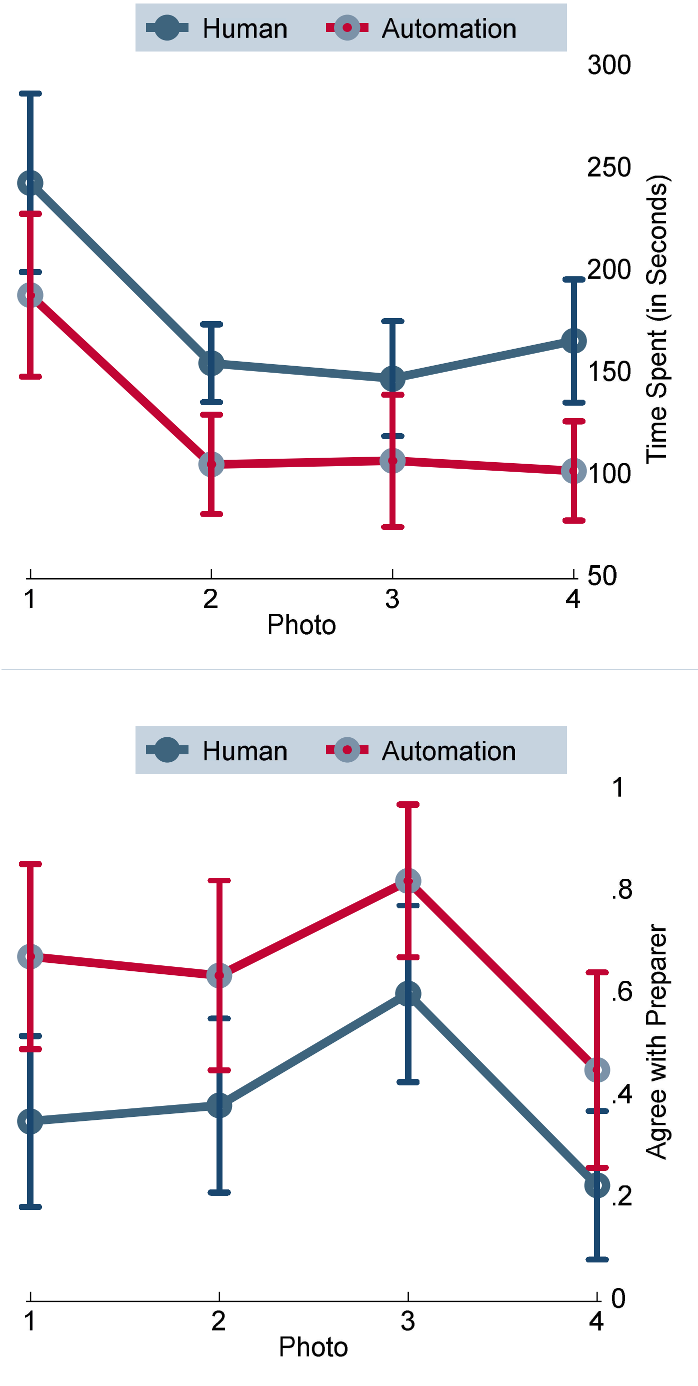
FIGURES

FIGURE 1
Instrument Flow



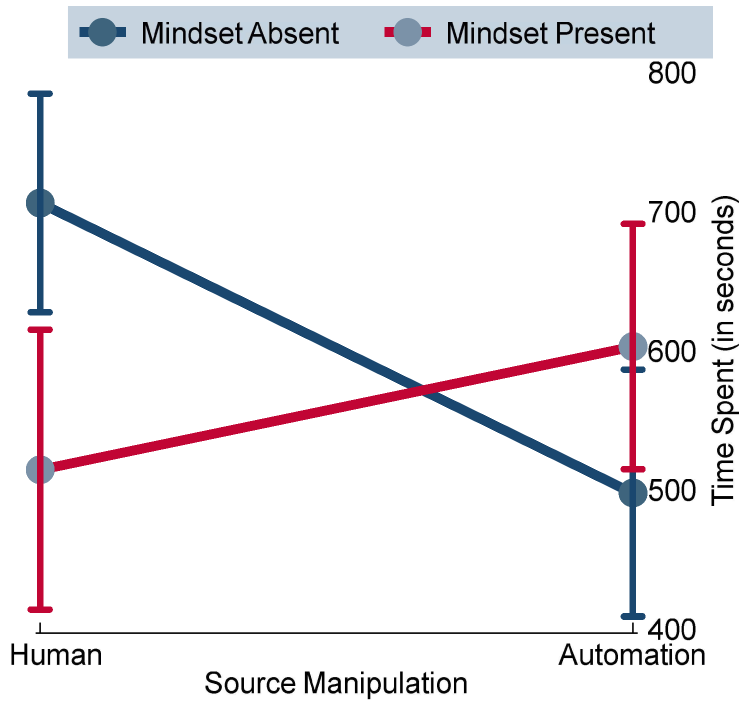
Notes: Figure 1 presents the instrument flow. The experimental case and procedures are described in Section 3.2. The *Counterarguing Manipulation* and *Source Manipulation* are described in Section 3.3. The manipulations are shown in Table 1.

