Received: 9 June 2021 | Accepted: 10 December 2024

DOI: 10.1111/1911-3846.13020

ARTICLE

CAAA ACPC CONTEMPORARY ACCOUNTING RESEARCH | RECHERCHE COMPTABLE CONTEMPORAINE

The effect of process monitoring on beyond-the-job process improvements

Ivo Schedlinsky[1] | Maximilian Schmidt[2] | Friedrich Sommer[3] | Arnt Wöhrmann[2]

[1]LMU Munich, Munich, Germany

[2]Giessen University, Giessen, Germany

[3]University of Bayreuth, Bayreuth, Germany

Correspondence
Friedrich Sommer, University of Bayreuth, Nuernberger Strasse 38, 95448 Bayreuth, Germany.
Email: friedrich.sommer@uni-bayreuth.de

Abstract

Although it has always been important for firms that employees innovate predefined processes, the working environment in which employees implement these processes has significantly changed. Currently, the working environment is often characterized by employee surveillance; that is, the way in which employees conduct a process is monitored. In the current study, we present the results of an experiment examining the effect of process monitoring on process improvements by employees. Although previous accounting literature has reported negative effects of monitoring techniques on several organizational outcomes, we show that process monitoring can have a positive effect on employees' implementation of process improvements in the absence, but not in the presence, of a firm's error avoidance policy. Without an error avoidance policy, employees are motivated to create a favorable impression in front of management by implementing process improvements. This finding has important implications for business practice. From a broader perspective, we show that the influence of action controls depends on the parameters of a cultural control.

KEYWORDS

error avoidance policy, process improvements, process innovation, process monitoring

Au-delà de l'emploi : l'impact du suivi des processus sur les améliorations des processus

Résumé

Il a toujours été crucial pour les entreprises que les employés mènent à bien les processus prédéfinis, mais l'environnement

Accepted by Clara Xiaoling Chen.

de travail dans lequel les employés mettent en œuvre ces processus a énormément changé. À l'heure actuelle, l'environnement de travail est souvent caractérisé par la surveillance des employés, leur façon de mener un processus faisant l'objet d'un suivi. Cette étude présente les résultats d'une expérience examinant l'impact du suivi des processus sur les améliorations des processus par les employés. Alors que la littérature comptable rapporte des effets négatifs des techniques de suivi sur plusieurs résultats d'entreprises, les auteurs montrent que le suivi des processus peut avoir un effet positif sur les projets d'amélioration des processus mis en place par les employés, notamment en l'absence, et non pas en présence, d'une politique de l'entreprise axée sur l'évitement des erreurs. En l'absence d'une telle politique, les employés sont motivés par le désir de produire une impression favorable auprès de leur gestionnaire en mettant en place des projets d'amélioration des processus. Ces constatations fournissent d'importants éclaircissements sur les pratiques entrepreneuriales. Dans une perspective plus large, les auteurs soulignent que l'influence des contrôles d'action dépend des paramètres liés à un contrôle culturel.

MOTS-CLÉS
améliorations des processus, innovation des processus, politique axée sur l'évitement des erreurs, suivi des processus

JEL CLASSIFICATION
M14, M41, M50

1 | INTRODUCTION

"Everyone has two jobs. First to build the car, second to find ways of doing the job better" (Caulkin, 1993, p. 38). This quote from John Towers, former owner and managing director of the British car manufacturer Rover, pinpoints the importance of process improvements—that is, new and improved ways to solve a problem—through employees executing predefined processes (Anderson et al., 2014). Process improvements—also referred to as process innovations (Crossan & Apaydin, 2010; Miron et al., 2004)—represent a type of innovation. We examine the effect of process monitoring on employees' identification and implementation of process improvements. We focus on the implementation of process improvements because they are an important antecedent of long-term organizational innovation and performance. One important source of innovation is employees executing a specific process because these employees have detailed knowledge about a process and, thus, are a potential source for process improvements—both incremental and radical—that can positively influence firm performance (Crossan & Apaydin, 2010; Webb et al., 2013). However, the tasks of searching for and implementing such improvements are usually not explicitly contained in the job description (Anderson et al., 2004; Axtell et al., 2000). Therefore, whether employees engage in such behavior is influenced by management controls that shape the working environment.

Previous research in this field has focused on the effect of results controls on process improvements. For example, prior research suggests that individuals discover and implement more process improvements (represented by shortcuts) under fixed-wage contracts and low-productivity targets than under bonus contracts and/or high productivity targets (Webb et al., 2013). Furthermore, relative performance information seems to increase the implementation of process improvements (Newman et al., 2022). Although results controls are widespread and motivate employees to produce more output, organizations increasingly use action controls, such as process monitoring. These controls make employees' process execution—and not only inputs or outputs—visible to the firm. For example, a recent survey conducted in 2022 reports that 60% of surveyed workers faced some form of monitoring (Skelton, 2023). In the current study, we examine the effect of process monitoring on process improvements by employees.

An important reason for monitoring process execution, and not simply input or output, is to ensure that employees produce output using a specific, predefined process (Subašić et al., 2011). This is especially important in areas in which following predefined processes is indispensable for ensuring product quality (e.g., in pharmaceutical manufacturing) or safety (e.g., in chemical manufacturing or aviation). In these contexts, organizations do not want employees to experiment to implement process improvements because the risks associated with such experimentation are too high.

Notably, process monitoring is also used when predefined processes are not essential for product or personnel safety. Examples include recording data from workers' handheld scanners in warehouses (Gruet, 2024), point-of-sale IT systems in restaurants that detect and report theft and fraud (Pierce et al., 2015), or tracking the position of delivery trucks (Levy, 2015). Importantly, process monitoring has also become popular in white-collar or administrative jobs in response to the COVID-19 pandemic and the subsequent increase in remote work and telecommuting (Negrón & Nguyen, 2023; Skopeliti, 2023). Monitoring employees in this context includes AI-driven email analysis to assess workplace culture, programs that capture keystrokes (Negrón & Nguyen, 2023), monitoring via webcams, and audio recording (Skopeliti, 2023). The motive behind process monitoring in these jobs is not only surveillance, but also an attempt to increase productivity (Skopeliti, 2023).

The conflicting views of process monitoring suggest that it cannot be examined in isolation. Instead, other—especially cultural—controls determine their effect. A firm's error avoidance policy is an important cultural control that shapes how employees perceive process monitoring and ultimately determines their willingness to engage in process improvements. Although prior research has emphasized the importance of tolerating (or even rewarding) early failure to motivate exploration and innovation (Manso, 2011), an error avoidance policy represents a frequent manifestation of a corporate culture that has no tolerance for failure. Whether an organization employs an error avoidance policy will most likely depend on its willingness to take risks. Although a firm's willingness to take risks differs during the firm's life cycle (e.g., Habib & Hasan, 2017), this also depends on the industry. For example, aviation or pharmaceutical and chemical manufacturing firms are likely to have an error avoidance policy. In this study, we combine these two aspects and investigate whether the effect of process monitoring on beyond-the-job process improvements depends on the absence versus the presence of an error avoidance policy.

We predict that process monitoring promotes process improvements when an error avoidance policy is absent, but not when an error avoidance policy is present. Although process monitoring increases employees' perceived visibility, whether monitoring is regarded as a control or an opportunity for self-presentation depends on the presence or absence of an error avoidance policy. The self-presentation aspect relies on the fact that process monitoring allows individuals to display their skills as the working process becomes visible. Because individuals' self-esteem is influenced by their imagined reaction of others to their performance (Leary & Kowalski, 1990), prior literature suggests that the visibility of the working process motivates employees to engage in working behaviors that create a favorable impression (Bolino et al., 2008; Tetlock & Manstead, 1985).

In particular, in the absence of an error avoidance policy, employees have an incentive to engage in experimentation if it offers them the chance to be perceived as a smart person who can identify and implement process improvements. For example, in an analytical study linking incentives and firm growth dynamics, Bennett and Levinthal (2017) argue that employees might strive to "shine more brightly" than their peers through beyond-the-job behavior, such as process improvements. However, in the presence of an error avoidance policy, the error avoidance policy affects employees' beliefs about how to create a positive impression, that is, by following established or prescribed process specifications rather than engage in process improvements.

To test our prediction, we conducted an experiment with a 2×2 full factorial between-subjects design.[1] We used a newly developed task based on the popular board game "Mastermind," which requires participants to identify a hidden color code in, at most, 12 steps. We slightly adjusted the traditional rules to improve the structure of the task, helping us provide a general solution guide that allows participants to identify any color code in exactly 12 steps. However, the process description provided in the solution guide was not optimal because several process improvements allowed the participants to identify color codes in fewer than 12 steps. In contrast to previous tasks, this allowed us to measure the implementation of process improvements in a situation in which creativity was not the main aspect of the task.

We manipulated *Process monitoring* at two levels (present vs. absent). In the *Process monitoring* present condition, process-oriented video surveillance captured participants' screens. In the absence of *Process monitoring*, there was no video surveillance. Furthermore, we manipulated *Error avoidance policy* at two levels (present vs. absent). Although we provided a definition of "defective codes" (codes that are not identified within the maximum of 12 steps) to all participants to ensure clean manipulation, only under the *Error avoidance policy* present condition did we inform participants that it was important to avoid producing such "defective codes." We did not provide this information to participants in the *Error avoidance policy* absent condition.[2]

In line with our prediction, we find that if an error avoidance policy is absent (but not if it is present), process monitoring motivates employees to implement more process improvements. Consequently, the implementation of process improvements is greatest in the *Process monitoring* present/*Error avoidance policy* absent condition. The participants in this condition also exhibit the greatest drive to make an impression by making improvements, thus supporting the self-presentation aspect of process monitoring in the absence of an error avoidance policy. By analyzing participants' choices regarding specific steps involved in the task, we also show that participants who are subject to process monitoring in the absence of an error avoidance policy engage more in radical and incremental experimentation. However, process monitoring does not automatically lead to superior performance in the short term.

