THE ACCOUNTING REVIEW
Vol. 95, No. 4
July 2020
pp. 51–72

American Accounting Association
DOI: 10.2308/accr-52610

The Unintended Consequences of Material Weakness Reporting on Auditors' Acceptance of Aggressive Client Reporting

Tim D. Bauer
University of Waterloo

Anthony C. Bucaro
Case Western Reserve University

Cassandra Estep
Emory University

ABSTRACT: Regulators are concerned that auditors do not sufficiently identify and report material weaknesses in internal control over financial reporting (ICFR). However, psychological licensing theory suggests reporting material weaknesses could have unintended consequences for acceptance of aggressive client financial reporting. In an experiment, we predict and find auditors accept more aggressive client reporting after they report a material weakness in ICFR than after they report no material weakness. We provide evidence licensing underlies this effect. In a second experiment, we investigate the efficacy of an intervention to reduce the identified licensing effects by prompting an audit quality goal. We find this prompt mitigates the unintended consequence when auditors report a material weakness. While regulators are concerned companies are undeservedly receiving clean ICFR audit opinions, our findings indicate adverse ICFR opinions may lead auditors to give companies undeservedly clean financial statement opinions. We provide a potential remedy to this unintended consequence.

Keywords: internal control over financial reporting; audit quality; material weaknesses; psychological licensing.

I. INTRODUCTION

Since the passage of Sarbanes-Oxley Section 404 (U.S. House of Representatives 2002), public companies that meet specific thresholds require an annual integrated audit; that is, an audit of internal control over financial reporting (ICFR) in conjunction with the audit of the financial statements. Effective ICFR provides a base level of assurance as to the reliability of management's financial reporting. Material weaknesses in ICFR suggest to financial statement users that ICFR is ineffective and that the related financial statements have an increased risk of being materially misstated, even if no specific misstatements exist (Public Company Accounting Oversight Board [PCAOB] 2007). Audit standards suggest that an auditor, upon identifying a material weakness to be reported in the auditor's ICFR opinion, should reevaluate audit risk and the audit plan (i.e., increase risk assessments and planned testing) in an effort to reduce the chance that a material misstatement goes undetected and uncorrected (PCAOB 2007, 2010). Thus, even if a material weakness leads to a material misstatement, the audited financial statements are not necessarily materially misstated because auditors should detect and have management

This paper has been improved by helpful comments from Elaine G. Mauldin (editor) and two anonymous reviewers, as well as Lindsay Andiola, Jeremy Bentley, Ken Bills, Emily Griffith, Jackie Hammersley, Pat Hurley, Kathryn Kadous, Yoon Ju Kang, Mark Peecher, Mike Ricci, Ken Trotman, Michael Williamson, Dan Zhou, and seminar participants at the 2016 Accounting Behavior and Organizations (ABO) Research Conference, the Goizueta Business School Experimental Brown Bag at Emory University, the University of Massachusetts Amherst, the 2017 Auditing Section Midyear Meeting, the Case Western Reserve University Brown Bag, and the 2019 Emory Accounting Research Conference.

We greatly appreciate the support of the CAQ and AAA Auditing Section Access to Auditors program and the time and effort of the audit firm participants. The views expressed in this article and its content are those of the authors alone and not those of the Center for Audit Quality.

Editor's note: Accepted by Elaine G. Mauldin, under the Senior Editorship of Mark L. DeFond.

Submitted: March 2017
Accepted: September 2019
Published Online: October 2019

correct the misstatement (Hogan and Wilkins 2008). Results of prior research, however, raise questions about the extent to which auditors successfully do so.

Hogan and Wilkins (2008) find that audit fees increase in the presence of material weaknesses and suggest that this is evidence auditors adjust their plans and increase testing in an attempt to detect potential misstatements. Other research, though, finds that audit plans, in practice, show limited adjustments for risk (Bedard 1989; Mock and Wright 1993, 1999). Further, Doyle, Ge, and McVay (2007) and Ashbaugh-Skaife, Collins, Kinney, and LaFond (2008) find that firms with (versus without) material weaknesses have lower financial reporting quality, suggesting that auditors may not sufficiently detect or require correction of misstatements resulting from weak ICFR (Hogan and Wilkins 2008; Joe, A. Wright, and S. Wright 2011). It begs the question why, in the presence of material weaknesses, is audit testing higher, but financial reporting quality lower?

One possibility is that while auditors increase testing in response to material weaknesses in ICFR, their changes are largely ineffective at *detecting* the corresponding misstatements (Hammersley, Johnstone, and Kadous 2011). We rely on psychological licensing theory (Monin and Miller 2001; Cain, Loewenstein, and Moore 2005, 2011) and two experiments with experienced auditors to investigate another possibility: that auditors who report material weaknesses can also fail to require management to *correct* subsequently detected misstatements. Specifically, we argue that reporting of material weaknesses identified during ICFR testing can inadvertently give auditors license to allow *more* aggressive, and potentially materially incorrect, client reporting. This potential unintended consequence is concerning given that regulators desire more reporting of material weaknesses in ICFR (e.g., Croteau 2013; Franzel 2014).

Whether auditors find and decide to report a material weakness in ICFR or not, they should be equally likely to have a client correct any aggressive reporting detected. However, psychological licensing theory suggests that once an individual dutifully performs an initial action, they may (even unknowingly) grant themselves a mental release or license to be less dutiful in a subsequent action (Miller and Effron 2010). Deciding to report a material weakness in ICFR may make auditors feel a subconscious release because they performed dutifully, since this action effectively warns investors to be wary of the risk of material misstatement to the financial statements (PCAOB 2007). This release could lead them to subsequently perceive an aggressive client-preferred accounting adjustment as legitimate and, thus, acceptable. In our study, we predict that auditors accept more aggressive client reporting when reporting a material weakness in ICFR than when reporting no such weakness.

We also investigate a potential remedy to mitigate this unintended consequence. Prior research shows that when goals are made salient, individuals are more likely to make a series of consistent decisions (Fishbach, Dhar, and Zhang 2006). Prompting auditors with an audit quality goal should lead to more consistency between ICFR and financial reporting decisions—that is, challenging more aggressive reporting (a dutiful act) when first reporting a material weakness in ICFR (a dutiful act)—and, thus, mitigate the licensing effects. We predict that prompting an audit quality goal mitigates licensing effects on auditors' acceptance of aggressive reporting.

In Experiment One, we manipulate the type of material weakness reported by the audit team for a hypothetical integrated audit in a 1×4 between-subjects design. Audit senior participants learn that they and their team have decided to either report a material weakness (we use three types of material weakness for robustness) or no material weakness. After receiving general engagement information and ICFR conclusion information, participants receive details about an inventory obsolescence issue subsequently detected by the audit team. They respond to questions that we use to evaluate their willingness to accept management's aggressive reporting of an ambiguous accounting estimate. In Experiment Two, we employ a 2×3 between-subjects design. Audit manager (and above) participants perform the same task as in Experiment One. We again manipulate whether a material weakness is reported. We also manipulate the presence versus absence of an audit quality goal prompt, and include two versions of the prompt.

In Experiment One, we find that auditors in material weakness conditions accept *more* aggressive reporting than those in the no material weakness condition. We replicate this effect in Experiment Two, in the conditions without the prompt, and provide evidence that psychological licensing is the underlying process in both experiments. In Experiment Two, we find that auditors who complete an audit quality goal prompt (versus not) accept less aggressive reporting when reporting a material weakness in ICFR. However, we find that auditors who complete the prompt (versus not) may accept more aggressive reporting when *not* reporting a material weakness. Thus, actively prompting this goal can mitigate the effects of psychological licensing, but careful consideration should be given to the situations in which it is effective to do so. We discuss such considerations in more detail below.

Our study has important implications. First, auditors have been required to attest to the effectiveness of certain public clients' ICFR since 2004. Empirical research on implications of ICFR evaluation and reporting is limited, despite calls from research and practice to examine auditor effectiveness in ICFR audit tasks (Asare et al. 2013; Center for Audit Quality [CAQ] 2018) and concerns from regulators that auditors have difficulty integrating ICFR and financial statement audits (PCAOB 2009; Bhaskar, Schroeder, and Shepardson 2019). We extend prior research that suggests that auditors do not always effectively modify substantive testing in response to ICFR deficiencies (Hammersley et al. 2011; Mauldin and Wolfe 2014) by showing that material weaknesses in ICFR affect *outcomes* of substantive testing, such as what corrective adjustments to require. We

also extend prior work that shows that auditors often fail to require such corrections of detected material misstatements (Church, Davis, and McCracken 2008) by showing evidence of psychological licensing as a potential explanation for this behavior.

Second, the Securities and Exchange Commission (SEC) and PCAOB have criticized auditors for reporting too few material weaknesses in ICFR (Croteau 2013; Franzel 2014). While evidence supports the concern that companies are undeservedly receiving clean ICFR audit opinions (Rice and Weber 2012; Asare et al. 2013), our findings suggest that regulators should carefully consider how to achieve improved ICFR reporting without sacrificing financial reporting quality. We do not advise regulators to stop encouraging the reporting of material weaknesses in ICFR, as material weaknesses provide valuable information to capital markets on financial reporting quality. Rather, interventions should be considered to avoid unintended consequences, such as the licensing effects we find.

Interventions like the prompt we examine in this study also require caution before implementing in practice. Prompting and making salient an audit quality goal may not be effective if implemented universally. For example, our prompted auditors received the prompt before encountering any audit engagement details. As predicted, the prompt helped mitigate licensing for auditors who initially encountered and reported a material weakness. They acted consistently, and in accordance with the audit quality goal, by subsequently limiting aggressive client reporting. But prompted auditors who initially encountered and reported *no* material weakness may *also* have acted consistently across their decisions. Initially, these auditors did not identify audit concerns because they reported no material weakness and, subsequently, they did not identify audit concerns because they allowed more aggressive reporting than unprompted auditors. In this case, consistency was not necessarily in accordance with the audit quality goal. Providing the prompt only *after* deciding to report an identified material weakness could, thus, be a way to benefit from the prompt without the undesirable effect that occurs when no material weakness is present. Future research can explore whether prompting an audit quality goal produces similar results when provided before or after the material weakness decision.

Finally, we contribute to accounting and psychology research on psychological licensing. Prior work in accounting (Griffin 2014) has shown effects that may be consistent with licensing, but our study is the first to demonstrate an underlying licensing process and investigate how to mitigate it. Prior work in psychology has also paid little attention to mitigators of licensing. Our two experiments show that licensing effects and the underlying process are robust across auditors of different ranks and material weakness types, even when the weaknesses have no direct effect on the potentially misstated financial statement account. Thus, our results suggest that auditors could exhibit licensing effects in a wide variety of settings—while we examine licensing in the integrated audit setting, its effects could occur across other areas of the audit. We encourage future research to examine other non-ICFR contexts that could lead auditors to feel dutiful and licensed and to consequently act less dutifully in the face of aggressive client reporting.

II. THEORY AND HYPOTHESIS DEVELOPMENT

Background: Integrated Audits of ICFR and Financial Statements

Within integrated audits of financial statements, an identified deficiency in ICFR during internal control testing signals the potential for increased risk of material misstatement (PCAOB 2007). Consistent with this, prior research has predicted and found that financial reporting quality is lower for companies with versus without material weaknesses in ICFR (Doyle et al. 2007; Ashbaugh-Skaife et al. 2008). However, while material weaknesses in ICFR may increase the risk of material misstatement, auditors should respond by increasing testing. Further, by increasing testing, the chance of a material misstatement remaining after the audit is performed should be low regardless of the level of risk before the audit is performed (PCAOB 2007, 2010). In essence, if auditors properly follow the Audit Risk Model, then financial reporting quality should not vary between companies with and without material weaknesses in ICFR.[1]

Auditors struggle to integrate ICFR results into their subsequent audit decisions (PCAOB 2009; Bhaskar et al. 2019). Prior research shows that auditors modify planned substantive testing in response to material weaknesses identified during ICFR testing, but their changes are ineffective at *detecting* the underlying misstatements (Hammersley et al. 2011; Mauldin and Wolfe 2014). We extend this research by exploring the possibility that auditors who report material weaknesses do not require management to *correct* misstatements the auditors effectively detect.

Prior research does provide evidence that when auditors detect material misstatements, they often fail to require (i.e., they waive) corrective adjustments (e.g., A. Wright and S. Wright 1997; Joe et al. 2011; for a review, see Church et al. 2008). Much

[1] