

Bauch, Kai A.; Weißenberger, Barbara E.

Article

The effects of accountability on favoritism in subjective performance evaluations: An eye-tracking study

Schmalenbach Journal of Business Research (SBUR)

Provided in Cooperation with:

Schmalenbach-Gesellschaft für Betriebswirtschaft e.V.

Suggested Citation: Bauch, Kai A.; Weißenberger, Barbara E. (2025) : The effects of accountability on favoritism in subjective performance evaluations: An eye-tracking study, Schmalenbach Journal of Business Research (SBUR), ISSN 2366-6153, Springer, Heidelberg, Vol. 77, Iss. 4, pp. 915-944, <https://doi.org/10.1007/s41471-025-00231-8>

This Version is available at:

<https://hdl.handle.net/10419/338956>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



The Effects of Accountability On Favoritism in Subjective Performance Evaluations: An Eye-Tracking Study

Kai A. Bauch · Barbara E. Weißenberger

Received: 4 February 2024 / Accepted: 30 October 2025 / Published online: 27 November 2025
© The Author(s) 2025

Abstract This study uses an eye-tracking experiment to investigate the effects of accountability on favoritism in subjective performance evaluations. We develop and test theory suggesting that the effectiveness of accountability to alleviate favoritism depends on managers' cognitive processing of performance information. We argue that accountability will induce managers to examine performance information more thoroughly, which may entail having to scrutinize information that reflects unfavorable on preferred subordinates, thus causing cognitive dissonance. To resolve this dissonance, managers can engage in motivated reasoning by cognitively discounting such information. The more managers reduce cognitive dissonance by discounting unfavorable information, the lesser is the favoritism-mitigating effect of accountability. Consistent with this prediction, our results show that the higher the proportion of cognitive processing effort managers devote to information that reflects unfavorable on a preferred subordinate, the lower the efficacy of accountability in mitigating favoritism. We contribute by enhancing our understanding of the effects of accountability in subjective performance evaluations.

Keywords Subjective performance evaluation · Favoritism · Accountability · Eye-tracking · Cognitive processing

JEL Classification M12 · M49 · M52

✉ Kai A. Bauch
University of Kaiserslautern-Landau, Kaiserslautern, Germany
E-Mail: kai.bauch@rptu.de

Barbara E. Weißenberger
Heinrich Heine University Düsseldorf, Düsseldorf, Germany
E-Mail: barbara.weissenberger@hhu.de

1 Introduction

This paper reports on an eye-tracking experiment that we conducted to investigate how *accountability* affects *favoritism* in subjective performance evaluations. Favoritism describes the managerial behavior of giving advantage to some subordinates over others. It can take numerous forms in organizations, such as promoting preferred subordinates, evaluating their performance more leniently, or granting them more budgetary slack. For example, at the U.S. State Department's human resources bureau, there have recently been claims that friends and family are preferred in promotions such that in one case, apparently a close family member without relevant experience ended up in a position where she was completely overwhelmed (Heckman 2022). Moreover, recent survey evidence suggests that a majority of employees list favoritism from management as the number one factor that makes them feel unappreciated (Kahn 2022).

In subjective performance evaluation, favoritism occurs when managers' personal preferences towards subordinates affect the evaluations subordinates receive. That is, managers exploit their discretion and inflate evaluations of certain, preferred subordinates (e.g., Carmona et al. 2014; Kaplan et al. 2007; Ittner et al. 2003). While there is considerable evidence on the existence of favoritism in subjective performance evaluation (e.g., Du et al. 2012; Ittner et al. 2003), less is known about *how* managers use their discretion and deal with performance information when exerting favoritism as well as possible control mechanisms firms may implement to mitigate its occurrence. Investigating this is important, because research shows that bias in performance evaluation often leads to negative consequences on the subordinate side, such as decreases in perception of fairness, motivation, and work effort (e.g., Arnold and Tafkov 2019; Voußem et al. 2016).

Extant literature recurrently suggests accountability—in the form of having managers justify their performance evaluations to superiors—as an organizational control mechanism to counteract biases in subjective performance evaluation (e.g., Libby et al. 2004; Mero et al. 2003; Mero and Motowidlo 1995). For example, recent research shows that it can mitigate halo biases (Maske et al. 2021). However, we develop theory suggesting that the efficacy of such accountability in mitigating favoritism depends on managers' cognitive processing of performance information during performance evaluation of a preferred subordinate.

In line with a solid literature (e.g., Donnelly and Donnelly 2023), we expect that managers under low accountability tend to exert lower levels of cognitive effort than under high accountability. Consequently, we expect managers under low accountability to be likely to have their personal preferences towards a subordinate guide their evaluations. When not exerting high cognitive effort on processing performance information, such managers also avoid having to scrutinize information that potentially reflects negative on a preferred subordinate, which would cause cognitive dissonance (Festinger 1957). Further consistent with prior research, we expect, that when managers are held accountable, they devote more cognitive processing effort to performance information, as examining the information exhaustively is potentially necessary to justify the evaluation (e.g., Dalla Via et al. 2019; Schlenker 1986). Thus, when information is mixed, managers cannot simply refrain from also

scrutinizing the part of the information that reflects *unfavorably* on a preferred subordinate—which then in turn causes cognitive dissonance. We argue that the more managers reduce such cognitive dissonance, the less accountability will deter them from exerting favoritism.

We base our argument on theory of motivated reasoning, which suggests that individuals can reduce cognitive dissonance by interpreting information in a biased way: accepting information which is consistent with a preferred outcome at face value but exerting substantial scrutiny on inconsistent information (e.g., Hales et al. 2011; Kunda 1990). That means individuals who engage in such motivated reasoning devote an increased proportion of their cognitive processing effort to information which is inconsistent with a preferred outcome because they try to refute or discount the dissonance-causing information. Drawing on this body of research, we predict an ordinal interaction between managers being held accountable and their relative cognitive processing effort devoted to information that reflects unfavorably on a preferred subordinate in terms of favoritism. Specifically, we predict that when a manager evaluates a preferred subordinate and is held accountable (as opposed to not being held accountable), the greater the proportion of cognitive processing effort managers devote to information that reflects unfavorably on the subordinate, the lesser is the favoritism-mitigating effect of accountability.

We use a scenario-based experiment to test our theory. Participants in our experiment assume the role of regional managers whose task is to subjectively evaluate the performance of subordinates. To quantify favoritism, we rely on a pretest-posttest design and ask participants to evaluate multiple subordinates in an initial step of the experiment. In a second step, a specific subordinate (the ‘preferred subordinate’) is portrayed as likeable and being friends with one. This preferred subordinate then has to be evaluated. The preferred subordinate shows similar performance to a subordinate in the first step, that is, everything else is kept constant other than the description of the preferred subordinate. We measure favoritism as the difference in participants’ evaluations of the preferred subordinate and the respective evaluation in the first step. During the task of evaluating the preferred subordinate, we manipulate accountability (high vs. low) between subjects by asking half of the participants to provide a written justification of their performance evaluation to headquarters, whereas the remaining participants only provide a performance evaluation. We employ eye-tracking technology to track how participants allocate their cognitive processing effort towards information that reflects favorable and unfavorable on the preferred subordinate.

Our results are consistent with our prediction. We provide evidence that the higher the proportion of cognitive processing effort managers devote to information that reflects unfavorably on the preferred subordinate, the lower the efficacy of accountability in mitigating favoritism. Supplemental analyses show that this interaction is robust to various alternative specifications of our hypothesis test and provide support for our underlying theory.

The findings of our paper have important implications for both research and practice. First, our results refine our scarce knowledge of the effects of accountability in subjective performance evaluation. Prior research finds that accountability can be used as a debiasing mechanism in case of various biases arising (in subjective

performance evaluation) due to managers' use of cognitive simplifying heuristics (e.g., Fehrenbacher et al. 2020; Maske et al. 2021). Our findings extend this stream of management accounting literature by providing evidence that its efficacy may be less pronounced in case of favoritism. This finding also contributes to the literature on favoritism in general. While it has been argued that accountability might also be effective in reducing favoritism (Du et al. 2012), our research is, to the best of our knowledge, the first to investigate this. By developing theory and providing evidence that the efficacy of accountability in mitigating favoritism crucially depends on managers' cognitive processing and their potential attempts to discount information that reflects unfavorable on a preferred subordinate, our findings support contrary claims that "favoritism is extremely difficult to constrain" (Prendergast and Topel 1993, p. 361). This is important for organizations in practice to consider when designing their management control systems, as favoritism may emerge in many settings, including business (e.g., Robertson and O'Reilly 2024), healthcare (e.g., De los Santos et al. 2020), and the public sector (e.g., Szucs 2024) and may entail various negative consequences, such as a negative effect on leader-member exchange (Bolino and Hsiung 2014) or a decline in individuals' willingness to cooperate with each other (Mollerstrom 2022). Since firms cannot control how managers process performance information, our results suggest that firms in practice should not only rely on accountability but also implement additional controls to combat favoritism.

Our study also contributes to the broader literature on accountability in various settings (e.g., Pan and Patel 2022; Aleksovska et al. 2019; Li et al. 2022; Pérez-Durán and Grimmelikhuijsen 2025), which has a diverging view on its consequences, providing evidence that its effects on judgment and decision makings are ambiguous (e.g., Kadous et al. 2003; Kruglanski and Freund 1983; Curley et al. 1986). That is, while holding individuals accountable can lead to less bias in some cases (e.g., Libby et al. 2004), it may also lead to or amplify bias in other cases (e.g., Bartlett et al. 2014; Gordon et al. 1988). Using eye-tracking technology, we take a step further in understanding the potential behavioral consequences of accountability. Specifically, while our finding that holding managers accountable increases cognitive processing effort is consistent with extant literature (e.g., Dalla Via et al. 2019), our results also suggest that the extent to which this leads to less bias depends crucially on how the cognitive processing effort is allocated, as individuals may expend high cognitive processing effort only to interpret information in a biased manner consistent with their preferences. Thus, our cognitive processing perspective can help to understand the controversial findings of accountability documented in the literature (Lerner and Tetlock 1999).

2 Background and Theory

2.1 Background

We address a setting in which a manager subjectively evaluates subordinate performance by weighting multiple performance measures. Although granting managers subjectivity should allow them to produce more accurate performance evaluations

(e.g., by considering performance dimensions that cannot be measured objectively), prior research also suggests that managers' susceptibility to biases can cause performance evaluations to be less accurate when determined subjectively (e.g., Ding and Beaulieu 2011; Feldman 1981). Importantly, subjective performance evaluations always have an inherent personal component (i.e., "human judges other humans", Milkovich and Wigdor 1991). Yet accounting research addressing how the personal relationships between managers and subordinates affect performance evaluations is scant. A solid finding so far, however, is that favoritism can lead to bias in performance evaluations (e.g., Du et al. 2012; Bol 2008; Ittner et al. 2003). This likely entails negative consequences for the personal relationships in organizations (Bolino and Hsiung 2014; Mollerstrom 2022).