Our study makes three contributions to accounting theory and practice. First, we contribute to the growing stream of research examining the effects of management controls on employee innovation. Prior research in this field differentiates among different types of innovation, including process innovation (Kahn, 2018), which is the focus of our study. Guo et al. conclude that, despite its largely acknowledged importance, "process innovation has remained out of focus" (Guo et al., 2019, p. 682), and because of its differences vis-à-vis other types of innovation, such as product innovation, prior findings cannot be generalized to process innovation. Cai et al. (2023) add that motivating process innovation is particularly challenging because it

[1]The research was conducted in an ethical manner. Specifically, subjects were treated anonymously in accordance with the relevant data protection regulations and were not exposed to specific risks. Furthermore, subjects were not deceived in any way or at any time. The institution at which the study was conducted does not have a review board to provide ethical clearance.

[2]This form of manipulation (present vs. absent) follows prior research that analyzes such value statements (e.g., Kachelmeier et al., 2016) because these statements are usually only meaningful (and hence applied) in one direction. For example, a firm that is open-minded with regard to the idea that employees learn most effectively from mistakes would nevertheless not tell them that "it is important to produce errors."

represents extra-role behavior that is difficult to directly incentivize. Therefore, examining process monitoring appears particularly fruitful in this context, but has not been addressed by prior research. For example, Guo et al. (2019) use survey data to show that input controls (e.g., training), behavioral controls (e.g., quality controls), and output controls (e.g., use of indicators of innovation outcomes) have a positive effect on process innovation. Pfister and Lukka (2019) present a case study showing that a results control, that is, stretch targets, may lead to productivity gains through process innovations. However, the positive effect depends on the personnel and culture controls used by the firm. Overall, we contribute to this stream of research by examining the effects of an action control of growing importance—process monitoring—and its interaction with the firm's error avoidance policy (a form of cultural control) on process improvements.

Second, our study highlights the self-presentation aspect of process monitoring, which can induce process improvements. In this vein, our study particularly relates to prior findings of Campbell et al. (2011), who show via a field experiment that employees engage in less experimentation to improve processes under tight versus loose monitoring. Particularly, they examine two different cultures of handling deviations from a standardized procedure, with process monitoring being present in all conditions. This result is essential, as it suggests that cultural controls can impact experimentation when process monitoring is present. However, their study is not intended to generate any inferences regarding situations without process monitoring. We build on this research by examining the effects of process monitoring under different cultural controls.

Third, we inform practitioners about the potential costs and benefits of process monitoring. Specifically, we reveal that—for firms without an error avoidance policy—control mechanisms that enhance the visibility of working processes (i.e., process monitoring) can serve as useful tools for fostering innovative behavior. In contrast, we show that—for firms with an error avoidance policy in place—process monitoring can come with hidden costs; that is, employees become reluctant to think about ways to improve their working processes.

The current paper proceeds as follows. Section 2 provides background information on monitoring. Section 3 elaborates on the theoretical background and develops the hypothesis. Section 4 describes our experimental method, and Section 5 presents the results. Section 6 concludes the paper.

2 | BACKGROUND

Nebeker and Tatum speculate that employees "have been monitored at work probably as long as people have been employed" (Nebeker & Tatum, 1993, p. 508). Because of technological innovations and trends such as telecommuting, monitoring has increased in recent years—particularly in the digital age (Kensbock & Stöckmann, 2021). The literature provides many examples of how firms monitor their employees and highlights the extent to which firms use monitoring. For example, research reports that 37% of firms track the employees they send to customers (e.g., by tracking their mobile devices) (Ante & Weber, 2013) and that cameras in a Las Vegas casino track the smiles of employees to measure customer service quality (Peck, 2013).

Although monitoring is pervasive in the business world, some researchers argue that monitoring "is much used and little understood" (Vorvoreanu & Botan, 2020, p. 3). Monitoring is known to have physiological and psychological effects. The former category of effects includes physiological arousal, increased heart rate, changes in skin conductance, and, ultimately, stress reactions (see Stanton (2000) for an overview). The latter category includes, for example, employees' intrinsic motivation or willingness to speak up (Kensbock & Stöckmann, 2021). Prior research suggests that the effect of monitoring depends on the context in which it is used

(Stanton, 2000). In other words, employees' perceptions and the characteristics of monitoring determine whether monitoring has a positive or negative effect on employee behavior. For example, although a certain level of stress or performance pressure may appear beneficial, negative performance effects become more likely when monitoring is too intense and/or perceived negatively. Stanton (2000) provides an overview of theories and findings explaining why the situation in which monitoring is employed matters. In this context, this author highlights the importance of social facilitation.

Social facilitation research—significantly influenced by Zajonc (1965)—examines the effect of the social presence of others (e.g., via monitoring) on individuals' behavior and their performance (Aiello & Douthitt, 2001). Guerin (1993) classifies the various existing theories that try to explain social facilitation effects into drive theories, cognitive process theories, and social conformity theories. Briefly, drive theories argue that the mere presence of others increases arousal that translates into performance. Cognitive process theories focus on the distraction caused by the presence of others. The final group of theories assumes that in the presence of others, an individual's awareness of which behavior is considered valuable is strengthened. Self-presentation theory belongs to this latter group of theories and argues that individuals care about their image and how they are evaluated by others (Aiello & Douthitt, 2001). This is discussed in more detail when developing our hypothesis in the next section.

A related study by Campbell et al. (2011) illustrates how anticipated evaluation by others affects behavior. These authors examine employee behavior in a setting in which exception reports are generated when employees use their discretion to deviate from the guidelines established by the firm; they find that when tight (as opposed to loose) monitoring of the exception reports is implemented, employees experiment less; that is, they deviate less from the guidelines. This finding is consistent with the claim that greater evaluative pressure decreases experimentation. This setting is different from the one that we investigate for an important reason. In Campbell et al. (2011), each deviation from the firm's guidelines is reported to management via an exception report. The handling of these exception reports by supervisors influences participants' pressure to adhere to established rules. In the setting we investigate, only the result matters, while deviations from a specified process have no consequences. In addition, monitoring is always present in Campbell et al. (2011), while we examine behavior when monitoring is present versus absent.

3 | HYPOTHESIS DEVELOPMENT

A common approach to the classification of management control systems differentiates action, results, personnel, and cultural controls (Merchant & van der Stede, 2023). Below, we derive a prediction regarding the effect of process monitoring (i.e., an action control) on process improvements that depends on a firm's error avoidance policy (i.e., a cultural control).

We focus on a setting where the outcome of a prescribed process is observed by the firm (i.e., a results control is in place). This allows us to derive a clear prediction for the effect of the action control (i.e., monitoring how a process is executed) that is not mixed with a results control effect (i.e., monitoring whether the process outcome is defective). When the observable outcome of an established process is defective, the firm can easily conclude that the employee has not followed the firm's guidelines. If, however, the outcome is error free, the employee might have followed the prescribed process or made process improvements.

To predict the effect of process monitoring on employee behavior in the context of process improvements, we base our approach on social facilitation. As outlined above, prior research uses different theories to explain social facilitation effects. We rely on self-presentation theory for an important reason: The effect of monitoring on employee behavior resulting from self-presentation depends on what employees assume is expected from them. Self-presentation was popularized by the work of Goffman (1959), who emphasizes individual's effort to create a positive image on others. Aiello and Douthitt (2001) summarize Baumeister's (1982) argument "that in the presence of others, people are motivated by a desire to please those who are observing them and to construct a certain public image of themselves" (Aiello & Douthitt, 2001, p. 169). This indicates that the effect of process monitoring on process improvements depends on employees' belief in what behavior is expected from them to create a favorable image. In the following section, we explain how this expectation is shaped by the presence or absence of an error avoidance policy.

In the absence of an error avoidance policy, engaging in process improvements under process monitoring might be appealing to employees. Considering that process monitoring is an action control, individuals' actions, that is, their working processes, become visible. Visibility enables employees to create a desired impression of others who observe how they execute a process (Leary & Kowalski, 1990). Leary and Kowalski argue that when "under others' intense scrutiny, for example, people find it difficult *not* to think about the impressions others are forming" (Leary & Kowalski, 1990, p. 36). Ariely et al. (2008) add that individuals work more productively if the task is meaningful, for example, because it is valued by others. This suggests that process monitoring might motivate individuals to engage in working behaviors, that is, developing process improvements, that they believe will create a favorable impression (Bolino et al., 2008; Tetlock & Manstead, 1985). Indeed, process monitoring provides a platform for employees to present their skills and abilities.

Whether this effect occurs and employees engage in process improvements depends on how process monitoring is perceived. When the firm does not emphasize an error-free process (i.e., without an error avoidance policy in place), employees do not experience performance pressure to behave in a specific way; that is, process monitoring is not perceived as a measure to control conformity with the firm's process guidelines. Thus, the positive self-presentation effect of process monitoring can materialize, and employees are motivated to take personal initiative and demonstrate extraordinary commitment by "going the extra mile" (Fay & Frese, 2001; Parker et al., 2010). In this situation, individuals are encouraged to take personal risks, to experiment, and to learn from feedback (Edmondson, 1996, 1999). Process monitoring provides the opportunity to broadcast skills and encourages individuals to select a strategy that focuses on process improvements (Roberts, 2005).[3]

In this situation, employees can follow two different strategies to improve processes: incremental and radical strategies (Dewar & Dutton, 1986). While incremental strategies describe changes within the boundaries of the current approach, radical strategies call the current approach into question and propose a completely new way of solving an existing task (Crossan & Apaydin, 2010). Although radical strategies have greater potential for improvement (e.g., concerning the time or cost saved), they also entail greater risk in terms of errors (Song & Thieme, 2009). Hence, process monitoring in an environment without an error avoidance policy provides benefits (in the form of a good impression) to employees when they implement process improvements that potentially outweigh the risk of producing errors. Because employees can benefit from the use of both incremental and radical strategies by creating a good impression, we assume that both approaches are triggered by process monitoring when an error avoidance policy is absent.

When an error avoidance policy is present, however, employees perceive process monitoring differently. Employees perceive the act of following predefined guidelines as more appreciated by the firm. Thus, self-presentation motivates them to convey the impression that they follow prescribed process specifications. Under these circumstances, experimentation is associated with

[3]Roberts (2005) uses the term *image construction* to describe a concept integrating impression management theory and social identity theory. Because social identity is not part of our investigation, that is, the observer was the same person in all experimental sessions, we refrain from distinguishing between these two concepts.

a high level of perceived personal risk (Edmondson, 1999). In particular, perceived personal risk can materialize through the creation of an unfavorable impression by producing errors, which is strongly undesirable. Thus, they engage in less (risky) experimentation and stick to the behavior suggested by their job descriptions (Campbell et al., 2011). Process monitoring is perceived as a way to control conformity with an error avoidance policy. As a result, the controlling aspect of monitoring becomes more salient.

In summary, the effect of process monitoring on process improvements is contingent on the presence or absence of an error avoidance policy. Process monitoring generally allows for self-presentation. If process monitoring is accompanied by an error avoidance policy, employees perceive that strictly following the firm's guidelines and not engaging in process improvements creates the most favorable impression. However, if no error avoidance policy is in place, employees engage in process improvements to broadcast their skills. Hence, we formally state our hypothesis as follows:

> **Hypothesis.** Process monitoring promotes process improvements when an error avoidance policy is absent but not when an error avoidance policy is present.

4 | EXPERIMENTAL METHOD

4.1 | Experimental task

To answer our research question, we developed a new experimental task based on the popular two-player board game Mastermind. In Mastermind, one player (the "codemaker") generates a secret color code ("hidden code") by choosing four out of six differently colored pegs; that is, the codemaker puts four colored pegs in four (hidden) holes at one end of a decoding board. The goal of the other player (the "codebreaker," which is the role of participants in our experiment) is to identify this hidden code (Laughlin et al., 1982). Therefore, the codebreaker makes (educated) guesses about the hidden code on the pegboard by putting colored pegs in the four empty spaces of a row. Next, the codebreaker receives feedback on the similarity between the guess and the hidden code. The feedback informs the codebreaker how many colored pegs in that particular guess are in the same position as in the hidden code and how many pegs have the correct color (but are not placed in the right position). The feedback does not indicate any position-specific information. The codebreaker's task is to use feedback to identify the hidden code within the maximum number of allowed guesses.

Importantly, in the original version of the game, a color could appear twice in the same code, and free spaces are allowed. In addition, the codebreaker is not restricted concerning the type or order of changes to make after one guess (i.e., from one row to another). Although prior literature examines several different strategies for identifying the hidden code as quickly as possible depending on the parameters in place (Berghman et al., 2009; Laughlin et al., 1982), we reduced these possibilities for better measurement of process improvements. Hence, we added the requirement that—for each guess—participants could add only one color. This additional requirement provided a clear structure for identifying the hidden codes, hence making identifying the codes more strategic and preventing random guesses. Thus, for their first guess, for example, the participants had to choose four pegs of the same color. If participants added more than one new color that did not appear in their last guess, they were reminded of the requirement and asked to alter their guess. In addition, the hidden code always consisted of exactly four out of six differently colored pegs in our setting. Furthermore, free spaces were not allowed, nor did any colors appear twice in the same code. A screen of the experimental task is depicted in Figure 1.

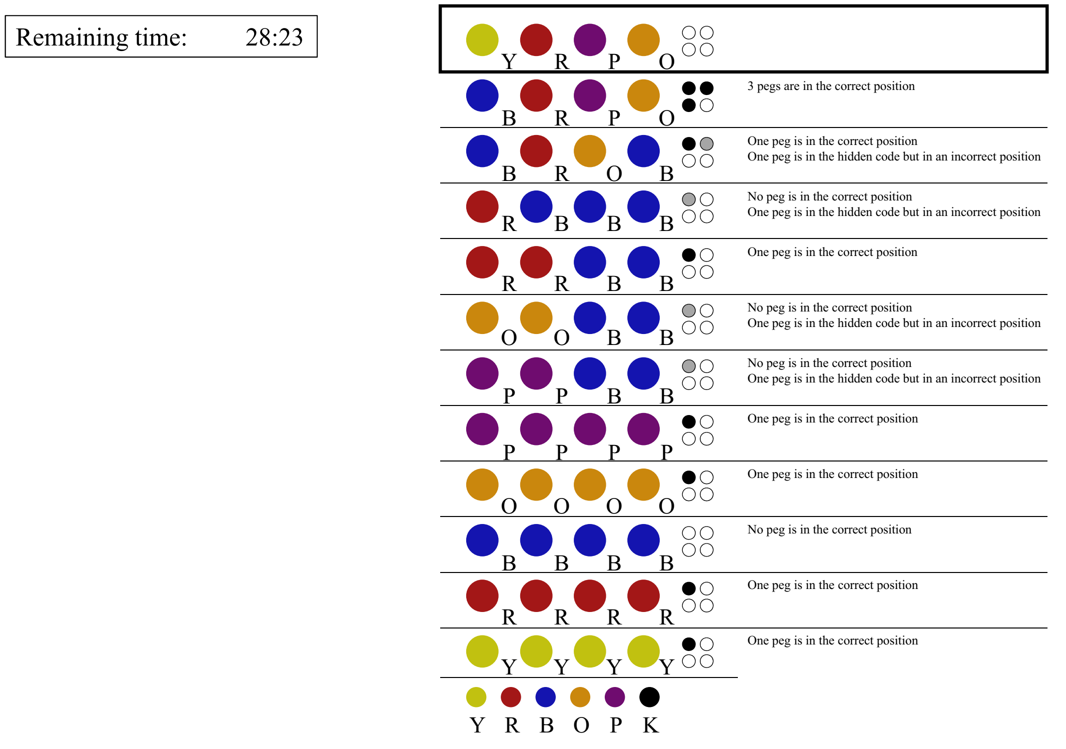

FIGURE 1 Screen of the experimental task. This figure shows a screen of the experiment. In Mastermind, one player (the "codemaker") generates a secret color code ("hidden code"), choosing four out of six differently colored pegs; that is, the codemaker puts four colored pegs in four (hidden) holes at one end of a decoding board. The goal of the other player (the "codebreaker," which is the role of participants in our experiment) is to identify this hidden code. The codebreaker makes (educated) guesses about the hidden code on the pegboard by putting colored pegs in the four empty spaces of a row. Next, the codebreaker receives feedback on the similarity between this guess and the hidden code. The feedback informs the codebreaker how many colored pegs in that particular guess are in the same position as in the hidden code and how many pegs have the correct color (but are not placed in the right position). The feedback does not indicate any position-specific information. To ensure comprehensibility in black and white printing, letters have been added to the original screen for the individual colors: Y (yellow), R (red), B (blue), O (orange), P (purple), and K (black).

In our setting, the hidden four-color codes were computer-generated, and participants took the role of the codebreaker. Each of the six available colors occurred only once in a code, resulting in $6!/(6-4)! = 360$ potential codes. The participants had a maximum of 12 guesses to identify a code. If participants identified a code correctly or were unable to identify a code within 12 guesses ("defective code"), they were informed of the outcome and then received a new (hidden) code. During the main task of the experiment, an unlimited number of hidden codes were available. All participants received the same codes in the same order. After each guess, the participants received feedback via a display of black and gray dots next to the row of their actual guess. For each colored peg in the correct position, a black dot was displayed. A gray dot appeared if the color existed in the hidden code but the position was incorrect. This feedback was also provided in written form (e.g., "One peg is in the correct position"). Because we aimed to investigate experimentation and process improvements, choosing a task that constantly provided feedback was an important design choice (Manso, 2011).

We provided a solution guide ("exemplary solution") to all participants that describes a structured process for identifying any hidden code within exactly 12 guesses—that is, within the maximum number of guesses allowed. Although this guide allows everyone to identify any code, it is not efficient: The participants could identify and implement process improvements and identify the hidden code in fewer than 12 guesses. Therefore, every deviation from the 12-step solution that also led to a correct answer represented a process improvement. Briefly, the exemplary solution followed a four-step strategy:

1. Rows 1–5 were used to identify all the colors contained in the hidden code.
2. Rows 6–8 were used to determine, for each color, if it was on the left or right side of the code.
3. Rows 9 and 10 served to identify the exact position of each color.
4. Rows 11 and 12 were used to insert all the colors in the correct places, as participants were allowed to add only one additional color per guess.

The pattern in Figure 1 displays a code that was solved using the process suggested by the solution guide.

Importantly, the exemplary solution allowed the user to solve any code without knowing how to interpret the feedback provided for every single step. Hence, the guide provided a structured process that—if understood and strictly applied—contained no risk of failing to solve any code within 12 guesses. Before the main task, all participants had to pass a tutorial in which they solved one code by precisely following the process proposed by the exemplary solution.

Process improvements in this task could take two forms, here based on the use of incremental or radical strategies. As noted in the hypothesis development section, incremental

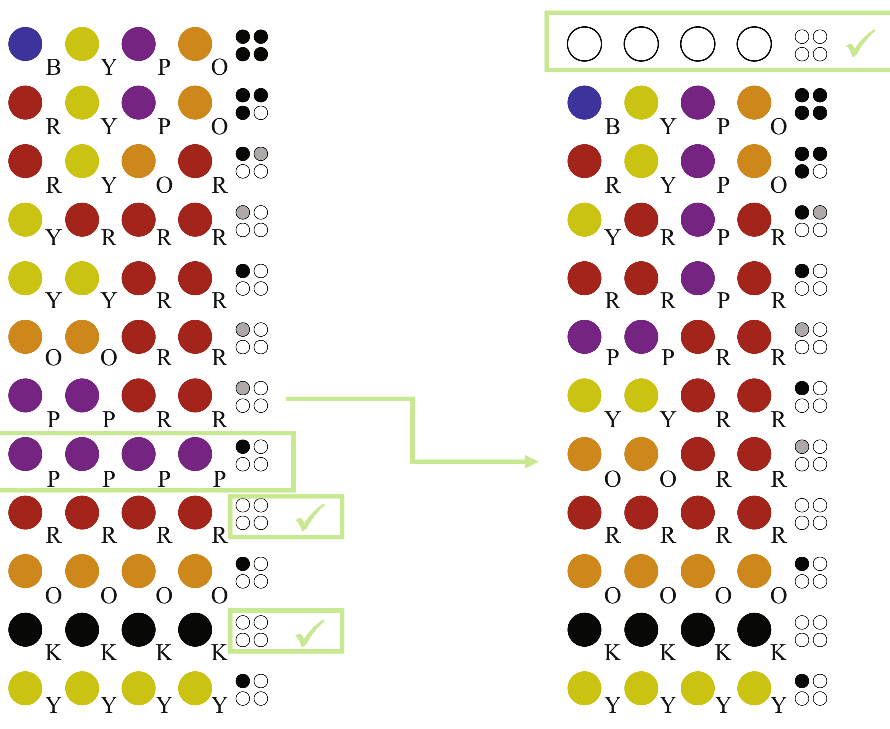

FIGURE 2 Process improvement example. This figure shows an example of process improvement. One can clearly identify which colors are contained in the color code as soon as the feedback ○○○○ appears twice when placing four pegs of the same color; at that point, one can deduce that all colors not yet tested are contained in the hidden code. To ensure comprehensibility in black and white printing, letters have been added to the original screenfor the individual colors: Y (yellow), R (red), B (blue), O (orange), P (purple), and K (black).

improvements represent specific ways of reducing the number of steps required to identify the hidden code without changing the overall strategy of the exemplary solution. An example of an incremental improvement in our task is depicted in Figure 2.

We define radical strategies as improvements that deviate from the core idea of solving hidden codes by using the pattern proposed by the exemplary solution (Steps 1–4 described above). An example of a radical strategy is using two different colors in the second row and adding further colors in the subsequent rows to gather information about their occurrence and their position simultaneously instead of solely identifying all hidden colors first. Using a radical strategy can help solve a hidden code problem using fewer guesses but contains a higher risk of failure (i.e., producing a defective code). Thus, the task allows for a broad spectrum of process improvements.

Because this task was new, all procedures, especially concerning the exemplary solution, were subjected to extensive pretesting to ensure their validity for the measurement of our variables.[4]

4.2 | Experimental manipulations

We tested our hypothesis via a computer-based laboratory experiment using a 2 × 2 full factorial design and SoPHIE software (Hendriks, 2012). We manipulated *Process monitoring* (absent vs. present) and *Error avoidance policy* (absent vs. present) between subjects.

When *Process monitoring* was present, participants worked under video surveillance. More precisely, we informed participants that a camera placed on their desks would capture their computer screens but not the participants themselves. The video was streamed to the computer of an observer in the same room. The observer could see and analyze participants' working processes. We further informed participants that a number affixed to their monitor allowed the observer to match the broadcast to their workspace. Importantly, the purpose of this video surveillance was to monitor the participants' working processes and analyze their work performance. However, participants knew that, in the context of the experiment, no further consequences (e.g., an influence on the participant's compensation or a confrontation with the observer) would result from the monitoring. Furthermore, we instructed participants to cover the camera on their desks with a felt bag provided at each workspace before they answered the post-experimental questionnaire. The participants in the *Process monitoring* absent conditions were not video-surveilled.

Concerning the second factor, we manipulated *Error avoidance policy* as being either present or absent. When *Error avoidance policy* was present, we instructed participants that it was important to avoid producing any defective codes (codes that are not solved within 12 guesses).[5] Notably, the provided exemplary solution depicted a process that allowed participants to solve any code within 12 guesses. Importantly, we did not assign any consequences (such as blame or other monetary or nonmonetary punishments) to the production of defective codes, here aiming to test the "pure" effect of process monitoring. The participants in the *Error avoidance policy* absent condition did not receive this instruction regarding defective codes. Importantly, to ensure information consistency, we defined defective codes for all participants in our experimental instructions. Thus, our manipulation is similar to that of Kachelmeier et al. (2016), who reemphasize the importance of accuracy, whereas we emphasized avoiding errors. As in those authors' research, our manipulation reemphasizes only the information that our incentive scheme inherently conveys.

[4]Pretesting helped design the materials in a way that made it easy for participants to understand the relatively complex task and especially the exemplary solution. Because it is vital for our research question that all participants fully understood the exemplary solution, the duration of the main task (30 min) was relatively short compared to the overall duration of the experiment (90 minutes). The main reason for this is that sufficient time was required for the tutorial, in which the participants were guided through the process of solving a color code following the exemplary solution. This insight was also gathered from pretesting.

[5]The exact wording is: "In this experiment, it is extremely important that you do not produce defective codes. Defective codes are color codes that are not identified within the maximum of 12 steps per code."

4.3 | Participants and procedure

In total, 112 business students participated in the experiment and were randomly assigned to one of the 4 treatment conditions.[6] Of these students, five who did not solve any code correctly were excluded from our analysis. The exclusion of these five participants was mandatory for our analysis because the measurement of process improvements was feasible only for solved codes.

On average, the participants were 23.74 years old, and 60.75% (39.25%) were male (female). There were no significant differences between conditions for *Age* ($p = 0.959$, two-tailed), *Gender* ($p = 0.744$, two-tailed), or *Academic degree* ($p = 0.506$, two-tailed). Additionally, we asked participants if they knew of the Mastermind game before the experiment. Twenty-four of the 107 participants (22.4%) indicated that they were familiar with the game, with no significant differences between conditions ($p = 0.595$, two-tailed). Hence, our randomization was successful.

After arriving at the laboratory, the participants had 10 min to read the written instructions. After that, they took a quiz to ensure that they understood the instructions. The participants then read the exemplary solution depicting a process that could be used to identify any hidden code within exactly 12 steps. To support the participants' understanding of this process, all participants had to pass a tutorial where they identified one hidden code by following the process given in the exemplary solution. During the tutorial, each step of the exemplary solution was also displayed on the screen. Moreover, participants received an interpretation of the feedback, such as "the feedback indicates that blue is located either at Position 1 or 2 within the hidden code." Only during this tutorial were participants required to apply the exact procedure suggested by the exemplary solution. The tutorial familiarized participants with the procedure for the main task and ensured that all participants solved at least one code using the process suggested by the exemplary solution before the main task started.

After the participants had passed the tutorial, they worked on the main task, which was identifying hidden color codes, for 30 min. Except for a notification that a code was identified correctly, we did not provide any further feedback during the task. After the main task, the participants had to answer a set of questions. We explicitly instructed participants in the treatments with process monitoring to cover the camera at their desk with a felt bag provided for this purpose because monitoring participants while answering the post-experimental questions could have influenced their answers. Afterward, participants were informed about their performance within the task and the resulting compensation and were allowed to leave the laboratory. In addition to a show-up fee of 1,600 lira (8 €), the participants received 150 lira (0.75 €) for each code identified correctly. On average, participants received 17.36 € for their participation.[7]

Notably, for all participants, irrespective of the treatment they received, our design created an incentive to avoid a simple trial-and-error strategy that leads to a high number of defective codes. First, the participants received extra compensation for each code they identified correctly. Second, the amount of time for identifying hidden color codes was limited to 30 min. Thus, there was an opportunity cost to invest time in the production of defective codes.

[6]Each treatment consisted of multiple sessions that were all administered within 4 days. To ensure all subjects participated under the same conditions, we instructed participants not to talk to others about the experiment, even after the end of the session. Nevertheless, to control for potential session effects, we test for differences between sessions within the different treatments. We find no significant differences for our main dependent variable, process improvements, and the number of codes solved correctly.

[7]All payment information given during the experiment was denominated in "lira," which is our experimental currency, and converted into euros at a rate of 200 lira per euro at the end of the experiment. At the time the experiment was conducted, 1.00 € equaled US $1.12.

5 | RESULTS

5.1 | Descriptive statistics

Table 1 presents the descriptive statistics. Our primary dependent variable is *Process improvements*. This variable represents the number of guesses a participant saved during the best try (compared with the 12 guesses required when following the exemplary solution). We rely on the best try to calculate our dependent variable (instead of total or average improvements) for an important reason: Both a participant's overall number of implemented process improvements in all codes and the average number of process improvements per color code contain a time aspect, that is, both measures are affected by the point of time when a process improvement is first identified and implemented. For example, identifying a process improvement when solving the first color code and constantly applying it when solving the following codes would lead to greater measures of overall and average improvements. Hence, these measures focus on performance at the individual level. Because we seek to examine the factors that influence whether process improvements are implemented in general, we use the best try of each participant to calculate our dependent variable. Thus, our variable represents the maximum number of guesses saved while solving one code and focusing on the implementation of process improvements.

The descriptive analysis reveals that participants working under *Process monitoring* in the absence of *Error avoidance policy* implemented the most process improvements. More precisely, the participants in this treatment saved on average 3.22 guesses during their best try, compared with 2.32 (*Process monitoring* absent/*Error avoidance policy* absent), 2.07 (*Process monitoring* absent/*Error avoidance policy* present), and 2.00 (*Process monitoring* present/*Error avoidance*

TABLE 1 Descriptive statistics per treatment.

		Error avoidance policy						
		Absent *Process monitoring*			Present *Process monitoring*			
		Absent	Present	Total	Absent	Present	Total	Total
Number of subjects (*n*)		25	27	**52**	27	28	**55**	**107**
Process improvements (12 less the number of guesses during the best try)	Mean	2.32	3.22	**2.79**	2.07	2.00	**2.04**	**2.40**
	SD	1.88	2.08	**2.02**	2.07	1.67	**1.87**	**1.97**
	Min	0	0	**0**	0	0	**0**	**0**
	Max	6	7	**7**	6	5	**6**	**7**
Performance								
Codes solved correctly	Mean	13.28	13.74	**13.51**	12.30	12.96	**12.63**	**13.07**
	SD	5.27	5.38	**5.28**	5.75	5.41	**5.54**	**5.41**
	Min	1	1	**1**	1	1	**1**	**1**
	Max	22	24	**24**	21	24	**24**	**24**
Codes solved incorrectly	Mean	0.60	1.19	**0.90**	1.07	0.82	**0.95**	**0.93**
	SD	0.96	1.49	**1.29**	1.66	1.33	**1.50**	**1.39**
	Min	0	0	**0**	0	0	**0**	**0**
	Max	3	5	**5**	7	5	**7**	**7**

Note: This table shows the main results per treatment and in total. The upper part presents the number of process improvements during the participants' best try. To compute this measure, we subtract the number of guesses a participant uses to identify the correct solution from the number of guesses necessary when following the exemplary solution (i.e., 12). We report arithmetic means (Mean), standard deviations (SD), minimum (Min), and maximum values (Max). The lower part of the table shows the number of codes solved correctly/incorrectly. The values in bold (Total) combine observations from different cells as indicated.

policy present) guesses saved in the other treatments. This pattern, which is depicted graphically in Figure 3, is in line with our hypothesis.

5.2 | Hypothesis test

To formally test our hypothesis, we follow a three-step approach. First, we conduct an analysis of variance (ANOVA) with *Process monitoring* (absent/present), *Error avoidance policy* (absent/present), and their interaction as the independent variables and *Process improvements* as the dependent variable. In addition, we conduct simple effects tests following the ANOVA to test the effect of partial variation against the total sample variance. Although this procedure follows the conventional approach used to analyze 2×2 full factorial designs, it can be seen as less powerful for testing ordinal interactions in which the predicted ordinal effect is based on the difference between one experimental cell and the three other experimental cells (Buckless & Ravenscroft, 1990). Our hypothesis complies with this pattern: When a firm does not have a strict error avoidance policy, employees engage more in process improvements when process monitoring is in place. In contrast, when a firm has a strict error avoidance policy, the level of process improvement is the same, irrespective of whether process monitoring is in place. Hence, in the second step, we conduct a planned custom contrast analysis to substantiate our hypothesis following the suggestions of Guggenmos et al. (2018). Because the need to specify contrast weights is subject to discussions in experimental accounting research, we ultimately verify our results by comparing regression coefficients, as suggested by Cohen (1983), in a third step.

The results of the ANOVA with *Process improvements* as our dependent variable and *Process monitoring*, *Error avoidance policy*, and their interaction as our independent variables are presented in Table 2, Panel A. As predicted, we find a significant interaction effect for our two independent variables on *Process improvements* ($F = 1.70$, $p = 0.0977$, one-tailed). Hence, we follow up by performing separate simple effects tests of the influence of *Process monitoring* in the *Error avoidance policy* absent/present conditions. In line with our hypothesis, *Process monitoring* leads

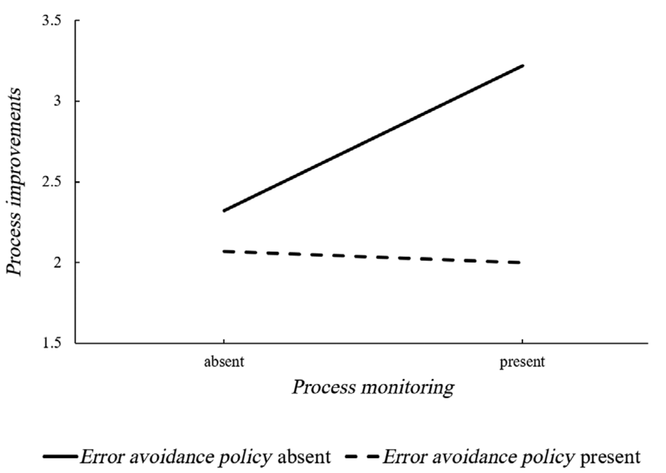

FIGURE 3 Graphical depiction of the effect on implemented *Process improvements* (best try) ($n = 107$). This figure graphically depicts the main results. When *Error avoidance policy* is absent, participants implement more process improvements (during their best try) when *Process monitoring* is present than when it is absent. However, when *Error avoidance policy* is present, the number of process improvements is similar, irrespective of the presence/absence of *Process monitoring*.

TABLE 2 Hypothesis test.

Panel A: ANOVA results for *Process monitoring* and *Error avoidance policy*

Dependent variable: *Process improvements* (best try) ($n = 107$)

Source	SS	df	MS	F-statistic	p-value[a]
Process monitoring	4.58	1	4.58	1.22	0.2716
Error avoidance policy	14.39	1	14.39	3.84	0.0527*
Process monitoring × *Error avoidance policy*	6.36	1	6.36	1.70	0.0977*
Total between-cells variance	25.33	3	8.44	2.25	0.0866*
Residuals	385.96	103	3.75		
Total	411.29	106			

Panel B: Simple effects following the ANOVA (Panel A)

Dependent variable: *Process improvements* (best try) ($n = 107$)

Source	t-statistic	p-value[b]
Error avoidance policy absent: *Process monitoring* (present vs. absent)	1.68	0.0482**
Error avoidance policy present: *Process monitoring* (present vs. absent)	−0.14	0.4438

Panel C: Planned contrast comparison between *Process monitoring* and *Error avoidance policy*

Dependent variable: *Process improvements* (best try) ($n = 107$)

Source	SS	df	MS	F-statistic	p-value[d]
Model custom contrast[c]	24.01	1	24.01	6.41	0.0129**
Residual between-cells variance	1.75	2	0.88	0.23	0.7919
Total between-cells variance	25.76	3	8.59	2.29	0.083*
Error	385.96	103	3.75		
Total between-cells variance	411.72	106			
Contrast variance residual, q^2	6.8%				

Note: [a]The *p*-value for the interaction effect is reported on a one-tailed basis because of the directional hypothesis concerning this effect. The *p*-values for the main effects are reported on a two-tailed basis because of the lack of a directional hypothesis regarding those effects.
[b]The *p*-values are reported on a one-tailed basis, given the directional hypothesis.
[c]The contrast coefficients are −1 for the *Process monitoring* absent/*Error avoidance policy* absent, −1 for the *Process monitoring* absent/*Error avoidance policy* present condition, +3 for the *Process monitoring* present/*Error avoidance policy* absent condition, and −1 for the *Process monitoring* present/*Error avoidance policy* present condition.
[d]The *p*-values are reported on a two-tailed basis.
* and ** represent significance levels of 0.10 and 0.05, respectively.

to more process improvements if *Error avoidance policy* is absent ($t = 1.68$, $p = 0.0482$, one-tailed), but not when it is present ($t = -0.14$, $p = 0.4438$, one-tailed) (Table 2, Panel B).

Second, we use planned custom contrasts to substantiate our findings, following the approach suggested by Guggenmos et al. (2018). We employ the following contrast weights: −1 for *Process monitoring* absent/*Error avoidance policy* absent, −1 for *Process monitoring* absent/*Error avoidance policy* present, +3 for *Process monitoring* present/*Error avoidance policy* absent, and −1 for *Process monitoring* present/*Error avoidance policy* present. These contrasts reflect our hypothesis, indicating a positive effect of *Process monitoring* in the absence of *Error avoidance policy* and a null effect of *Process monitoring* if *Error avoidance policy* is present.[8] As a first step, Guggenmos et al. (2018) suggest a visual test of the fit of the contrast

[8]As outlined before, we assume a setting where the process outcome is observable. Thus, the firm concludes from a defective code that the employee deviated from the prescribed process. Thus, *process monitoring* does not have an additional (negative) effect on *process improvements*.

weights to the descriptive data. Because our descriptive data indicates greater process improvements in the *Process monitoring* present/*Error avoidance policy* absent condition and only marginal differences among the three other cells, the contrasts representing our hypothesis match the descriptive pattern, allowing us to consider the actual significance of the planned contrast analysis. In support of our hypothesis, Table 2, Panel C, indicates that our data match the planned contrast weights ($F = 6.41$, $p = 0.0129$, two-tailed) (Table 2, Panel C). Third, our q^2 is 6.8%, suggesting that our contrast explains 93.2% of the between-cell variance.

Third, as another substantiation of the interaction without the need to specify contrast weights, we test whether the slopes for the effect of *Process monitoring* differ depending on the absence or presence of *Error avoidance policy* (Cohen, 1983). The effect size of *Process monitoring* is significantly greater ($z = 1.84$, $p = 0.017$, one-tailed) when *Error avoidance policy* is absent ($\beta_1 = 3.22$; standard error [SE] $= 0.492$) than when it is present ($\beta_2 = 2.00$; SE $= 0.448$) (untabulated). Again, we find support for our hypothesis.

5.3 | Test of mechanisms

In our hypothesis development, we argue that—in the absence of an error avoidance policy—process monitoring enforces the self-presentation effect; that is, process improvements are viewed as a superior strategy to create a positive impression compared with following the exemplary solution. Notably, Leary and Kowalski (1990) distinguish among three different motives for such self-presentation: social and material outcomes, the development of an identity, and self-esteem maintenance. The former two motives rely on building a reputation. Emler defines reputation as the "set of judgments a community makes about the personal qualities of one of its members" (Emler, 1990, p. 171). In the first case (social and material outcomes), these judgments or the reputation lead to desired outcomes, for example, approval by others, promotions, or salary increases. In the second case, reputation is built by executing identity-related activities (Gollwitzer, 1986), with identity being "ultimately derived from society" (Leary & Kowalski, 1990, p. 38). These reputation-related drivers of self-presentation cannot materialize in our (one-shot) setting, in which the participant remains anonymous to the observer and will thus not expect to receive any desired outcomes from this observer.[9] However, the third motive of self-presentation, that is, self-esteem maintenance, can materialize even in the absence of reputation-building incentives, as noted by Leary and Kowalski: "[S]elf-esteem is affected by people's self-evaluations of their performances and others' **imagined** reactions to them. People may perceive they have made a good or bad **impression** and experience resulting changes in self-esteem **in the absence** of explicit or implicit feedback from others"[10] (Leary & Kowalski, 1990, p. 37; emphases added). As this self-presentation effect materializes even without the possibility of building a reputation, we build on self-presentation motivated by self-esteem maintenance as the driving force in our setting.

Thus, we asked the participants to report on a 7-point Likert scale how much they agreed that (1) solving a code using fewer guesses than the number suggested by the exemplary solution and

[9]Several design choices prohibit the development of a reputation effect: First, although the observer could identify the participants within the experiment through numbers affixed to the screen, the experiment consisted of only one round, and, hence, not enough time to actually build a "reputation." In addition, the observer could not grant any social or material outcomes, such as bonuses or praise, to the participants, which would make building a reputation less attractive for the participants.

[10]This claim is also supported by empirical research. For example, Ariely et al. (2008) show that individuals work more productively if the task is meaningful—for example, because it is valued by another person who has no influence on compensation or the possibility of confronting the participants. The authors note that recognition "does not have to be linked to any financial incentives or to any nontangible rewards such as praise or appreciation" (Ariely et al., 2008, p. 672).

(2) following the exemplary solution creates a positive impression (1 = *do not agree*, 7 = *fully agree*). We calculated a variable by subtracting participants' indications regarding the exemplary solution from their indications regarding improving processes ((1)–(2)). We refer to this variable as *Impression through improvement*. The results are presented in Table 3. When *Error avoidance policy* is absent, *Impression through improvement* is greater when *Process monitoring* is present (1.56) than when it is absent (0.68). In the presence of *Error avoidance policy*, however, the participants indicated less perceived *Impression through improvement* under the *Process monitoring* present condition (0.25) than under the *Process monitoring* absent condition (1.33). We conduct an ANOVA with *Process monitoring*, *Error avoidance policy*, and their interaction as our independent variables and *Impression through improvement* as our dependent variable. In line with our argument, we find a significant interaction effect of *Process monitoring* and *Error avoidance policy* ($F = 4.54$, $p = 0.0355$, two-tailed), while the ANOVA does not indicate any main effects. Moreover, perceived *Impression through improvement* is significantly greater than 0 ($t = 4.11$, $p < 0.001$, one-tailed). Hence, process monitoring indeed enforces a self-presentation effect regarding process improvements in the absence of an error avoidance policy.

Similarly, we argue that process monitoring induces performance pressure and that this pressure is especially pronounced when an error avoidance policy is present. To test whether participants perceived more performance pressure, we asked the participants on a 7-point Likert scale how pressured they felt by monitoring while working on the task. The results of an ANOVA with *Process monitoring*, *Error avoidance policy*, and their interaction as the independent variables indicate a positive main effect of process monitoring (1.76 vs. 1.25, $F = 4.90$, $p = 0.0291$, two-tailed). A simple effects test following ANOVA further indicates that performance pressure is greater under *Process monitoring* when *Error avoidance policy* is present rather than absent (2.07 vs. 1.44, $p = 0.0522$, two-tailed). Hence, the performance pressure induced by an error avoidance policy influences the implementation of process monitoring.

TABLE 3 Test of mechanisms: *Impression through improvement*.

Panel A: Descriptive statistics (mean [standard deviation])

	Process monitoring		
Error avoidance policy	**Absent**	**Present**	**Total**
Absent	0.68 [2.10]	1.56 [2.53]	1.13 [2.35]
Present	1.33 [2.42]	0.25 [2.41]	0.78 [2.45]
Total	1.02 [2.27]	0.89 [2.54]	0.95 [2.40]

Panel B: ANOVA results for *Process monitoring* and *Error avoidance policy*

Dependent variable: *Impression through improvement* (*n* = 107)

Source	**SS**	**df**	**MS**	***F*-statistic**	***p*-value**
Process monitoring	0.29	1	0.29	0.05	0.82
Error avoidance policy	2.84	1	2.84	0.50	0.48
Process monitoring × *Error avoidance policy*	25.62	1	25.62	4.54	0.04**
Total between-cells variance	29.41	3	9.80	1.74	0.16
Residuals	581.36	103	5.64		
Total	610.77	106			

Note: All *p*-values are reported on a two-tailed basis.
** represents a significance level of 0.05.

5.4 | Additional analysis

Our additional analysis consists of four parts. First, we conduct robustness tests and a manipulation check. Second, we analyze the use of incremental and radical strategies for process improvements. Third, we analyze short-term performance effects. Finally, we discuss alternative explanations.

5.4.1 | Robustness tests and manipulation check

First, we provide a robustness test for our dependent variable. Instead of relying on the best try, we repeat our main analysis using the number of guesses saved in the best two (*TOP2*) and best three (*TOP3*) tries. Considering the *TOP2* (*TOP3*) tries of each participant, we find that participants saved 4.00 (5.52; *Process monitoring* absent/*Error avoidance policy* absent), 3.70 (5.04; *Process monitoring* absent/*Error avoidance policy* present), 5.41 (7.33; *Process monitoring* present/*Error avoidance policy* absent), and 3.68 (5.18; *Process monitoring* present/*Error avoidance policy* present) guesses. This pattern is quite similar to the pattern of our main dependent variable and is in line with our hypothesis. We conduct a planned custom contrast analysis according to the suggestions of Guggenmos et al. (2018), employing the same weights as those used in our main test. The results are statistically significant regarding the alternative dependent variables *TOP2* ($F = 4.35$, $p = 0.0395$, two-tailed) and *TOP3* ($F = 3.73$, $p = 0.0563$, two-tailed).

To validate our setting, we investigate whether individuals subjected to process monitoring perceived their working process as more visible. Accordingly, we asked the participants to indicate on a scale ranging from 1 (*do not agree*) to 7 (*fully agree*) whether they felt observed while working on the task. We find that individuals under *Process monitoring* conditions felt more observed (2.76) than individuals whose working process was not monitored (1.35) ($F = 19.21$, $p < 0.001$, two-tailed), with no significant differences between the two *Error avoidance policy* conditions. Hence, process monitoring increases the perceived visibility of the working process, irrespective of an error avoidance policy, and our manipulation of process monitoring worked as intended, which is the basis for the self-presentation and control effects.

5.4.2 | Incremental and radical process improvements

Although our hypothesis test provides evidence suggesting that process monitoring increases the implementation of process improvements in the absence of an error avoidance policy, it does not capture how process improvements are identified and implemented. According to our theory, process improvements can be incremental or radical, and the risk of producing defective codes is greater for radical improvements. To analyze how incremental and radical process improvements may be influenced by our independent variables, we conduct two analyses to compare participants' deviations from the exemplary solution between treatments. We do this by examining participants' second- and fifth-row guesses.

The exemplary solution suggested that participants should first identify all hidden colors by testing one color in each of the first five guesses. However, the participants could deviate from the exemplary solution and use a maximum of two colors in their second-row guess. Using two different colors in the second-row guess reflects a major deviation from the exemplary solution, which is (as participants have almost no information regarding the correct code at this stage) associated with high risk. This strategy represents a radical approach. We conduct an ANOVA featuring *Process monitoring*, *Error avoidance policy*, and their interaction as independent variables and *Average number of second-row guesses containing two colors* as the dependent variable.

The results suggest no main effects (all *p*-values >0.40, two-tailed) but indicate a significant interaction effect ($F = 2.98$, $p = 0.0875$, two-tailed). Using simple effects tests, we find that, if *Error avoidance policy* is absent, *Process monitoring* encourages participants to add a second color to their second guess more often (5.9%) than when it is absent (0.0%). This difference is statistically significant ($F = 3.04$, $p = 0.0841$, two-tailed). In contrast, when *Error avoidance policy* is in place, there is no significant difference between the presence and absence of *Process monitoring* (0.9% vs. 3.1%, $F = 0.46$, $p = 0.4976$, two-tailed). These analyses support our theory and suggest that process monitoring also fosters radical approaches to process improvements in the absence but not in the presence of an error avoidance policy.