FIGURE 2
Simple Effect of Automation on Professional Skepticism in *Mindset Absent* Conditions



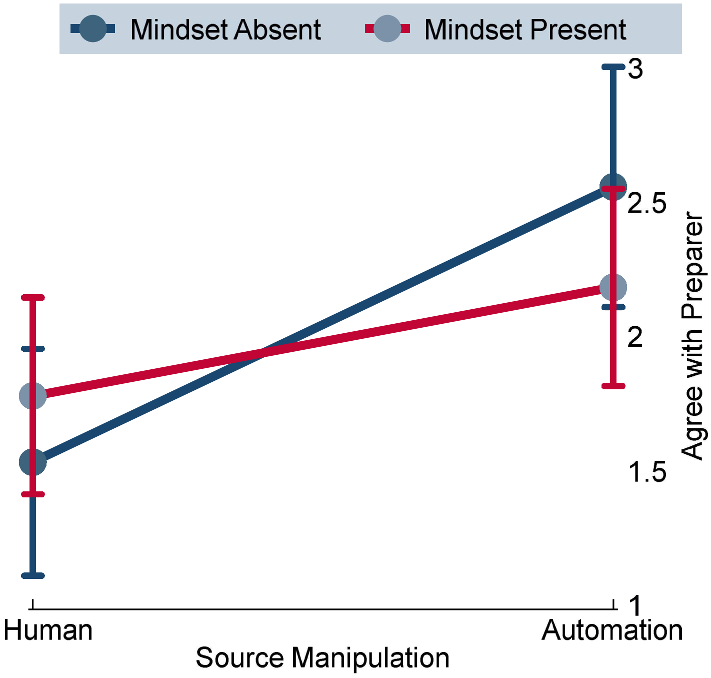
Notes: Figure 2 displays *Time Spent* (top figure) and *Agree with Preparer* (bottom figure) for both the *Human* and *Automation* conditions nested within the *Mindset Absent* conditions (i.e., the simple effect of *Automation* in the *Mindset Absent* conditions).

FIGURE 3
Observed Interaction Plot for Auditors' Time Spent Reviewing Counting Procedures



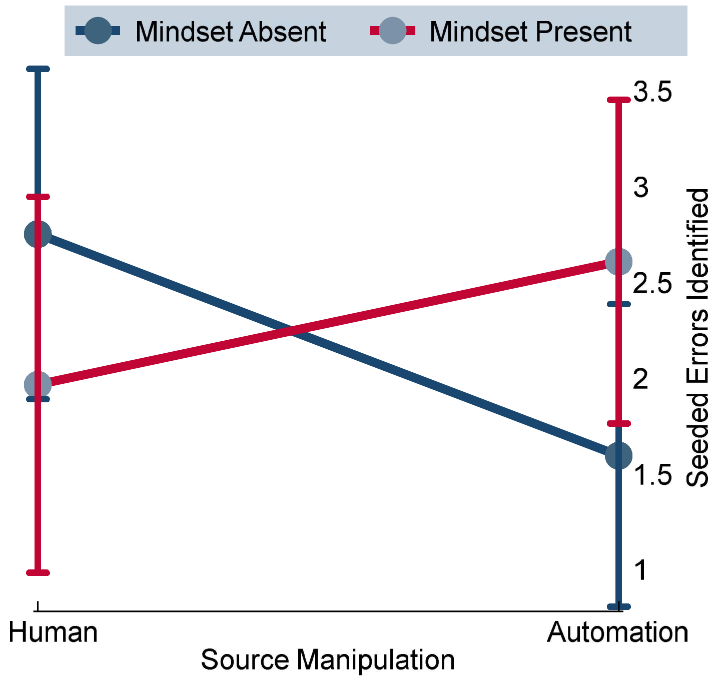
Notes: Figure 3 shows the interaction plot of the *Source* and *Mindset* manipulations. The dependent variable is *Time Spent*, the time that auditors spent on reviewing the inventory counts, measured in seconds. The blue (darker dots) line indicates the *Mindset Absent* conditions (i.e., those that were not prompted with a counterarguing mindset). The red (lighter dots) line indicates *Mindset Present* conditions (i.e., those that were prompted with a counterarguing mindset). 95 percent confidence intervals are provided. Robust standard errors are used. See Appendix B for variable definitions.

FIGURE 4
Observed Interaction Plot for Agree with Preparer



Notes: Figure 4 shows the interaction plot of the *Source* and *Mindset* manipulations. The dependent variable is *Agree with Preparer*, An indicator variable equal to "1" if auditors judged the initial count by the workpaper preparer to be correct, and equal to "0" if not. The blue (darker dots) line indicates the *Mindset Absent* conditions (i.e., those that were not prompted with a counterarguing mindset). The red (lighter dots) line indicates *Mindset Present* conditions (i.e., those that were prompted with a counterarguing mindset). 95 percent confidence intervals are provided. Robust standard errors are used. See Appendix B for variable definitions.

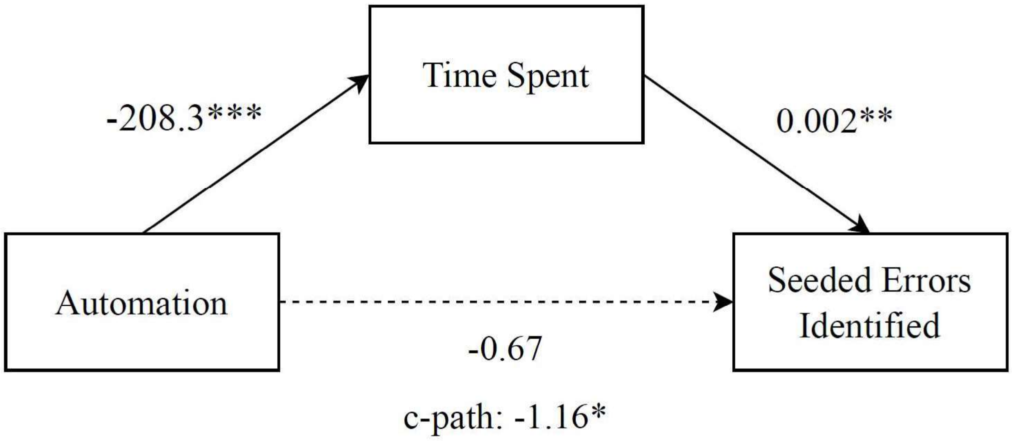
FIGURE 5
Observed Interaction Plot for Seeded Errors Identified



Notes: Figure 5 shows the interaction plot of the *Source* and *Mindset* manipulations. The dependent variable is *Seeded Errors Identified*, the number of seeded errors identified by auditors in the inventory counting procedures. Bounded by zero (lower-bound) and six (upper-bound). The blue (darker dots) line indicates the *Mindset Absent* conditions (i.e., those that were not prompted with a counterarguing mindset). The red (lighter dots) line indicates *Mindset Present* conditions (i.e., those that were prompted with a counterarguing mindset). 95 percent confidence intervals are provided. Robust standard errors are used. See Appendix B for variable definitions.

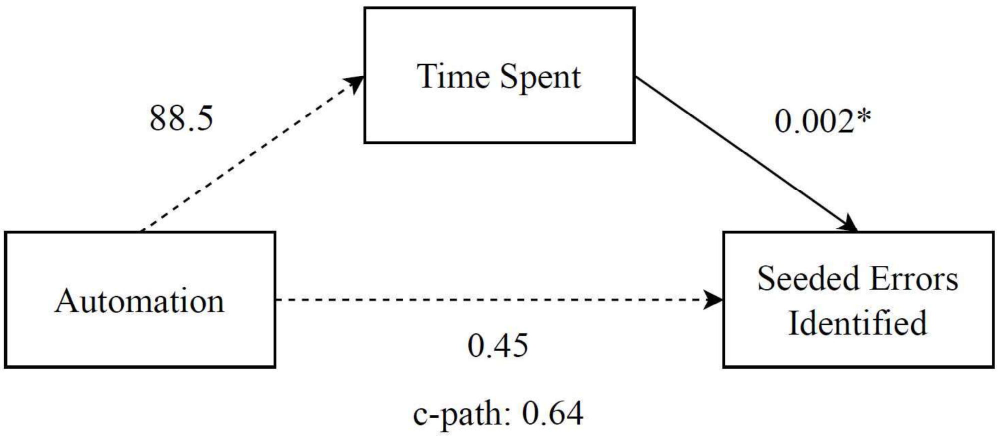
FIGURE 6
Path Analyses

Panel A: Mediation Analysis in the *Mindset Absent* Conditions



Indirect effect: β = -0.48, z = -1.56, p = 0.12, two-tailed, n = 59.