Favoritism in performance evaluation arises when managers, for example because of similarities, likeability, or other personal preferences, do preferred subordinates a favor by inflating their performance evaluations beyond their 'true' performance (Kaplan et al. 2007; Xu and Tuttle 2005). In other words, such preferred subordinates receive higher performance evaluations compared to other, non-preferred subordinates with similar performance.¹ While there is empirical evidence on the existence of such favoritism (e.g., Du et al. 2012), less is known about how managers use their discretion and how they deal with performance information when exerting favoritism as well as possible control mechanisms firms may implement to mitigate its occurrence. We investigate one such control mechanism, accountability, and how managers' cognitive processes when dealing with performance information under accountability, affect favoritism.

Accountability refers to "the implicit or explicit expectation that one may be called on to justify one's beliefs, feelings, and actions to others" (Lerner and Tetlock 1999, p. 255). Accountability can take various forms in performance evaluation systems. While outcome accountability describes managers being responsible for the result of their evaluation judgments only, process accountability entails managers being able to explain and justify their evaluation processes (Dalla Via et al. 2019). Regarding the 'recipient', managers can be held accountable towards a superior, a subordinate, or even multiple persons—for example in a calibration committee (Arshad et al. 2018; Mero and Motowidlo 1995). While managers being accountable for the outcome of their evaluations to their subordinates to some extent is the prevailing norm in practice (Hermans et al. 2021), we examine accountability in the form of explicitly having to justify the evaluation process to a superior. Prior literature shows that this form of accountability may mitigate some biases in performance evaluation. Libby et al. (2004) shows that managers overreliance on common (versus unique) performance measures can be reduced by holding them accountable. In a more recent study, Maske et al. (2021) show that halo bias in the form of good performance in one area spilling over to affect managers' evaluations of other areas can be mitigated by accountability. In contrast, the findings of Bartlett et al. (2014) suggest that accountability can also lead to dilution effects

¹ While at the heart of subjective performance evaluations is the fact that there is likely no single 'true' (i.e., correct) evaluation, this understanding rests on the fact that there is some consistency in performance evaluations when terming differences in evaluations between similar performing employees bias.

in performance evaluations, whereby managers put too much weight on irrelevant performance measures. Hence, accountability can have positive and negative effects on biases in performance evaluation. In our hypothesis development below, we argue that the efficacy of accountability in mitigating *favoritism* is limited and crucially depends on managers' cognitive processing.

Importantly, accountability is likely a continuum in practice ranging from virtually no accountability to full accountability. While we do not aim to focus on the very endpoints of this continuum, we develop theory that relates to environments in which accountability substantially differs and refer to these in the following as environments with low- or high accountability, respectively. In developing our theory, we connect components of the management (accounting) and the psychology literature. Regarding the latter, we draw from the broad field of research on cognitive dissonance and motivated reasoning (e.g., Festinger 1957; Kunda 1990). We connect this stream of literature with research on the effects of accountability on individuals' decision-making, mostly drawing on prior work in management and accounting, but also to some extent early psychology research on accountability (e.g., Kennedy 1995; Lerner and Tetlock 1999). We connect these literatures along a common claim in both strands of literature, that is, that the effects operate through individuals' level of cognitive processing effort, as outlined below.

2.2 Hypothesis Development

As explained above, favoritism in subjective performance evaluations originates from a manager's goal to do a preferred subordinate a favor by providing an inflated evaluation. We reason that managers under low accountability that have such a goal likely let their preferences towards a subordinate guide their evaluations. Multiple studies have shown that under low accountability, individuals tend to exert lower levels of cognitive effort than under high accountability (e.g., Donnelly and Donnelly 2023). Consequently, we also expect that when evaluating a preferred subordinate, managers that operate under low accountability will exert comparatively little cognitive effort. Moreover, in performance evaluation, a vast literature on cognitive biases shows that managers that are not held accountable tend to be susceptible to effort-reducing heuristics when subjectively evaluating subordinates' performance (e.g., Lipe and Salterio 2000). With regard to favoritism, the findings of Kaplan et al. (2007) suggest that the affect towards a preferred subordinate can guide managers' evaluations. Consistent with this, we expect that managers under low accountability exert comparatively little cognitive effort, superficially process performance information, and let their preference towards a subordinate guide their performance evaluations. As a byproduct, that way, managers also do not have to scrutinize information that potentially reflects negative on a preferred subordinate, which would cause cognitive dissonance (Festinger 1957).

We reason that when managers are held accountable, however, they likely exert higher cognitive effort. Accountability has been termed an "effort-inducing incentive[s]" (Kennedy 1993, p. 233) and research repeatedly suggests that accountability causes more effortful cognitive processing because information to justify a judgment or decision has to be studied exhaustively (e.g., Dalla Via et al. 2019; Schlenker

1986) and managers have to be able to refer to given information and justify the (non-)use of it in their performance evaluations.² While holding managers accountable likely leads to greater cognitive processing effort, we argue that accountability will not necessarily lead to reduced favoritism *in general*. When managers exert greater cognitive processing effort, it follows that they will oftentimes also have to thoroughly examine information which reflects less positive or even negative on a preferred subordinate. As a result, managers cannot easily inflate evaluations while avoiding cognitive dissonance to arise from examining performance information indicating that a worse evaluation might be warranted (Festinger 1957). We reason that the effect of accountability on favoritism crucially depends on how managers allocate their cognitive processing effort in order to deal with cognitive dissonance-causing information.³ Specifically, we argue that the mitigating effect of accountability is reduced if managers devote a higher proportion of their cognitive processing effort to information that reflects unfavorable on a preferred subordinate.

When individuals pursue a directional goal such as evaluating a preferred subordinate favorably, they often tend to *interpret* information in an asymmetric manner: they accept information which confirms their beliefs at face value but exert substantial scrutiny on disconfirming evidence (Hales et al. 2011; Kunda 1990).⁴ Specifically, a solid body of psychology research suggest that individuals engage in motivated reasoning by devoting increased scrutiny to disconfirming information (e.g., Ditto et al. 1998; Ditto and Lopez 1992). Individuals do so because they try to refute or discount the dissonance-causing information, a behavior which has been shown effective in reducing cognitive dissonance that evokes from having to process disconfirming information, especially when the amount of disconfirming information is low (Fischer et al. 2008; Ditto et al. 1998; Ditto and Lopez 1992). Thus, based on this evidence, we argue that when a manager evaluates a preferred subordinate and is held accountable, the higher the proportion of cognitive processing effort

² Consequently, prior work also finds that accountability can counteract biases that evoke due to managers' reliance on 'cognitive simplifying heuristics', whereby managers take simple shortcuts to avoid exerting high cognitive effort when engaging in subjective performance evaluations and having to process multiple performance measures (e.g., Libby et al. 2004). In contrast to favoritism, these heuristics evoke predominantly unconsciously and are not the result of a manager's directional goals.

³ For this very reason, we do not predict a general main effect of accountability on favoritism. Such an unconditional main effect is exactly what our theory put forward in this section contradicts, respectively nuances.

⁴ Such motivated reasoning behavior may be especially prompted under high accountability. Specifically, prior research implies that holding individuals accountable may cause a defensive mindset, motivating managers to defend and justify their existing opinions (Palmer and Feldman 2005; Kennedy 1995; Tetlock and Boettger 1989). In this vein, in auditing, for example, Donnelly and Donnelly (2023, p. 43) recently outline that "the increased effort accountable auditors put forth in the decision-making process does not necessarily result in decision outcomes that significantly differ from auditors who were not accountable." Holding individuals accountable may even exacerbate the effect of directional goals (Boiney et al. 1997). That is, prior work indicates that factors designed to encourage individuals to exert more cognitive effort can in fact increase the amount of bias caused by underlying directional preferences (Peecher et al. 2010; Kunda 1990). For example, prior work in auditing finds that auditors which have a directional preference to accept an aggressive, client-preferred accounting method are even more likely to accept the method when asked to formally identify the most appropriate method and provide a written justification previously (Kadous et al. 2003).

devoted to performance information that reflects unfavorable on the subordinate, the more she is likely able to reduce cognitive dissonance and self-justify inflating the evaluation and hence the lesser is the favoritism-mitigating effect of accountability.

An additional factor that might foster that accountability mitigates favoritism is that holding managers accountable can cause them to adjust their behavior in favor of the preferences of the superior to whom they are held accountable (Lerner and Tetlock 1999; Tetlock 1983). The reason is that such adjustment of behavior increases the probability of being perceived as a good manager by one's superior, which, in turn, leads to positive feelings such as self-esteem maintenance (Tetlock et al. 1989). Being perceived as a good manager may also entail other, extrinsic benefits such as a higher probability of being promoted or eligibility for additional bonuses. Managers are likely aware that from a superior perspective, favoritism is undesired. Thus, when held accountable, managers' perceptions about their potential *own benefits* when refraining from favoritism likely work against their goal to do a preferred subordinate a favor, which, in turn, contributes to the mitigating effect of accountability on favoritism. Still, we conjecture this is eventually unlikely to be a major driver in subjective performance evaluations. The reason is that once managers have self-justified their favoritism by engaging in motivated reasoning and cognitively discounting information that reflects unfavorable on a preferred subordinate, actually justifying a biased evaluation to superiors is likely less of a hurdle, as in multi-layered organizational settings, managers usually have an informational advantage over their own superiors with regard to their subordinates. Thus, even when managers are held accountable and have to justify their performance evaluations to their superiors, managers can potentially disguise favoritism in their evaluations from their superiors, for example by referring to alleged private, non-verifiable information (Prendergast and Topel 1993).

Taken together, the above discussion suggests an ordinal interaction between managers being held accountable and their proportion of cognitive processing effort devoted to unfavorable information. That is, whether the higher cognitive processing effort managers exert under high accountability (compared to low accountability) leads to less favoritism depends on how managers deploy their increased cognitive processing effort, that is, the extent to which they engage in motivated reasoning. We summarize this discussion in our formal prediction below:

Hypothesis: The greater the proportion of cognitive processing effort managers devote to information that reflects unfavorable on a preferred subordinate, the lesser will be the mitigating effect of accountability on favoritism in their subjective performance evaluations.

3 Method

3.1 Experimental Design and Task Overview

To test our hypothesis, we use an experiment in which we manipulate accountability between subjects (on two levels: high vs low) and measure the proportion of

Performance measure	Target	Actual	+/-
Sales growth	15%	14%	-
Sales margins	66%	73%	+
Percentage of sales from new products	24%	24%	=
Inventory turnover	7,2	8,3	+
Customer satisfaction	85%	94%	+
Employee satisfaction	84%	92%	+

Fig. 1 Performance Report. Notes: This figure presents a performance report containing target and actual values as well as performance markers (+/=/-) attached to these values as it has been presented to participants in the experiment

cognitive processing effort devoted to unfavorable information using eye-tracking technology.⁵ Participants assume the role of a regional manager of a hypothetical firm tasked with evaluating subordinate performance. We provide participants with performance reports which consist of multiple performance measures. Participants have discretion in weighting these measures. Figure 1 depicts a performance report as used in our study. The dependent variable is *favoritism* and is measured using a pretest-posttest design as described in detail below. During the task of evaluating a specific, likeable subordinate, we manipulate *accountability* (high vs. low) between subjects by asking half of the participants to additionally provide a written justification of their performance evaluation to headquarters whereas the other half provides a performance evaluation only.⁶ We track participants' eye-movements throughout the whole experiment using eye-tracking technology to measure their cognitive processing unobtrusively.

3.2 Participants and Procedures

In total, 111 students from a large, public university participated in the experiment. Similar to related research, we deem students appropriate participants for our study because our theory focusses on general human cognitive processes (e.g., Dalla Via

⁵ The authors were not required to obtain ethical approval and there was no institutional review board at their school at the time data was collected. However, the authors adhered to all applicable ethical standards in the field.

⁶ As outlined above, accountability is likely a continuum in practice ranging from virtually *no* accountability to full accountability. While our low accountability condition might appear to be an operationalization of no accountability at first glance, we argue that it is actually more reflective of low accountability. The reason is that our complete experimental task is focused on evaluations, and it is thus reasonable to believe that participants—on average—expected to some extent that they might be called to justify their prior behavior (e.g., in the post experimental questionnaire), similar to managers in practice that might oftentimes expect to be asked to justify an evaluation, although this accountability is not a *formalized* process.

et al. 2019; Libby et al. 2004, 2002). Participants were recruited from a self-enrolled participant pool of the university's behavioral laboratory. They were invited to participate via email and informed that the experiment features eye-tracking. They signed up for one of multiple sessions unaware of any treatments and were thus randomly allocated to conditions. Participants are on average 26.2 years old and 56.4% are female. Participants received a fixed payment of € 10 for their participation. This design choice was guided by extant literature and also represents the fact that managers in practice are typically not directly incentivized for their subjective performance evaluations. Similar to related eye-tracking research (e.g., Kramer and Maas 2020; Dalla Via et al. 2019), the eye-movements of three participants could not be tracked and thus these participants are excluded from all analyses, leading to a usable sample size of 108.⁷

The experiment was conducted in a laboratory. Participants were seated in separate cubicles, which were equipped with a computer that had an eye-tracker mounted beneath the monitor. Each session started out with an instructor reading aloud the basic instructions that informed participants on the eye-tracking specifics. Before proceeding, participants had to give their written consent. General instructions followed on the computer screen. In a second preceding step, participants were calibrated by fixating their eyes on various points appearing on the computer screen to ensure reliability of the eye-tracking measurement (cf. Holmqvist et al. 2011).

The actual experiment then started by exposing participants to the scenario and the managerial role to be assumed as well as the performance evaluation task and the respective performance measures. All materials were provided on the computer screen. Before participants were allowed to proceed to the experimental task, they had to pass a short quiz to reinforce their understanding. Participants were not allowed to proceed until they answered the quiz correctly. After all participants of a session had finished the task and completed the post-experimental questionnaire, they collected their payment separately. In sum, procedures took approximately 45 min.

3.3 Task and Materials

Participants in the experiment assume the role of a regional manager in a fictional company (named "TDC") and are tasked with evaluating subordinate employees'

⁷ The demographic data further shows that student participants are on average in their seventh semester, have three years of (full- or part-time) work experience. 69% of them have German as their native tongue and 27% pursue a major in business or economics. There are no significant differences between the accountability low and high condition for the demographic variables as well as general leniency (i.e., participants' average evaluation of the 12 hypothetical subordinates in the initial evaluation period), or attention towards the performance reports in the initial evaluation period (all p 's > 0.1, two-tailed) but gender ($p = 0.06$, two-tailed) and term of study ($p = 0.02$, two-tailed) do differ across conditions. However, our hypothesis test reported below is robust to including term of study and/or gender as covariates, while term of study is insignificant ($p = 0.98$, two-tailed) and gender (coded 1 for females) has a marginally significant negative effect on favoritism in the regression ($p = 0.09$, two-tailed). The finding that women tend to exert less favoritism when evaluating the preferred subordinate is consistent with prior research in psychology that shows that when evaluating male subordinates, female managers tend to be harsher than male managers (Furnham and Stringfield 2001). We further discuss this finding in the results section.

performance. The performance reports that they receive as an information basis are adopted from Cardinaels and van Veen-Dirks (2010). In particular, the performance measures employed are ‘sales growth’, ‘sales margins’, ‘percentage of sales from new products’, ‘inventory turnover’, ‘customer satisfaction’, and ‘employee satisfaction’.⁸ Right next to the target and actual values for the performance measures, we add performance markers (+, =, –, indicating above target, on target, and below target, respectively) similar to Cardinaels and van Veen-Dirks (2010). The performance markers enable participants to easily (i.e., using little cognitive effort) distinguish positive from negative performance information.

The actual experimental task consists of four steps. In an initial evaluation period, participants are asked to evaluate 12 hypothetical subordinates based on their performance reports.⁹ The initial step serves two purposes. First, it allows participants to become familiar with the evaluation task. Second and more important, it is crucial for our measurement of *favoritism*, as we will outline in more detail during the variable measurement section. Participants have to rate the subordinates on a continuous 101-point scale that has been used by prior research (e.g., Ding and Beaulieu 2011; Lipe and Salterio 2000), ranging from very poor performance (0) to excellent performance (100).

In a second step, a specific, likeable subordinate (‘Michael Schmitz’) is portrayed by presenting a short profile and outlining various personal incidents that indicate that one is friends with him. The respective part of the materials is presented in the appendix.¹⁰ Since we do not manipulate the presentation of this subordinate between subjects, all participants evaluate him. We conducted pretests to ensure that Michael Schmitz is perceived as likeable on average. This is corroborated by post-experimental questionnaire data as outlined in the Results section.¹¹

Subsequently, we announce our accountability manipulation similar to Bol et al. (2016) by providing participants with an email from the company’s CEO. The email for participants in the *accountability high* condition outlines that headquarters decided to request written justifications for upcoming performance evaluations.¹²

⁸ Similar to Maas and Verdoorn (2017), we picked measures from Cardinaels and van Veen-Dirks (2010) that seemed easy to understand and slightly altered the wording where deemed necessary.

⁹ We adopt the approach from Maas and Verdoorn (2017) to adapt the target and actual values from Cardinaels and van Veen-Dirks (2010): target values of the initial evaluation period subordinates’ performance measures vary within –20% to +20% of Cardinaels and van Veen-Dirks’ (2010) targets while actual values vary between –15% and +15% of targets. This variation only refers to the 12 hypothetical subordinates. It was held constant for *all participants*, both in terms of order and target/actual performance values.

¹⁰ The original language of the materials was German.

¹¹ Prompting favoritism using a scenario is challenging and likely yields lower levels of favoritism than potentially observed in real-world settings, where managers factually interact with their subordinates. However, this likely results in a lower level of favoritism *on average* and as such cannot explain the interaction we predict. It is also a conservative design choice, as a lower level of favoritism on average works against us finding support for our prediction.

¹² At this stage, another between-subjects manipulation was performed by providing a message that the CEO acknowledges managers’ efforts in providing the performance evaluations or not. However, this manipulation did not significantly affect favoritism or interact with our other independent variables in our hypothesis test (all p ’s > 0.1). Finally, fully crossing this variable with our other independent variables does not affect any of the inferences that we draw from our hypothesis test. In line with extant research

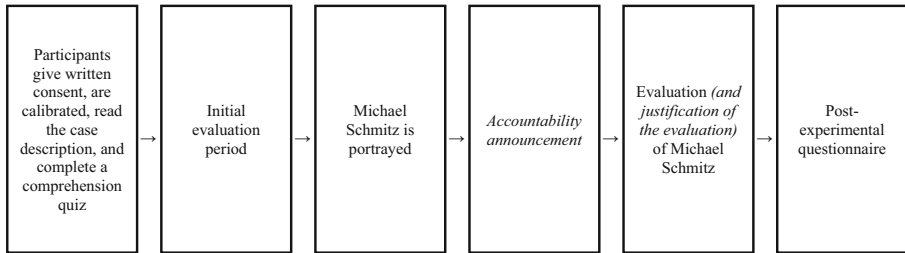


Fig. 2 Timeline of the Experiment. Notes: This figure provides a breakdown of the experimental steps in the order they were given to the participants. In italics are features of the experiment that were only present in the accountability high condition. The entire participation (including payout) took about 45 min

Participants in the *accountability low* condition are not asked to provide a written justification. Finally, participants review the performance report for Michael Schmitz containing the same type of performance measures as in the first step. Participants are asked to evaluate him on the same 101-point scale as in the first step. Participants in the *accountability high* conditions are additionally asked to write down their justification on the same screen (i.e., simultaneously to providing their evaluation). In Fig. 2 we present a timeline of the experiment.

3.4 Measurement of Variables and Use of Eye-Tracking

3.4.1 Dependent Variable

Our main dependent variable is *favoritism*. To measure favoritism, we rely on a specific pretest-posttest design (e.g., Asay and Hales 2018; Kramer et al. 2016; Christ et al. 2012). During the initial evaluation period, participants evaluate one subordinate (i.e., the sixth out of the twelve subordinates) with the same performance as Michael Schmitz, that is, everything else was kept *constant* other than the (absence of the) description of Michael Schmitz. We calculate favoritism as the difference between the two evaluations. Specifically, favoritism is calculated by subtracting this prior evaluation from participants' evaluation of Michael Schmitz. Thus, the measure takes positive values when participants exert favoritism.¹³

3.4.2 Eye-Tracking Data

In addition to our accountability manipulation described above, our second independent variable of interest is the *proportion of cognitive processing effort devoted*

(e.g., Fehrenbacher et al. 2018; Nelson and Rupa 2015; Clor-Proell and Maines 2014) and for brevity, we therefore drop the analysis of this variable and do not discuss it further.

¹³ Such a procedure, in addition to randomization, helps “remove any differences that could have been caused by factors other than the manipulations” (e.g., differences in leniency) (Christ et al. 2012, p. 442). We obtain qualitatively similar results if we instead use the evaluation of Michael Schmitz as dependent variable in our hypothesis test. That is, the interaction term used as our main hypothesis test is positive ($p = 0.06$, one-tailed) and highly significant when controlling for the evaluation of the previous employee with the same performance ($p < 0.01$, one-tailed).

to unfavorable information. This variable is measured instead of manipulated which can be warranted when the key research question focusses on an underlying cognitive process (Libby et al. 2002). Using such measured variables to increase our understanding of the mechanisms during subjective performance evaluations is also consistent with recent research (Fehrenbacher et al. 2018). Specifically, *proportion of cognitive processing effort devoted to unfavorable information* is measured using eye-tracking technology. As described in more detail below, it is the ratio of cognitive processing effort devoted to unfavorable information to the cognitive processing effort devoted to favorable information.

From a cognitive psychology perspective, cognitive processing describes a series of cognitive operations, including reasoning and synthesizing, whereby input is transformed to mental representations of information (Krcch 2018). Simply speaking, cognitive processes are operations whereby an individual deals with information. Literature has long been arguing that eye-movements reflect such cognitive processing (Rayner 1998). In this vein, the use of eye-tracking to measure cognitive processes has also been increasing in accounting recently (e.g., Kramer and Maas 2020; Dalla Via et al. 2019; Chen et al. 2016). To capture eye-movements, we use Tobii® Pro X2-60 eye-trackers (mounted to 23-inch monitors), which have a sampling rate of 60Hz. The sampling rate implies that participants' eye positions are recorded every 16.67 milliseconds. To analyze the eye-tracking data, we define so-called "areas of interest" (AOI). An AOI is a specific, definite area on the computer screen. Using AOIs allows to analyze whether a participant spent his or her attention on the information that lies within the AOI. Based on Kramer and Maas (2020), we define AOIs around the target and actual values and include the markers of each of the six performance measures that were presented in the performance reports when participants were evaluating Michael Schmitz. These AOIs are presented in Fig. 3.

In general, eye-tracking data can be separated into saccades and fixations. Saccades are rapid eye movements and usually last around 30 to 40 milliseconds. When the eye stands still for longer than that, research frequently entitles this a fixation (Holmqvist et al. 2011). According to the eye-mind hypothesis, when an individual fixates on information, the individual also cognitively processes it (Holmqvist et al. 2011; Just and Carpenter 1980). Hence, the number of fixations within specific AOIs is frequently used as a proxy for participants' cognitive processing of the information that is presented within the AOI. Moreover, our use of fixation counts as a proxy for cognitive processing is in line with Fehrenbacher et al. (2018), who also argue that such gaze data is the product of the eye-brain communication and can proxy for individuals' cognitive processing effort, referring to Horstmann, Ahlgrimm, and Glöckner (2009). We use Tobii Studio 3.2.1 software and adopt a fixation threshold of 60 milliseconds.¹⁴ This means that a fixation is recorded within an AOI whenever the eye fixates within the AOI for at least 60 milliseconds, which is a common minimum threshold for classifying fixations (e.g., Fehrenbacher et al. 2018; Holmqvist

¹⁴ Tobii Studio 3.2.1 is developed by the producers of the used eye-trackers and features mathematical algorithms to identify fixations. We used it to process the data, which allowed us to calculate specific variables (outlined below) for each participant from the raw data.

Performance measure	Target	Actual	+/-
Sales growth	15%	16%	+
Sales margins	60%	51%	-
Percentage of sales from new products	30%	32%	+
Inventory turnover	6,0	6,1	+
Customer satisfaction	85%	77%	-
Employee satisfaction	80%	82%	+

Fig. 3 Performance Report with Areas of Interest (AOIs). Notes: This figure shows the performance report for Michael Schmitz. The colored rectangles represent the AOIs used to measure participants' fixations on favorable and unfavorable performance information as captured by the eye-tracker. Participants did not see the AOIs on their screen. We use Tobii Studio 3.2.1 and adopt a fixation threshold of 60ms to calculate the variable *proportion of cognitive processing effort devoted to unfavorable information* as the sum of all fixations on AOIs that contain unfavorable information (blue color) divided by the sum of all fixations within AOIs that contain favorable information (red color). The higher the value on the measure, the relatively more cognitive processing participants devoted to the unfavorable information

et al. 2011). We extract both fixation counts and fixation duration for the AOIs, in line with Fehrenbacher et al. (2018).

As outlined above, our main variable of interest is *proportion of cognitive processing effort devoted to unfavorable information*. It is calculated as the sum of all fixations on AOIs that contain unfavorable information divided by the sum of all fixations within AOIs that contain favorable information.¹⁵ The higher the value on the measure, the relatively more cognitive processing effort participants devoted to the unfavorable information. As displayed in Fig. 3, the unfavorable performance measures for Michael Schmitz are 'sales margins' and 'customer satisfaction' whereas the remaining four performance measures are positive. Having more information which is consistent with participants' potential goal to provide a favorable performance information than inconsistent information is a design choice that is reflective of corporate practice, where targets are usually set rather readily achievable, thus potentially leading to little negative performance information in subordinates' performance reports (Merchant and Manzoni 1989).

Additionally, we use the eye-tracking data to calculate *cognitive processing effort*. Similar to Fehrenbacher et al. (2018), it is calculated as the sum of all fixations on both the favorable and the unfavorable performance information. The intuition behind this measure is that the more fixations on the overall performance information, the more cognitive processing effort participants exert on examining the information at all (i.e., the less superficial is their information processing).

¹⁵ In supplemental analyses, we show that our results are robust to using alternative ways of calculating *proportion of cognitive processing effort devoted to unfavorable information*.

4 Results

4.1 Preliminary Analyses and Descriptive Statistics

To confirm that our setting motivated participants to exert favoritism (i.e., they experienced positive affect towards Michael Schmitz), we asked participants during the post-experimental questionnaire to state their agreement (on a 7-point Likert scale anchored 1 = “Fully disagree” and 7 = “Fully agree”) towards the statement “I liked store manager Michael Schmitz.”. The data indicates that as intended, participants in both the accountability low condition (mean = 5.53, sd = 0.88) and the accountability high condition (mean = 5.34, sd = 1.47) liked Michael Schmitz (both values are significantly higher than the scale midpoint, p 's < 0.01, two-tailed).^{16,17} These means do not differ statistically significantly across conditions ($t = 0.85$, $p = 0.40$, two-tailed).

Table 1 reports descriptive statistics on our main dependent variable *favoritism* as well as *proportion of cognitive processing effort devoted to unfavorable information* and *cognitive processing effort*. Taking into account that favoritism is the difference between participants' evaluation of Michael Schmitz and their evaluation of the similar subordinate in the first step, we first note that the table shows that in the accountability low condition, favoritism significantly exceeds zero (mean = 3.52, sd = 11.93, $t = 2.25$, $p = 0.01$, one-tailed, $d = 0.29$), indicating that participants that were

Table 1 Descriptive Statistics

	Accountability		Overall ($n = 108$)
	Low ($n = 58$)	High ($n = 50$)	
Favoritism	3.52 (11.93)	1.00 (12.01)	2.35 (11.98)
Proportion of cognitive processing effort devoted to unfavorable information	0.76 (0.40)	0.70 (0.30)	0.73 (0.36)
Cognitive processing effort	47.98 (32.16)	102.00 (65.14)	72.99 (56.80)

Notes: The table presents descriptive statistics for the measured variables. We present mean values as well as standard deviations (in brackets)

Accountability is a between-subjects factor manipulated at two levels. In the accountability high condition, participants were asked to provide a written justification for their performance evaluation, whereas participants in the accountability low condition did not have to provide such a justification

Favoritism is our main dependent variable and is measured as the difference between participants' evaluation of Michael Schmitz and their evaluation of a similar performing subordinate during the initial evaluation period

Proportion of cognitive processing effort devoted to unfavorable information is calculated as the sum of all fixations on AOIs that contain unfavorable information divided by the sum of all fixations within AOIs that contain favorable information as captured by the eye-tracker

Cognitive processing effort is calculated as the sum of all fixations on both the favorable and the unfavorable performance information as captured by the eye-tracker

¹⁶ In this paper, p -levels are one-tailed for directional expectations and two-tailed otherwise.

¹⁷ We acknowledge that this is not a classical manipulation check. However, as exclusion of participants that score low on such a question might jeopardize results (Kotzian et al. 2020) and not excluding such participants biases against us finding support for our prediction, we deem our approach of not excluding any participants from our analyses conservative.

not held accountable exerted favoritism and inflated Michael Schmitz' performance evaluation.

Recall that *proportion of cognitive processing effort devoted to unfavorable information* is calculated as the sum of all fixations on AOIs that contain unfavorable information divided by the sum of all fixations within AOIs that contain favorable information. Thus, the higher the value on the measure, the relatively more cognitive processing effort participants devoted to the unfavorable information. The table shows that *proportion of cognitive processing effort devoted to unfavorable information* is not significantly higher *on average* (but even slightly lower) in the accountability high condition (mean = 0.70, sd = 0.30) than in the accountability low condition (mean = 0.76, sd = 0.40, $t = 0.75$, $p = 0.45$, two-tailed, $d = 0.15$). Finally, consistent with superficial information processing under low accountability, the table shows that *cognitive processing effort* is substantially lower under low accountability (mean = 47.98, sd = 32.16) than under high accountability (mean = 102.00, sd = 65.14, $t = -5.58$, $p < 0.01$, one-tailed, $d = -1.09$). In Appendix B, we further present a heatmap to additionally visualize the eye-tracking data.

4.2 Hypothesis Testing

In our hypothesis, we predict that the greater the proportion of cognitive processing effort managers devote to information that reflects unfavorable on a preferred subordinate, the lesser will be the mitigating effect of accountability on favoritism in their subjective performance evaluations. We test our hypothesis using a regression model. Specifically, we regress *favoritism* on an indicator variable *accountability* (equal to 1 when accountability is high and 0 when it is low), the continuous variable *proportion of cognitive processing effort devoted to unfavorable information* described above and the interaction of the two variables using robust standard errors. As the most conservative approach and in line with related research, we specified the model without any control variables. However, we present our hypothesis test including control variables further below. Data was analyzed using Stata 18.

As reported in Table 2 (Model 1), both the coefficient of accountability is negative ($\beta = -9.56$, $p = 0.02$, one-tailed) and the coefficient of the interaction term is positive and significant ($\beta = 10.11$, $p = 0.05$, one-tailed). Taken together, these results support our hypothesis as they indicate that the favoritism mitigating effect of accountability is attenuated the greater the proportion of cognitive processing effort devoted to unfavorable information. Figure 4 illustrates these findings. In the subsequent analyses, we further explore this interaction.