Next, we analyze the implementation of incremental improvements in detail. The exemplary solution suggests that the fifth row is the last row that is filled with four pegs of the same color. Thus, using more colors in row five indicates experimentation, which may result in process improvements. However, because the participants already received feedback on several steps in the preceding guesses, the risk of deviating from the exemplary solution at this point is comparatively low, hence identifying this approach as an incremental approach to the implementation of process improvements. Using our main independent variables, *Process monitoring* and *Error avoidance policy*, we investigate *Average number of colors participants use in row five*. The ANOVA results indicate a significant interaction effect between *Process monitoring* and *Error avoidance policy* ($F = 3.37$, $p = 0.0693$, two-tailed). Simple effects tests underscored our prediction that *Process monitoring* prompts participants to use different colors in row five if *Error avoidance policy* is absent (1.50 vs. 1.16, $F = 8.81$, $p = 0.0037$, two-tailed), but not if it is present (1.26 vs. 1.22, $F = 0.18$, $p = 0.6752$, two-tailed). Hence, the increased number of process improvements when participants are faced with *Process monitoring* in the absence of *Error avoidance policy* is based not only on radical but also on incremental experimentation.

5.4.3 | Short-term performance effects

Next, we investigate participants' learning behavior over time and performance effects. Our theory suggests that process improvements stem from experimentation and learning. As part of the task validation, we can show that the number of guesses saved per code increases with the number of codes attempted ($p < 0.001$); that is, we observe learning behavior.

In the real world, process improvements should ultimately lead to increased performance. Thus, although this is not the focus of our study, we also investigate performance effects. Table 1 presents data regarding participants' *Performance* (i.e., the number of correctly solved codes). In addition, we report the number of incorrectly solved codes. The participants solved, on average, 13.07 codes correctly and 0.93 codes incorrectly. To investigate the effect of our manipulations on *Performance*, we conduct an ANOVA featuring the two independent variables of *Process monitoring* and *Error avoidance policy*, and their interaction. We find neither main effects nor an interaction effect (all *p*-values >0.40, two-tailed). Thus, process improvements do not lead to (short-term) performance effects in our setting. This finding may seem surprising at first. However, the self-presentation effect prompts individuals to engage in continuous experimentation. Although incremental changes build upon prior successful experimentation, especially radical changes in subsequent rounds are likely to produce worse results compared with prior rounds. In this context, employees identify process improvements and assess their relative performance potential over a longer period before selecting a preferred approach for the future. Hence, we assume that the (identified and tested) process improvements in our setting would (if proven successful) be continuously used and increase performance. However, such an effect is unlikely to materialize in a relatively short experiment. Another way to perpetuate improvements is for managers to implement them as part of the (formal) process specifications. We intentionally did not implement the second opportunity in the experimental design.

5.4.4 | Alternative explanations

To rule out alternative explanations, we analyze several mechanisms that could have affected the effect of monitoring on process improvements in our study.

To do this, we first look at the role of the observer when process monitoring is present. The effects of observation are a central part of our theory and reflected by our design. The videos in the *Process monitoring* present treatments were streamed to the screen of an observer in the same room, which could give rise to the possibility of an experimenter effect. To manipulate process monitoring by the inclusion of an observer *without* inducing an experimenter effect, we implemented a variety of procedures in our experimental design. First, although the observer was located in the same room, we established a clear physical distance between the workplace of the observer and the experimenter. The experimenter did not have the chance to look at the observer's screen. Accordingly, pleasing the experimenter was virtually impossible. Second, we used the different affiliations of the researchers involved in the current study. Although the experimenter was affiliated with the university at which the experiment was conducted, the observer (who was the same across all sessions) was not. Hence, the participants did not know the observer, and they would not expect to be evaluated by this observer in the future. The observer did not introduce himself (thus, the subjects were not aware of the fact that he was a researcher) and remained silent throughout the experiment; in contrast, the experimenter guided the sessions and provided instructions. Waag et al. (1973) report that observation effects are significantly weaker when subjects do not know the observer; hence, the reported effects are likely to be even greater in practice. We additionally analyze participants' answers in the post-experimental questionnaire regarding different aspects of observation. More precisely, we investigate whether the process monitoring employed in our experiment would influence perceived accountability and coercion of control. To analyze this aspect, we asked the participants to report on a 7-point Likert scale how important it was for them to perform well on the task (*Perceived accountability*) and how intensely they experienced monitoring while working on the task (*Coercion of control*). The results of an ANOVA featuring the independent variables of *Process monitoring*, *Error avoidance policy*, and their interaction do not indicate any significant influence of *Perceived accountability* (all p-values >0.40, two-tailed). This nonsignificant finding contradicts the presence of a possible experimenter effect because accountability would most likely be directed at "the person in charge." We find a positive main effect of *Process monitoring* with regard to *Coercion of control* (1.71 vs. 2.47, $F = 6.12$, $p = 0.015$, two-tailed), indicating that our manipulation of process monitoring evokes typical observation effects. However, we do not find positive interaction effects (all p-values >0.17, two-tailed). Hence, although we do find that the monitoring employed in the experiment influences *Coercion of control*, this effect does not interact with the absence or presence of an error avoidance policy and, therefore, does not serve as an alternative explanation for our findings.

Another source of noise could be the influence of trust between the experimenters and participants. Christ et al. (2008) argue that behavior control might harm trust. The participants in the *Process monitoring* present condition could perceive monitoring as a form of distrust, mistrust the experimenters and, hence, be less motivated to engage in process improvements. To rule out such an effect, we determine whether participants experienced mistrust in any way during the experiment. First, we asked participants whether they (1) felt trusted while working on the task and (2) trusted the experimenter. The results of two separate ANOVAs, including our main independent variables and their interaction, neither indicate a significant effect of *Process monitoring* on *Feeling of being trusted* (5.94 vs. 6.09, $F = 0.48$, $p = 0.4904$, two-tailed) nor *Trust in the experimenter* (6.27 vs. 6.55, $F = 2.36$, $p = 0.1279$, two-tailed) nor any other significant effect. Hence, we conclude that trust does not affect the results of our study.

Finally, because one could argue that our results might be influenced by individual characteristics, such as skill or risk preferences, we conduct additional randomization checks for these factors. First, we verify the randomization of skill. We do not find significant differences between treatments for the *Time participants needed to finish the tutorial* ($F = 0.49$, $p = 0.690$, two-tailed). Second, we asked participants to report their general *Risk preference* because implementing process improvements may arguably imply certain risks. We also do not find any differences in *Risk preference* between treatments ($F = 0.41$, $p = 0.745$, two-tailed).

6 | CONCLUSION

We investigate the effect of process monitoring on the implementation of process improvements via a laboratory experiment. In line with our prediction, we find that process monitoring motivates employees to engage in experimentation and implement process improvements beyond their job description only in the absence of an error avoidance policy.

Because of technological progress in the past decade, process monitoring has become popular for guaranteeing contractual agreements or providing additional services to customers. However, it entails hidden costs and benefits when it comes to improving processes within a firm. We offer guidance to firms that use or intend to use process monitoring. Acknowledging that strict error avoidance policies are necessary in certain industries, such as aviation or health care, we show that firms that do not require error avoidance might want to debate how to implement some form of process monitoring to exploit the hidden benefits we reveal.

Furthermore, we develop and validate a task allowing us to experimentally measure the beyond-the-job implementation of process improvements. Notably, in contrast to previous tasks that measured creativity by instructing participants to be creative, our task does not require any emphasis on the possibility of process improvements. Thus, the task creates opportunities for future research that aims to further exert beyond-the-job innovative behavior in different settings, an important dependent variable that previous accounting research could scarcely capture.

Future research could explore the generalizability and boundary conditions of our results. Although our theory builds on the self-presentation effect of process monitoring, this effect can be weakened or strengthened. For example, we employ a moderate form of performance-contingent pay. If the compensation per solved code was significantly increased, the opportunity cost of self-representation would also increase (and thus process improvements would decrease) because not following the sample solution increases the risk of “giving up” compensation. However, if the opportunity cost of self-representation was decreased by using fixed compensation, process improvements might increase. Furthermore, we use a loose type of process monitoring. Using tighter process monitoring, where deviations from a firm’s guidelines are still possible but must be explained, could also weaken the effects of our study.

Other future research opportunities arise from our design choices: Choosing a rather weak manipulation for an error avoidance policy provides a strong test for our theory but is also subject to an important limitation. When aiming to harness the positive effects of process monitoring on process improvements, firms should also consider other environmental factors (within their management control system). In particular, additional (nonmonetary) incentives, leadership style, or a low level of trust may also elicit employees to perceive process monitoring as controlling. Future research may want to investigate this further. Moreover, we use a simple piece-rate compensation scheme for all of our participants to hold the compensation constant across treatments. As a result, we cannot draw any conclusions about how process monitoring interacts with different rewards (or punishments) for errors. Hence, examining the link between rewards and process monitoring appears to be a promising avenue for future research.

Finally, laboratory experiments are limited by nature to a short duration. Hence, we do not predict or observe how process improvements turn into performance and thereby into company profits in the long run. Certain organizational structures, such as a company suggestion system, may facilitate the exploitation of the full potential of process improvements through experimentation. Analyzing these structures and their influence via a longitudinal study also seems to be a promising path for future research.

ACKNOWLEDGMENTS

We thank Clara Xiaoling Chen (editor), two anonymous reviewers, Katlijn Haesebrouck (discussant), Christoph Hörner (discussant), William Timothy Mitchell (discussant), Michael Williamson, participants at the 2019 Experimental Research in Management Accounting (EXRIMA) Summer School, the 2019 Accounting Behavior and Organizations Research Conference, the 2020 Management Accounting Section Midyear Meeting, the 17th Annual Conference for Management Accounting Research (ACMAR) and the VHB Annual Meeting 2020 for their helpful comments and suggestions. We appreciate the financial support provided by Dr. Werner Jackstädt-Stiftung and the Chartered Institute of Management Accountants (CIMA). Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

Data are available from the corresponding author upon request.

ORCID

Ivo Schedlinsky https://orcid.org/0000-0003-1031-9894
Maximilian Schmidt https://orcid.org/0000-0002-0166-5026
Friedrich Sommer https://orcid.org/0000-0002-7697-2864
Arnt Wöhrmann https://orcid.org/0000-0003-4157-5401

REFERENCES

Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, *5*(3), 163–180.

Anderson, N., De Dreu, C. K. W., & Nijstad, B. A. (2004). The routinization of innovation research: A constructively critical review of the state-of-the-science. *Journal of Organizational Behavior*, *25*(2), 147–173.

Anderson, N., Potočnik, K., & Zhou, J. (2014). Innovation and creativity in organizations: A state-of-the-science review, prospective commentary, and guiding framework. *Journal of Management*, *40*(5), 1297–1333.

Ante, S. E., & Weber, L. (2013, October 22). Memo to workers: The boss is watching. *The Wall Street Journal*. https://www.wsj.com/articles/SB10001424052702303672404579151440488919138

Ariely, D., Kamenica, E., & Prelec, D. (2008). Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization*, *67*(3–4), 671–677.

Axtell, C. M., Holman, D. J., Unsworth, K. L., Wall, T. D., Waterson, P. E., & Harrington, E. (2000). Shopfloor innovation: Facilitating the suggestion and implementation of ideas. *Journal of Occupational and Organizational Psychology*, *73*(3), 265–285.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, *91*(1), 3–26.

Bennett, V. M., & Levinthal, D. A. (2017). Firm lifecycles: Linking employee incentives and firm growth dynamics. *Strategic Management Journal*, *38*(10), 2005–2018.

Berghman, L., Goossens, D., & Leus, R. (2009). Efficient solutions for Mastermind using genetic algorithms. *Computers & Operations Research*, *36*(6), 1880–1885.

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, *34*(6), 1080–1109.

Buckless, F. A., & Ravenscroft, S. P. (1990). Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review*, *65*(4), 933–945.

Cai, W., Gallani, S., & Shin, J.-E. (2023). Incentive contract design and employee-initiated innovation: Evidence from the field. *Contemporary Accounting Research*, *40*(1), 292–323.

Campbell, D., Epstein, M. J., & Martinez-Jerez, F. A. (2011). The learning effects of monitoring. *The Accounting Review*, *86*(6), 1909–1934.

Caulkin, S. (1993, May 8). British firms resurrected by the courtesy of Japan. *The Guardian*.

Christ, M. H., Sedatole, K. L., Towry, K. L., & Thomas, M. A. (2008). When formal controls undermine trust and cooperation. *Strategic Finance*, *89*(7), 39–44.

Cohen, A. (1983). Comparing regression coefficients across subsamples: A study of the statistical test. *Sociological Methods & Research*, *12*(1), 77–94.
Crossan, M. M., & Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies*, *47*(6), 1154–1191.
Dewar, R. D., & Dutton, J. E. (1986). The adoption of radical and incremental innovations: An empirical analysis. *Management Science*, *32*(11), 1422–1433.
Edmondson, A. C. (1996). Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *The Journal of Applied Behavioral Science*, *32*(1), 5–28.
Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383.
Emler, N. (1990). A social psychology of reputation. *European Review of Social Psychology*, *1*(1), 171–193.
Fay, D., & Frese, M. (2001). The concept of personal initiative: An overview of validity studies. *Human Performance*, *14*(1), 97–124.
Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday-Anchor.
Gollwitzer, P. M. (1986). Striving for specific identities: The social reality of self-symbolizing. In R. F. Baumeister (Ed.), *Public self and private self* (pp. 143–159). Springer.
Gruet, S. (2024, January 23). *Amazon fined for "excessive" surveillance of workers*. BBC News. https://www.bbc.com/news/business-68067022
Guerin, B. (1993). *Social facilitation*. Cambridge University Press.
Guggenmos, R. D., Piercey, M. D., & Agoglia, C. P. (2018). Custom contrast testing: Current trends and a new approach. *The Accounting Review*, *93*(5), 223–244.
Guo, B., Paraskevopoulou, E., & Santamaría Sánchez, L. (2019). Disentangling the role of management control systems for product and process innovation in different contexts. *European Accounting Review*, *28*(4), 681–712.
Habib, A., & Hasan, M. M. (2017). Firm life cycle, corporate risk-taking and investor sentiment. *Accounting & Finance*, *57*(2), 465–497.
Hendriks, A. (2012). *SoPHIE: Software platform for human interaction experiments*. University of Osnabrueck.
Kachelmeier, S. J., Thornock, T. A., & Williamson, M. G. (2016). Communicated values as informal controls: Promoting quality while undermining productivity? *Contemporary Accounting Research*, *33*(4), 1411–1434.
Kahn, K. B. (2018). Understanding innovation. *Business Horizons*, *61*(3), 453–460.
Kensbock, J. M., & Stöckmann, C. (2021). "Big brother is watching you": Surveillance via technology undermines employees' learning and voice behavior during digital transformation. *Journal of Business Economics*, *91*(4), 565–594.
Laughlin, P. R., Lange, R., & Adamopoulos, J. (1982). Selection strategies for Mastermind problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(5), 475–483.
Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*(1), 34–47.
Levy, K. E. C. (2015). The contexts of control: Information, power, and truck-driving work. *The Information Society*, *31*(2), 160–174.
Manso, G. (2011). Motivating innovation. *The Journal of Finance*, *66*(5), 1823–1860.
Merchant, K. A., & van der Stede, W. A. (2023). *Management control systems: Performance measurement, evaluation and incentives*. Pearson.
Miron, E., Erez, M., & Naveh, E. (2004). Do personal characteristics and cultural values that promote innovation, quality, and efficiency compete or complement each other? *Journal of Organizational Behavior*, *25*(2), 175–199.
Nebeker, D. M., & Tatum, B. C. (1993). The effects of computer monitoring, standards, and rewards on work performance, job satisfaction, and stress. *Journal of Applied Social Psychology*, *23*(7), 508–536.
Negrón, W., & Nguyen, A. (2023). *The long shadow of workplace surveillance*. Stanford Social Innovation Review. https://doi.org/10.48558/CE5C-CA69
Newman, A. H., Stikeleather, B. R., & Waddoups, N. J. (2022). How relative performance information affects exploration-exploitation decisions. *Journal of Management Accounting Research*, *34*(1), 75–95.
Parker, S. K., Bindl, U. K., & Strauss, K. (2010). Making things happen: A model of proactive motivation. *Journal of Management*, *36*(4), 827–856.
Peck, D. (2013, December). They're watching you at work. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681/
Pfister, J. A., & Lukka, K. (2019). Interrelation of controls for autonomous motivation: A field study of productivity gains through pressure-induced process innovation. *The Accounting Review*, *94*(3), 345–371.
Pierce, L., Snow, D. C., & McAfee, A. (2015). Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science*, *61*(10), 2299–2319.
Roberts, L. M. (2005). Changing faces: Professional image construction in diverse organizational settings. *The Academy of Management Review*, *30*(4), 685–711.

Skelton, S. K. (2023, June 5). *Workers "deeply uncomfortable" with digital surveillance at work*. ComputerWeekly.com. https://www.computerweekly.com/news/366539253/Workers-deeply-uncomfortable-with-digital-surveillance-at-work

Skopeliti, C. (2023, May 30). "I feel constantly watched": The employees working under surveillance. *The Guardian*. https://www.theguardian.com/money/2023/may/30/i-feel-constantly-watched-employees-working-under-surveillance-monitorig-software-productivity

Song, M., & Thieme, J. (2009). The role of suppliers in market intelligence gathering for radical and incremental innovation. *Journal of Product Innovation Management*, *26*(1), 43–57.

Stanton, J. M. (2000). Reactions to employee performance monitoring: Framework, review, and research directions. *Human Performance*, *13*(1), 85–113.

Subašić, E., Reynolds, K. J., Turner, J. C., Veenstra, K. E., & Haslam, S. A. (2011). Leadership, power and the use of surveillance: Implications of shared social identity for leaders' capacity to influence. *The Leadership Quarterly*, *22*(1), 170–181.

Tetlock, P. E., & Manstead, A. S. (1985). Impression management versus intrapsychic explanations in social psychology: A useful dichotomy? *Psychological Review*, *92*(1), 59–77.

Vorvoreanu, M., & Botan, C. H. (2020). *Examining electronic surveillance in the workplace: A review of theoretical perspectives and research findings*. https://www.cerias.purdue.edu/assets/pdf/bibtex_archive/2000-14.pdf

Waag, W. L., Tyler, D. M., & Halcomb, C. G. (1973). Experimenter effects in monitoring performance. *Bulletin of the Psychonomic Society*, *1*(6), 387–388.

Webb, R. A., Williamson, M. G., & Zhang, Y. M. (2013). Productivity-target difficulty, target-based pay, and outside-the-box thinking. *The Accounting Review*, *88*(4), 1433–1457.

Zajonc, R. B. (1965). Social facilitation: A solution is suggested for an old unresolved social psychological problem. *Science*, *149*(3681), 269–274.

How to cite this article: Schedlinsky, I., Schmidt, M., Sommer, F., & Wöhrmann, A. (2025). The effect of process monitoring on beyond-the-job process improvements. *Contemporary Accounting Research*, *42*(2), 866–889. https://doi.org/10.1111/1911-3846.13020