Panel B: Mediation Analysis in the *Mindset Present* Conditions



Indirect effect: β = 0.19, z = 0.91, p = 0.36, two-tailed, n = 60.

Notes: Figure 6 shows the path models that demonstrate how *Time Spent* mediates between *ATT* and *Seeded Errors Identified*. Two models are reported: Panel A (B) reports the mediation analysis in the *Mindset Absent* (*Mindset Present*) conditions. The c-path represents a univariate regression of *Seeded Errors Identified* on *Automation*. Robust standard errors are used. See Appendix B for variable definitions. Standardized path coefficients provided. Nonsignificant results are denoted by a dashed line. $^{*}\ p \leq 0.10$, $^{**}\ p \leq 0.05$, and $^{***}\ p \leq 0.01$, all p-values are two-tailed.

TABLES

TABLE 1
Manipulations
(differences between treatments in italics)

Panel A: Source Manipulation

Human	Automation
The inventory on the photos was counted by *your firm's proprietary counting software*.	The inventory on the photos was counted by *an audit team member*.
The counting algorithm is based on the Countthings (machine-learning based) algorithm, and applies firm-approved methodologies to conduct the inventory count.	*The audit team member* applies firm-approved methodologies to conduct the inventory count.
the counting algorithm [used throughout]	*the audit team member* [used throughout]

Panel B: Mindset Manipulation

Mindset Present	Mindset Absent
Before you continue the audit engagement, you are asked to write some *arguments about one of the propositions, testing your ability to articulate arguments.*	Before you continue the audit engagement, you are asked to write some *facts about one of the topics, testing your general knowledge*.
Topic 1: *For a business, it is acceptable to do anything to make a profit*.	Topic 1: *The pyramids of Egypt.*
Topic 2: *Higher education should not be available to all, but only to a small minority of selected students.*	Topic 2: *The solar system.*
Topic 3: *Human activity has no major impact on the environment.*	Topic 3: *The First World War*.
Think about one of the above *propositions that you have the strongest feeling about.* Write a short essay *indicating why you agree or disagree with it.* You have three to four minutes to mention *a couple of arguments*.	Think about one of the above *topics that you know most about*. Write a short essay *about this topic*. You have three to four minutes to mention *a couple of facts*.

TABLE 2
Time Spent by Condition

Panel A: Descriptive Statistics: Mean (Standard Deviation) [N]

	Mindset Intervention		Collapsed across Mindset
	Absent	Present	
Human	705.72	514.21	618.08
	(226.16)	(266.82)	(261.75)
	[32]	[27]	[59]
Automation	497.46	602.73	555.36
	(235.63)	(257.95)	(251.68)
	[27]	[33]	[60]
Collapsed across Source	610.42	562.89	586.46
	(251.35)	(263.50)	(257.56)
	[59]	[60]	[119]

Panel B: Planned Simple Effects

Simple Effect	t	Two-Sided p-value
Effect of Automation on Time Spent in the Mindset Absent Conditions	3.46	**< 0.01**
Effect of Automation on Time Spent in the Mindset Present Conditions	-1.30	0.20

Panel C: Analysis of Variance (ANOVA)

Source	MS	F	Two-Sided p-value
Automation	105724.49	1.73	0.19
Counterarguing	54851.10	0.90	0.35
Automation × Counterarguing	649391.71	10.65	**< 0.01**
Residual	60951.18		

Notes: Table 2 reports descriptive statistics and hypotheses tests for *Time Spent*. Panel A provides the descriptive statistics by condition. Panel B reports the planned simple effects of automation depending on the *Counterarguing* conditions. Panel C reports an analysis of variance (ANOVA). See Appendix B for variable definitions.

TABLE 3
Agree with Preparer by Condition

Panel A: Descriptive Statistics: Mean (Standard Deviation) [N]

	Mindset Intervention		**Collapsed across Mindset**
	Absent	**Present**	
Human	1.53	1.78	1.64
	(1.22)	(0.97)	(1.11)
	[32]	[27]	[59]
Automation	2.56	2.18	2.35
	(1.19)	(1.07)	(1.13)
	[27]	[33]	[60]
Collapsed across Source	2.00	2.00	2.00
	(1.30)	(1.04)	(1.17)
	[59]	[60]	[119]

Panel B: Planned Simple Effects

Simple Effect	**t**	**Two-Sided p-value**
Effect of Automation on Agree with Preparer in the Mindset Absent Conditions	-3.29	**< 0.01**
Effect of Automation on Agree with Preparer in the Mindset Present Conditions	-1.51	0.14

Panel C: Analysis of Variance (ANOVA)

Source	**MS**	**F**	**Two-Sided p-value**
Automation	15.04	12.00	< 0.01
Counterarguing	0.12	0.90	0.76
Automation × Counterarguing	2.84	10.65	**0.14**
Residual	1.25		

Notes: Table 3 reports descriptive statistics and hypotheses tests for *Agree with Preparer*. Panel A provides the descriptive statistics by condition. Panel B reports the planned simple effects of automation depending on the *Counterarguing* conditions. Panel C reports an analysis of variance (ANOVA). See Appendix B for variable definitions.

TABLE 4
Seeded Errors Identified by Condition

Panel A: Descriptive Statistics: Mean (Standard Deviation) [N]

	Mindset Intervention		**Collapsed across Mindset**
	Absent	**Present**	
Human	2.75	1.96	2.39
	(2.49)	(2.61)	(2.55)
	[32]	[27]	[50]
Automation	1.59	2.61	2.15
	(2.10)	(2.47)	(2.35)
	[27]	[33]	[60]
Collapsed across Source	2.22	2.32	2.27
	(2.37)	(2.53)	(2.44)
	[59]	[60]	[119]

Panel B: Planned Simple Effects

Simple Effect	**t**	**Two-Sided p-value**
Effect of Automation on Time Spent in the Mindset Absent Conditions	1.91	**0.06**
Effect of Automation on Time Spent in the Mindset Present Conditions	-0.98	0.33

Panel C: Analysis of Variance (ANOVA)

Source	**MS**	**F**	**Two-Sided p-value**
Automation	1.95	0.33	0.57
Counterarguing	0.38	0.06	0.80
Automation × Counterarguing	23.90	4.05	**0.05**
Residual	5.91		

Notes: Table 4 reports descriptive statistics and hypotheses tests for *Seeded Errors Identified*. Panel A provides the descriptive statistics by condition. Panel B reports the planned simple effects of automation depending on the *Counterarguing* conditions. Panel C reports an analysis of variance (ANOVA). See Appendix B for variable definitions.

TABLE 5
Test of Spillover Effect (H3)

Panel A: Descriptive Statistics: Mean (standard deviation)

Condition	*n*	Reason-ableness	Surface issues	Deep issues	Total issues	Contact directly
Human / Absent	32	5.28	0.20	0.48	0.69	0.44
		(1.40)	(0.54)	(0.64)	(0.85)	(0.50)
Human / Present	27	6.07	0.43	0.37	0.80	0.52
		(1.30)	(0.58)	(0.58)	(0.86)	(0.51)
Automation / Absent	27	5.33	0.28	0.57	0.85	0.44
		(1.49)	(0.45)	(0.78)	(0.73)	(0.51)
Automation / Present	33	5.33	0.56	0.42	0.98	0.52
		(1.53)	(0.70)	(0.60)	(1.03)	(0.51)

Panel B: Negative binomial regression model: Z (robust standard errors)

Variable	Reason-ableness	Surface issues	Deep issues	Total issues	Contact directly
Automation	0.01	0.31	0.17	0.21	0.02
	(0.07)	(0.55)	(0.35)	(0.27)	(0.30)
Counterarguing	0.14**	0.74	-0.27	0.15	0.17
	(0.06)	(0.53)	(0.38)	(0.30)	(0.27)
Automation * Counterarguing	-0.14	-0.04	-0.03	0.00	-0.02
	(0.10)	(0.65)	(0.52)	(0.38)	(0.39)
N	119	119	119	119	119

Notes: Table 5 reports descriptive statistics and hypotheses tests for the spillover effect (H3). Panel A provides the descriptive statistics by condition. Panel B reports a negative binomial regression model. Robust standard errors are used. * $p \leq 0.10$, ** $p \leq 0.05$, and *** $p \leq 0.01$, all p-values are two-tailed.