4.3 Supplemental Analyses

4.3.1 Further Exploration of the Interaction

To further explore the nature of the interaction, in Table 2, we first run an alternate model inspired by the literature (Maas and Verdoorn 2017) that controls for gender (coded 1 for female) and work experience (measured in months), with both being self-indicated by participants on the post-experimental questionnaire. As reported in

Table 2 Effects of Accountability and Proportion of Cognitive Processing Effort devoted to Unfavorable Information on *Favoritism*

Favoritism	Model 1	Model 2	Model 3
Accountability (1 = high)	−9.56 (4.63)	−8.67 (4.61)	−0.40 (4.63)
	0.02**	0.03**	0.02**
Proportion of cognitive processing effort devoted to unfavorable information	1.41 (3.30)	0.65 (3.27)	0.04 (3.30)
	0.67	0.84	0.67
Accountability × Proportion of cognitive processing effort devoted to unfavorable information	10.11 (6.03)	9.75 (5.80)	0.34 (6.03)
	0.05**	0.05**	0.05**
Gender		−4.25 (2.48)	
		0.09*	
Work experience		−0.01 (0.03)	
		0.81	
Constant	2.45 (2.78)	9.66 (5.55)	
	0.38	0.09*	
R ²	0.05	0.08	0.05
N	108	108	108

Notes: The table displays results of an OLS regression using heteroscedasticity-robust standard errors. We report beta coefficients for Model 1 and Model 2 and standardized beta coefficients for Model 3, standard errors in brackets, and *p*-levels. *, **, *** indicate significance at the 0.10, 0.05 and 0.01 levels, respectively. *P*-levels are one-tailed for directional expectations and two-tailed otherwise

Favoritism is our main dependent variable and is measured as the difference between participants' evaluation of Michael Schmitz and their evaluation of a similar performing subordinate during the initial evaluation period

Accountability is a between-subjects factor manipulated at two levels. In the accountability high condition, participants were asked to provide a written justification for their performance evaluation, whereas participants in the accountability low condition did not have to provide such a justification

Proportion of cognitive processing effort devoted to unfavorable information is calculated as the sum of all fixations on AOIs that contain unfavorable information divided by the sum of all fixations within AOIs that contain favorable information as captured by the eye-tracker

Gender (dummy variable coded 1 for Female) and *Work experience* (measured in months) represent control variables and are self-indicated by participants on the post-experimental questionnaire

Table 2 (Model 2), the coefficient of *accountability* ($\beta = -8.67$, $p = 0.03$, one-tailed) as well as the interaction term ($\beta = 9.75$, $p = 0.05$, one-tailed) remain largely unchanged, that is, they do not change in direction or significance. While work experience has no significant effect ($\beta = -0.01$, $p = 0.81$, two-tailed), gender has a negative effect ($\beta = -4.25$, $p = 0.09$, two-tailed) on favoritism. This indicates that female participants exerted less favoritism when evaluating Michael Schmitz and is consistent with psychology research that shows that when evaluating male subordinates, female

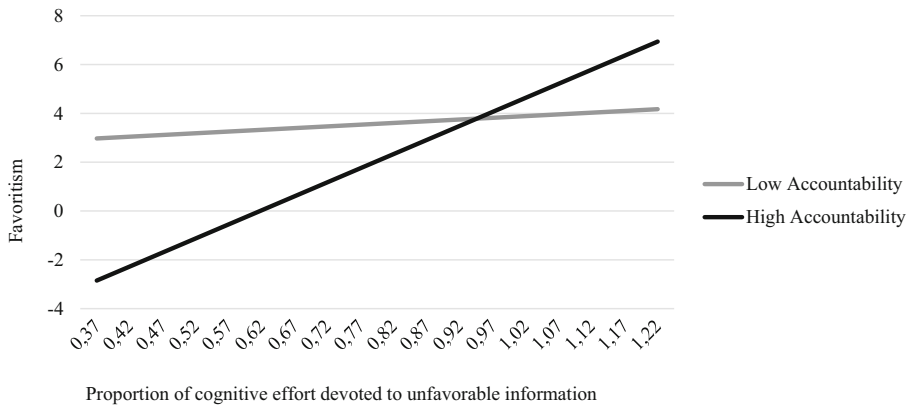


Fig. 4 Effects of Accountability and proportion of cognitive processing effort devoted to unfavorable information on *Favoritism*. Notes: The figure displays how proportion of cognitive processing effort devoted to unfavorable information affects favoritism under low accountability (grey line) and high accountability (black line). *Favoritism* is our main dependent variable and is measured as the difference between participants' evaluation of Michael Schmitz and their evaluation of a similar performing subordinate during the initial evaluation period. *Accountability* is a between-subjects factor manipulated at two levels. In the accountability high condition, participants were asked to provide a written justification for their performance evaluation, whereas participants in the accountability low condition did not have to provide such a justification. *Proportion of cognitive processing effort devoted to unfavorable information* is calculated as the sum of all fixations on AOIs that contain unfavorable information divided by the sum of all fixations within AOIs that contain favorable information as captured by the eye-tracker

managers tend to be harsher than male managers (Furnham and Stringfield 2001).¹⁸ In addition, within Table 2, we include Model 3, in which we present standardized beta coefficients for our regression used to test our hypothesis.

Finally, we examine whether the interaction we predict and observe is robust to the use of alternate ways to calculate our *proportion of cognitive processing effort devoted to unfavorable information* measure. First, we calculate it as the sum of all fixations on unfavorable information divided by the sum of all fixations on favorable and unfavorable information. Re-running the model reported in Table 2 with this measure (untabulated), we again find a negative coefficient for accountability ($\beta = -13.76$, $p = 0.03$, one-tailed) and a positive and marginally significant interaction term ($\beta = 28.41$, $p = 0.06$, one-tailed). Similarly, when we calculate our *proportion of cognitive processing effort devoted to unfavorable information* measure using the absolute time spent (i.e., the so-called total dwell time, cf. Lynch and Andiola 2019) on unfavorable information divided by the absolute time spent on both favorable and unfavorable information and re-run the model reported in Table 2 (untabulated), we again find that the coefficient of accountability is negative ($\beta = -13.29$, $p = 0.03$,

¹⁸ Whether this effect would alleviate when evaluating a preferred subordinate that is female (or female managers are less prone to favoritism in general) is beyond the scope of our study. Future research addressing this is very welcome.

one-tailed) and the coefficient of the interaction term is positive and marginally significant ($\beta = 27.38$, $p = 0.06$, one-tailed).¹⁹

4.3.2 Favoritism Under Low Accountability

In our theory development, we argue that when managers are held accountable, the proportion of their cognitive processing effort they devote to unfavorable information drives the efficacy of accountability in mitigating favoritism, but when managers are *not* held accountable, they rather not scrutinize information. Instead, they engage in superficial information processing and let their preference towards a subordinate guide their performance evaluation. As outlined above and consistent with our theory, we find that *cognitive processing effort* is substantially lower when accountability is low than when it is high ($p < 0.01$, one-tailed).²⁰ Next, we regress *favoritism* on *cognitive processing effort* and participants' response to the previously discussed post-experimental questionnaire statement as to whether they liked Michael Schmitz *for the subsample of participants in the accountability low condition*. We find that favoritism directionally increases the less *cognitive processing effort* participants exerted ($\beta = 0.06$, $p = 0.11$, one-tailed) and marginally significantly so the more they stated that they liked Michael Schmitz ($\beta = 2.78$, $p = 0.06$, one-tailed). Taken together, these results provide some support for our theory.

4.3.3 Managers' Concerns About Their Own Benefits

We also argue that an additional driver of the efficacy of accountability might be that managers that are held accountable may want to appeal to their superiors as being a good manager because this can entail monetary and non-monetary benefits for them. These perceptions about their potential own benefits when refraining from favoritism in turn likely work against their goal to do a preferred subordinate a favor. To corroborate this, we asked participants with regard to their performance evaluation of Michael Schmitz, to indicate, on the post-experiment questionnaire, their agreement to the statements "I think doing this activity could help me to be a good manager at TDC.", "I believe doing this activity could be beneficial to me.", and "I think this is important to do because it can make me a better manager at TDC." (on a Likert scale from 1 (fully disagree) to 7 (fully agree)). We calculate an index variable *perceived own benefit* as the average response to these three questions (Cronbach's alpha: 0.89). We regress *favoritism* on *proportion of cognitive processing effort devoted to unfavorable information*, *perceived own benefit* and the interaction of both variables for the subsample of participants in the accountability high condition. Consistent with our analyses reported beforehand, we find a positive coefficient for *proportion of cognitive processing effort devoted to unfavorable information* ($\beta = 31.20$, $p = 0.04$, one-tailed) and we also find a negative coefficient

¹⁹ We also find that our hypothesis test is robust to controlling for participants' overall *cognitive processing effort*.

²⁰ We again obtain similar results if we instead use absolute time spent on the performance information ($p < 0.01$, one-tailed).

for the interaction term ($\beta = -3.86$, $p < 0.10$, one-tailed). These results provide some support for our theory as they indicate that managers' perceptions about their potential own benefits work against the effect of their biased cognitive processing on their performance evaluations.

4.3.4 Managers' References to Non-Verifiable Information

Finally, however, we argue that when managers are held accountable and successfully self-justify a biased performance evaluation for a preferred subordinate, they likely disguise their favoritism by justifying their performance evaluation referring to alleged private, non-verifiable information. To corroborate this argument, we had an independent research assistant content-analyze participants' written justifications. The coder was instructed to classify whenever a participant justified his/her evaluation by not directly referring to Michael Schmitz' performance as captured by the performance measures. Participants did so by referring to Michael Schmitz' kindness, commitment, general productivity, determination, trust in him, future potential and praising him for his work. Using this data, we create a variable labeled *reference to non-verifiable information* that takes the value 1 whenever a participant made such references and 0 otherwise. Next, we regress *favoritism* on *proportion of cognitive processing effort devoted to unfavorable information*, *reference to non-verifiable information* and the interaction of both variables for the subsample of participants in the accountability high condition. Again, consistent with our analysis reported beforehand, we find a positive coefficient for *proportion of cognitive processing effort devoted to unfavorable information* ($\beta = 8.33$, $p < 0.01$, one-tailed) and a positive coefficient for the interaction term ($\beta = 14.52$, $p = 0.03$, one-tailed). These results support our theory as they indicate that engaging in motivated reasoning by referring to non-verifiable information even magnifies the effects of participants' biased cognitive processing on their favoritism.

5 Discussion

Many firms employ subjective performance evaluations as part of their control system. Yet research shows that subjectivity can enable managers to exert favoritism and lead to bias in performance evaluations (e.g., Ittner et al. 2003). While there is empirical evidence on the existence of favoritism, less is known about how managers use their discretion and how they deal with performance information when exerting favoritism as well as possible control mechanisms firms may implement to mitigate its occurrence. This paper presents the results of an eye-tracking experiment we conducted to investigate how accountability affects favoritism in subjective performance evaluation. We argue that accountability will cause managers to examine performance information more thoroughly. It follows that managers likely also have to scrutinize information which reflects unfavorable on preferred subordinates. This causes cognitive dissonance and managers can resolve this dissonance by engaging in motivated reasoning and cognitively discounting such information. The more managers reduce cognitive dissonance by discounting unfavorable information, the

less is the favoritism-mitigating effect of accountability. Hence, we predict and find that when a manager evaluates a preferred subordinate and is held accountable, the greater the proportion of cognitive processing effort devoted to performance information that reflects unfavorable on the subordinate, the lesser is the favoritism-mitigating effect of accountability.

5.1 Theoretical Implications

Favoritism is a topic that receives continuous scholarly attention (e.g., Dickens 2018; Bao and Huang 2024). Recent literature documents its continuing prevalence across many areas, including sports (e.g., Gift 2018; Audrino 2020), business (e.g., Robertson and O'Reilly 2024), healthcare (e.g., De los Santos et al. 2020), and the public sector (e.g., Szucs 2024). In addition to showing the prevalence of favoritism across these very diverse settings, this literature also links favoritism to various negative consequences, such as a negative effect on leader-member exchange (Bolino and Hsiung 2014) or a decline in individuals' willingness to cooperate with each other (Mollerstrom 2022). The literature is, however, largely silent on mechanisms to counteract favoritism. Our findings add to this literature, as we document—in a performance evaluation setting—that the higher the proportion of cognitive processing effort managers devote to information that reflects unfavorable on a preferred subordinate, the lower is the efficacy of accountability in mitigating favoritism. In the light of our theory, we interpret these findings as evidence on a motivated reasoning mechanism, by which managers discount information that reflects negatively on a preferred subordinate to self-justify exerting favoritism when held accountable. Taken together with the considerable lack of evidence on mechanisms that can successfully counteract favoritism, our findings imply that early claims in the literature that favoritism is hard to constrain might still be justified (Prendergast and Topel 1993). While it has been argued that accountability might be effective in reducing favoritism (Du et al. 2012), our cognitive processing approach using eye-tracking technology allows us to refine our understanding by showing that the efficacy of accountability crucially hinges on managers' cognitive processing of performance information.

Our findings also hold theoretic implications for the literature on accountability. As with favoritism, there is continuous scholarly interest in the effects of holding managers accountable in various settings (e.g., Donnelly and Donnelly 2023; Pan and Patel 2022; Aleksovska et al. 2019; Li et al. 2022; Pérez-Durán and Grimmelikhuijsen 2025) and this literature shows that it can entail both positive and negative effects. In addition, this literature discusses various organizational control mechanisms and design choices that can increase accountability. One example is transparency, which has received considerable attention recently (e.g., Hecht et al. 2023; Reynaers and Grimmelikhuijsen 2015). For example, the work of Bol et al. (2016) suggests that increasing transparency about managers' performance evaluation likely increases the extent to which managers feel accountable towards their subordinates. In contrast, a prominent example that can increase the extent to which managers feel accountable towards superiors in an organizational setting is having

managers formally justify their behavior towards a higher level, for example a direct superior (Hall et al. 2017).

In terms of holding managers accountable towards superiors that way, there is some specific literature that shows that some biases in subjective performance evaluations can be mitigated thereby. For example, some unconscious biases in performance evaluation that are the result of heuristics can be mitigated by accountability, such as common measure bias or halo bias caused by objectively measured performance dimension (e.g., Maske et al. 2021; Libby et al. 2004). However, there is also literature that suggests that accountability can amplify biases (Aleksavska et al. 2019; Lerner and Tetlock 1999; Bartlett et al. 2014). For example, Gordon et al. (1988) find in a hiring setting that accountability causes individuals' judgments to be more biased towards age stereotypes and work by Tetlock and Boettger (1989) shows that dilution bias is amplified by accountability. Our findings refine theory by introducing a boundary condition for the presumed debiasing effect of accountability (Du et al. 2012). Our results suggest that the extent to which holding managers accountable may lead to less bias may crucially depend on how managers allocate their cognitive processing effort, as they might engage in high cognitive processing effort to interpret information in a biased way consistent with their preferences. Thus, taking a cognitive processing perspective can help to understand the controversial findings of accountability documented in the literature (Donnelly and Donnelly 2023; Lerner and Tetlock 1999) and extends the literature on subjective performance evaluation (e.g., Rotaru et al. 2020; Fehrenbacher et al. 2018; Haas and Speckbacher 2017; Borchering and Grund 2003) by shedding more light on the role of managers' cognitive processes when exerting favoritism as a determinant for the efficacy of accountability as a debiasing instrument. In doing so, our study bridges psychological literature on motivated reasoning (e.g., Epley and Gilovich 2016) with the management control literature by highlighting that formal control mechanisms like accountability can have ambiguous effects depending on managers' cognitive processes.

Finally, we contribute to the eye-tracking literature. While originating from the field of psychology, eye-tracking is widely applied by researchers in management and accounting by now (e.g., Ursu et al. 2024; Gajewski et al. 2025; Zhang et al. 2024). Especially in accounting, the use of eye-tracking to investigate cognitive processes is growing in recent years and the so-obtained data is used to proxy various constructs (e.g., Sirois et al. 2018; Lynch and Andiola 2019; Panggabean et al. 2022; Ko et al. 2024). In line with recent calls from Lynch and Andiola (2019), we add to this small but growing stream of research by showing how eye-tracking can be used as a measure of cognitive processing effort and thus applied in a way that provides new and valuable insight to the effects of accountability on individuals' judgment and decision making beyond what could be inferred from a typical experiment.

5.2 Practical Implications

In addition to the scholarly discussion, our findings inform firms in practice that employ subjective performance evaluations on the possibilities and boundaries of

including accountability to their control systems to mitigate favoritism. Recently, Shi and Van Triest (2025) show that in performance evaluations of rank-and-file employees, managers place an average weight of 75% on subjective information. Their survey respondents come from a wide variety of industries, highlighting the relevance favoritism may play especially at such hierarchical levels across diverse managerial settings. While findings of Shi and Van Triest (2025) also suggest that accountability may lead superiors to rely less on subjective information, our findings suggest that when it comes to *preferred* employees, the efficacy of accountability is highly dependent on managers' cognitive processes.

Our findings may be particularly relevant in industries, organizations, and environments in which managers have significant discretion over the use of subjective information and there is limited transparency over objective employee performance or individual contributions—for example, in R&D teams or creative industries. In such settings, managers' personal preferences can more easily influence evaluations, and firms might attempt to use formal accountability structures (e.g., by explicitly requiring justification memos) to combat favoritism. Our results suggest that such attempts may be insufficient if managers self-justify biased evaluations through motivated reasoning. Therefore, firms might complement accountability mechanisms with additional control mechanisms. Our supplemental results provide some first indication: as these findings indicate that managers' perceptions about their potential own benefits work against their motivated reasoning attempts to exert favoritism, one approach could be to install calibration committees and link managers' own performance evaluations to the outcome of these committees and thus provide some stronger incentives to refrain from exerting favoritism (Grabner et al. 2020; Demeré et al. 2019).

5.3 Limitations and Future Research

Like all research, our study has limitations, which must be considered when interpreting its results while also providing opportunities for future research. First, we did not actively address evaluating managers' incentive structure in our experiment. Prior archival research has shown the prevalence of favoritism and related evaluation distortions although there likely are some incentives for managers to exert effort during evaluations and not to engage in such distortions in these settings (e.g., Moers 2005; Bol 2011; Du et al. 2012). We thus believe that while there are settings in practice that might somewhat motivate managers to exert more effort on their performance evaluations and generally mute the occurrence of favoritism (i.e., shift the *general* level of favoritism), the *directional* effects we predict and find in our experiment do generalize above our experimental setting, in which we are silent on evaluating managers' own incentives. Yet, it is eventually an interesting empirical question for future research how making managers' pay more dependent on their evaluations might affect the effectiveness of accountability to mitigate favoritism.

Second, similar to related research (e.g., Dalla Via et al. 2019), we test our research question using a student sample. The respective literature makes some arguments in favor of this approach, including the fact that professionals' specific experiences in performance evaluation might lead to more noise in the data (Libby

et al. 2004) and that results to theory based on general cognitive process (e.g., motivated reasoning) should not vary with experience (e.g., Libby et al. 2002). Thus, given that “there is no reason to believe a priori that motivational concerns [...] would be different for more experienced or knowledgeable subjects” (Libby et al. 2004, p. 1085) and that our sample fosters comparability with related research (e.g., Fehrenbacher et al. 2018), we deem our choice of participants appropriate for our research question. However, it might still be possible that experienced professionals assess the consequences of favoritism for a firm differently, and our student participants have limited work experience and thus also little experience in performance evaluation. While this should only lead to a different level of favoritism in general (i.e., represent a main effect) and not affect the effects of accountability on favoritism or the interaction with the focus on unfavorable information, which are at the heart of our study, this is eventually an empirical question. Given that eye-tracking technology in mobile devices becomes more and more convenient and affordable, we deem future research that investigates the robustness of our findings with a professional sample very valuable.

Finally, prior research shows that individuals’ preferences for fairness can affect their behavior in performance evaluation (e.g., Maas et al. 2012). In our post-experimental questionnaire, we did not capture participants’ individual fairness preferences as our study’s aim was not to test theory explicitly involving the effects of fairness. It seems, however, reasonable to expect that fairness preferences also affect favoritism. Moreover, as the literature implies that individuals have preferences for appearing fair and trustworthy to other individuals (e.g., Abdel-Rahim et al. 2022), including superiors, fairness preferences could also strengthen the effects of holding individuals accountable towards their superiors. Exploring fairness preferences thus appears to be a fruitful avenue for future research on favoritism.

Notwithstanding the above, our study provides new insight into how accountability affects favoritism in performance evaluations by uncovering the crucial role of managers’ cognitive processing of performance information. Using eye-tracking technology, we show that accountability alone does not eliminate favoritism, but its effectiveness depends on how managers allocate their cognitive processing effort. These findings enhance our understanding of the effects of accountability in subjective performance evaluation and suggest that cognitive processing patterns may moderate the effectiveness of formal control mechanisms.

6 Appendix

6.1 Appendix A

6.1.1 *Wording Used to Portray Michael Schmitz*

Recently you met Michael Schmitz again. Michael Schmitz is one of your store managers, whom you personally always liked very much. You have the feeling that you’re often on the same wavelength with him. Maybe that’s because you have many common interests and the same sense of humor and therefore always have nice

conversations with him when you meet him. You often talk about private matters and he always shows a lot of interest in your opinion. By now, you thus know a lot about Michael Schmitz's private life: For example, that he has two brothers and that his girlfriend works as a nurse. He himself trains the youth table tennis team in the community where he lives. Only when you last met him did he tell you that they are preparing once again for the upcoming regional championships. At the last work party of the TDC you sat next to Michael Schmitz and everyone at the table had a really pleasant and funny evening, which you still remember with pleasure. As always, you said goodbye to Michael very warmly.

6.2 Appendix B

Figure B. 1 displays a heatmap to additionally visualize the eye-tracking data; it shows the screen when participants interpreted the performance of the preferred subordinate and is generated from the fixation data from the accountability low condition. It can be seen that all areas containing information received a certain level of attention, while some areas received more attention than others. Specifically, as can be expected, it shows that participants allocated most of their attention to the numerical values and indicators of the various performance measures, which correspond to the AOIs we examine. Interestingly, the red areas already indicate that participants paid particular attention to the performance measures that reflected negative on the preferred subordinate.

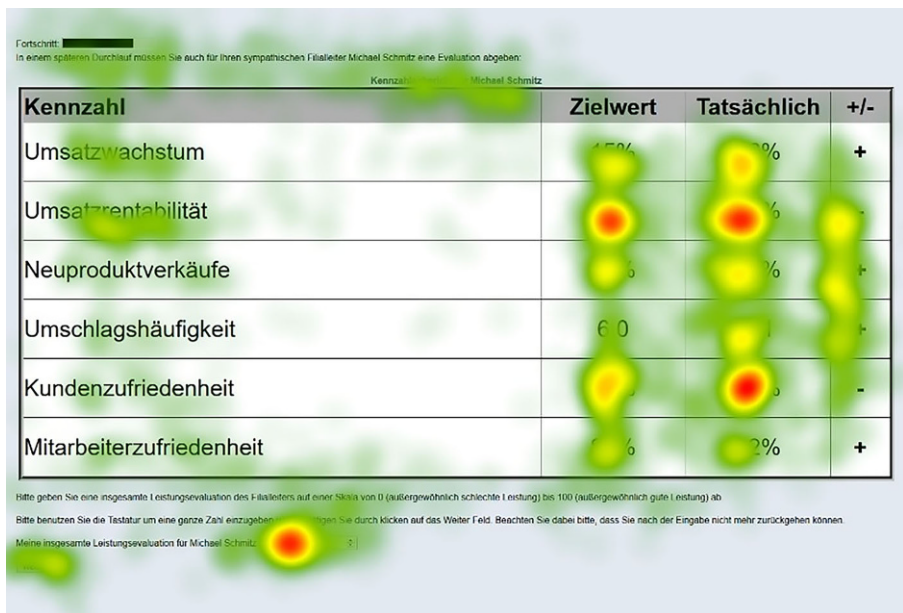


Fig. B.1 Heat Map Example

Acknowledgements We are grateful for the comments and suggestions of Markus Arnold, Florian Behrend, Jeremy Bentley, Christian Brück, Florian Elsinger, Chrisoph Hörner, Christian Huber (discussant), Kun Huo, Michael Majerczyk, Alice Muncy, Adam Presslee, Matt Sarji, Ivo Schedlinsky, Matt Sooy, Steve Sutton (discussant), Todd Thornock, Adam Vitalis, Dominique Wasna, Dimitri Yatsenko, participants in the Managerial Accounting Experimental Research Brown Bag, 2020 ACMAR conference, and the 2018 Nordic Accounting Conference Doctoral Colloquium.

Funding This study was funded by Heinrich Heine University Duesseldorf.

Availability of data The data and instrument are available from the authors upon request.

Declarations

Conflict of interest K.A. Bauch and B.E. Weißenberger declare that they have no competing interests.

Ethical standards We include a letter of Heinrich Heine University's dean (both authors were affiliated with that institution at the time the research was conducted) stating that there was no formal IRB and that we were not required to obtain any further approval for the experiment included in the paper. However, we confirm that we adhered to ethical standards. Participants gave their consent before participating and gave consent that the collected data will be published.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdel-Rahim, H.Y., J. Hales, and D.E. Stevens. 2022. How far will managers go to look like a good steward? An examination of preferences for trustworthiness and honesty in managerial reporting. *Contemporary Accounting Research* 39(2):1023–1053.
- Aleksovska, M., T. Schillemans, and S. Grimmellikhuijsen. 2019. Lessons from five decades of experimental and behavioral research on accountability: A systematic literature review. *Journal of Behavioral Public Administration* 2(2):1–18.
- Arnold, M.C., and I.D. Tafkov. 2019. Managerial discretion and task interdependence in teams. *Contemporary Accounting Research* 36(4):2467–2493.
- Arshad, F., E. Cardinaels, and B. Dierynck. 2018. *Facing a calibration committee: The impact on costly information collection and subjective performance evaluation*. working paper.
- Asay, H.S., and J. Hales. 2018. Disclaiming the future: Investigating the impact of cautionary disclaimers on investor judgments before and after experiencing economic loss. *The Accounting Review* 93(4):81–99.
- Audrino, F. 2020. Strongest team favoritism in European national football: Myth or reality? *International Journal of Sport Finance* 15(3):137–151.
- Bao, Z., and D. Huang. 2024. Gender-specific favoritism in science. *Journal of Economic Behavior & Organization* 221:94–109.
- Bartlett, G., E. Johnson, and P. Reckers. 2014. Accountability and role effects in balanced scorecard performance evaluations when strategy timeline is specified. *European Accounting Review* 23(1):143–165.
- Boiney, L.G., J. Kennedy, and P. Nye. 1997. Instrumental bias in motivated reasoning: More when more is needed. *Organizational Behavior and Human Decision Processes* 72(1):1–24.
- Bol, J.C. 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature* 27:1–24.

- Bol, J.C. 2011. The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review* 86(5):1549–1575.
- Bol, J.C., S. Kramer, and V.S. Maas. 2016. How control system design affects performance evaluation compression: The role of information accuracy and outcome transparency. *Accounting, Organizations and Society* 27(4–5):303–345.
- Bolino, M.C., and H.H. Hsiung. 2014. Playing favorites: How leader favoritism undermines leader-member exchange relationships. *Academy of Management Proceedings* 2014:10521.
- Borcherding, M., and C. Grund. 2003. Tariforientierung, Humankapitalinvestitionen und subjektive Leistungsbewertung – eine empirische Studie. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung* 55(5):482–509.
- Cardinaels, E., and P.M.G. van Veen-Dirks. 2010. Financial versus non-financial information: The impact of information organization and presentation in a balanced scorecard. *Accounting, Organizations and Society* 35(6):64–73.
- Carmona, S., G. Iyer, and P.M.J. Reckers. 2014. Performance evaluation bias: A comparative study on the role of financial fixation, similarity-to-self and likeability. *Advances in Accounting* 30(1):9–17.
- Chen, Y., J. Jermias, and T. Panggabean. 2016. The role of visual attention in the managerial judgment of balanced-scorecard performance evaluation: insights from using an eye-tracking device. *Journal of Accounting Research* 54(1):113–146.
- Christ, M.H., S.A. Emmett, S.L. Summers, and D.A. Wood. 2012. The effects of preventive and detective controls on employee performance and motivation. *Contemporary Accounting Research* 29(2):432–452.
- Clor-Proell, S.M., and L.A. Maines. 2014. The impact of recognition versus disclosure on financial information: A preparer's perspective. *Journal of Accounting Research* 52(3):671–701.
- Curley, S.P., J.F. Yates, and R.A. Abrams. 1986. Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes* 38(2):230–256.
- Dalla Via, N., M. van Rinsum, and P. Perego. 2019. How accountability type influences information search processes and decision quality. *Accounting, Organizations and Society* 75:79–91.
- De los Santos, J.A.A., R.A. Rosales, C.C. Falguera, C.N. Firmo, K. Tsaras, and L.J. Labrague. 2020. Impact of organizational silence and favoritism on nurse's work outcomes and psychological well-being. *Nursing Forum* 55(4):782–792.
- Deméré, B.W., K.L. Sedatole, and A. Woods. 2019. The role of calibration committees in subjective performance evaluation systems. *Management Science* 65(4):1562–1585.
- Dickens, A. 2018. Ethnolinguistic favoritism in African politics. *American Economic Journal: Applied Economics* 10(3):370–402.
- Ding, S., and P. Beaulieu. 2011. The role of financial incentives in balanced scorecard-based performance evaluations: Correcting mood congruency biases. *Journal of Accounting Research* 49(5):1223–1247.
- Ditto, P.H., and D.F. Lopez. 1992. Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology* 63(4):568–584.
- Ditto, P.H., J.A. Scepansky, G.D. Munro, A.M. Apanovitch, and L.K. Lockhart. 1998. Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology* 75(1):53–69.
- Donnelly, A.M., and D.P. Donnelly. 2023. A systematic review of experimental research on accountability in auditing. *Behavioral Research in Accounting* 35(2):37–72.
- Du, F., G. Tang, and S.M. Young. 2012. Influence activities and favoritism in subjective performance evaluation: Evidence from Chinese state-owned enterprises. *The Accounting Review* 87(5):1555–1588.
- Epley, N., and T. Gilovich. 2016. The mechanics of motivated reasoning. *Journal of Economic Perspectives* 30(3):133–140.
- Fehrenbacher, D.D., A.K.-D. Schulz, and K. Rotaru. 2018. The moderating role of decision mode in subjective performance evaluation. *Management Accounting Research* 41(1):1–10.
- Fehrenbacher, D.D., S.E. Kaplan, and C. Moulang. 2020. The role of accountability in reducing the impact of affect on capital budgeting decisions. *Management Accounting Research* 47:100650.
- Feldman, J. 1981. Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology* 66(2):127–148.
- Festinger, L. 1957. *A theory of cognitive dissonance*, 2nd edn., Stanford: Stanford University Press.
- Fischer, P., S. Schulz-Hardt, and D. Frey. 2008. Selective exposure and information quantity: How different information quantities moderate decision makers' preference for consistent and inconsistent information. *Journal of Personality and Social Psychology* 94(2):231–244.
- Furnham, A., and P. Stringfield. 2001. Gender differences in rating reports: Female managers are harsher raters, particularly of males. *Journal of Managerial Psychology* 16(4):281–288.

- Gajewski, J.F., M. Heimann, P.M. Léger, and P. Teye. 2025. Enhancing auditors' professional skepticism through nudges: An eye-tracking experiment. *Accounting and Business Research* 55(4):426–444.
- Gift, P. 2018. Performance evaluation and favoritism: Evidence from mixed martial arts. *Journal of Sports Economics* 19(8):1147–1173.
- Gordon, R.A., R.M. Rozelle, and J.C. Baxter. 1988. The effect of applicant age, job level, and accountability on the evaluation of job applicants. *Organizational Behavior and Human Decision Processes* 41(1):20–33.
- Grabner, I., J. Künneke, and F. Moers. 2020. How calibration committees can mitigate performance evaluation bias: An analysis of implicit incentives. *The Accounting Review* 95(6):213–233.
- Haas, N., and G. Speckbacher. 2017. Everything under my control: CEO characteristics and the evaluation of middle manager performance in small and medium-sized firms. *Schmalenbach Business Review* 18(2):109–128.
- Hales, J., X. Kuang, and S. Venkataraman. 2011. Who believes the hype? An experimental examination of how language affects investor judgments. *Journal of Accounting Research* 49(1):223–255.
- Hall, A.T., D.D. Frink, and M.R. Buckley. 2017. An accountability account: A review and synthesis of the theoretical and empirical research on felt accountability. *Journal of Organizational Behavior* 38(2):204–224.
- Hecht, G., V.S. Maas, and M. Van Rinsum. 2023. The effects of transparency and group incentives on managers' strategic promotion behavior. *The Accounting Review* 98(7):239–260.
- Heckman, L. 2022. State Dept. IG raises 'favoritism' concerns within Foreign Service promotion boards. <https://federalnewsnetwork.com/hiring-retention/2022/05/state-dept-ig-raises-favoritism-concerns-within-foreign-service-promotion-boards/>.
- Hermans, T., M. Cools, and A. Van den Abbeele. 2021. The role of information accuracy and justification in bonus allocations. *Journal of Management Control* 32(2):197–223.
- Holmqvist, K., M. Nyström, R. Andersson, R. Dewhurst, J. Halszka, and J. van de Weijer. 2011. *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Horstmann, N., A. Ahlgrimm, and A. Glöckner. 2009. How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes. *Judgment and Decision making* 4(5):335–354.
- Ittner, C.D., D.F. Larcker, and M.W. Meyer. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review* 78(3):725–758.
- Just, M.A., and P.A. Carpenter. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review* 87(4):329–354.
- Kadous, K., J. Kennedy, and M.E. Peecher. 2003. The effect of quality assessment and directional goal commitment on auditors' acceptance of client-preferred accounting methods. *The Accounting Review* 78(3):759–778.
- Kahn, V. 2022. Survey says: Appreciation matters more than you think. <https://bonusly.com/post/employee-appreciation-survey>.
- Kaplan, S.E., M.J. Petersen, and J.A. Samuels. 2007. Effects of subordinate likeability and balanced scorecard format on performance-related judgments. *Advances in Accounting* 23(1):85–111.
- Kennedy, J. 1993. Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research* 31(2):231–245.
- Kennedy, J. 1995. Debiasing the curse of knowledge in audit judgment. *The Accounting Review* 70(2):249–273.
- Ko, J.C.W., M.M. Cheng, and W.J. Green. 2024. Stakeholder focus or strategy focus? An eye-tracker study on the effect of presentation format on nonprofessional investors' information processing patterns. *European Accounting Review* 33(4):1521–1548.
- Kotzian, P., T. Stöber, F. Hoos, and B.E. Weißenberger. 2020. To be or not to be in the sample? On using manipulation checks in experimental accounting research. *Accounting Research Journal* 33(3):469–482.
- Kramer, S., and V.S. Maas. 2020. Selective attention as a determinant of escalation bias in subjective performance evaluation judgments. *Behavioral Research in Accounting* 32(1):87–100.
- Kramer, S., V.S. Maas, and M. van Rinsum. 2016. Relative performance information, rank ordering and employee performance: A research note. *Management Accounting Research* 33:16–24.
- Krch, D. 2018. Cognitive processing. In *Encyclopedia of clinical neuropsychology*, 2nd edn., ed. J.S. Kreutzer, J. DeLuca, and B. Caplan, 859.
- Kruglanski, A.W., and T. Freund. 1983. The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology* 19(5):448–468.
- Kunda, Z. 1990. The case for motivated reasoning. *Psychological Bulletin* 108(3):480–498.

- Lerner, J.S., and P.E. Tetlock. 1999. Accounting for the effects of accountability. *Psychological Bulletin* 125(2):255.
- Li, Y., W. Liu, and G. Yu. 2022. Dancing in shackles: The double-edged sword effect of felt accountability on work outcomes and individual wellbeing. *Frontiers in Psychology* 13:904946.
- Libby, R., R. Bloomfield, and M.W. Nelson. 2002. Experimental research in financial accounting. *Accounting, Organizations and Society* 27(8):775–810.
- Libby, T., S.E. Salterio, and A. Webb. 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review* 79(4):1075–1094.
- Lipe, M.G., and S.E. Salterio. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures? *The Accounting Review* 75(3):283–298.
- Lynch, E.J., and L.M. Andiola. 2019. If eyes are the window to our soul, what role does eye-tracking play in accounting research? *Behavioral Research in Accounting* 31(2):107–133.
- Maas, V.S., and N. Verdoorn. 2017. The effects of performance report layout on managers' subjective evaluation judgments. *Accounting and Business Research* 47(7):731–751.
- Maas, V.S., M. Van Rinsum, and K.L. Towry. 2012. In search of informed discretion: An experimental investigation of fairness and trust reciprocity. *The Accounting Review* 87(2):617–644.
- Maske, M.K., M. Sohn, and B. Hirsch. 2021. How managerial accountability mitigates a halo effect in managers' ex-post bonus adjustments. *Management Accounting Research* 51:100738.
- Merchant, K.A., and J.-F. Manzoni. 1989. The achievability of budget targets in profit centers: A field study. *The Accounting Review* 64(3):539–558.
- Mero, N.P., and S.J. Motowidlo. 1995. Effects of rater accountability on the accuracy and the favorability of performance ratings. *Journal of Applied Psychology* 80(4):517–524.
- Mero, N.P., S.J. Motowidlo, and A.L. Anna. 2003. Effects of accountability on rating behavior and rater accuracy. *Journal of Applied Social Psychology* 33(12):2493–2514.
- Milkovich, G.T., and A.K. Wigdor. 1991. *Pay for performance: Evaluating performance appraisal and merit pay*. Washington: National Academy Press.
- Moers, F. 2005. Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations and Society* 30(1):67–80.
- Mollerstrom, J. 2022. Favoritism and cooperation. *Public Choice* 191(3):293–307.
- Nelson, M.W., and K.K. Rupar. 2015. Numerical formats within risk disclosures and the moderating effect of investors' concerns about management discretion. *The Accounting Review* 90(3):1149–1168.
- Palmer, J.K., and J.M. Feldman. 2005. Accountability and need for cognition effects on contrast, halo, and accuracy in performance ratings. *The Journal of Psychology* 136(2):119–137.
- Pan, P., and C. Patel. 2022. Formal accountability, perceived accountability and aggressive reporting judgements. *Accounting and Business Research* 52(1):67–93.
- Panggabean, T., Y. Chen, and J. Jermias. 2022. Do you see what i see? The effects of dissenting opinion on information acquisition and decision quality: An eye-tracking study. *Advances in Accounting Behavioral Research* 25:249–277.
- Peecher, M.E., M.D. Piercey, J.S. Rich, and R.M. Tubbs. 2010. The effects of a supervisor's active intervention in subordinates' judgments, directional goals, and perceived technical knowledge advantage on audit team judgments. *The Accounting Review* 85(5):1763–1786.
- Pérez-Durán, I., and S. Grimmelikhuijsen. 2025. Watching care behind closed doors: Upward and downward accountability mechanisms in nursing home services. *Public Management Review* 27(5):1469–1488.
- Prendergast, C., and R.H. Topel. 1993. Discretion and bias in performance evaluation. *European Economic Review* 37(2–3):355–365.
- Rayner, K. 1998. Eye movements in reading and information processing: 20 years of research? *Psychological Bulletin* 124(3):372–422.
- Reynaers, A.M., and S. Grimmelikhuijsen. 2015. Transparency in public-private partnerships: Not so bad after all? *Public Administration* 93(3):609–626.
- Robertson, K.M.P., and J. O'Reilly. 2024. Playing the favorite game: A contextual examination of workplace favoritism. *Academy of Management Proceedings* 2024:17386.
- Rotaru, K., D.D. Fehrenbacher, M.H. Liang, and A.K.-D. Schulz. 2020. Causal inference in judgment using the balanced scorecard. *Journal of Management Accounting Research* 32(2):201–224.
- Schlenker, B.R. 1986. *Personal accountability: Challenges and impediments in the quest for excellence*. San Diego: Naval Personnel Research and Development Center.
- Shi, B., and S. Van Triest. 2025. What determines managers' use of subjective performance information? *Accounting and Business Research* . forthcoming.

- Sirois, L.P., J. Bédard, and P. Bera. 2018. The informational value of key audit matters in the auditor's report: Evidence from an eye-tracking study. *Accounting Horizons* 32(2):141–162.
- Szucs, F. 2024. Discretion and favoritism in public procurement. *Journal of the European Economic Association* 22(1):117–160.
- Tetlock, P.E. 1983. Accountability and the complexity of thought. *Journal of Personality and Social Psychology* 45(1):74–83.
- Tetlock, P.E., and R. Boettger. 1989. Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology* 57(3):388–398.
- Tetlock, P.E., L. Skitka, and R. Boettger. 1989. Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology* 57(4):632–640.
- Ursu, R.M., T. Erdem, Q. Wang, and Q. Zhang. 2024. Prior information and consumer search: Evidence from eye tracking. *Management Science* 70(12):8685–8708.
- Voußem, L., S. Kramer, and U. Schäffer. 2016. Fairness perceptions of annual bonus payments. The effects of subjective performance measures and the achievement of bonus targets. *Management Accounting Research* 30:32–46.
- Xu, Y., and B.M. Tuttle. 2005. The role of social influences in using accounting performance information to evaluate subordinates: A causal attribution approach. *Behavioral Research in Accounting* 17(1):191–210.
- Zhang, Z., X. Jia, R.L. Greenbaum, and N. Zaman. 2024. Time is hunting: Investigating time pressure and bottom-line mentality with an eye-tracking approach. *Academy of Management Proceedings* 2024:13162.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.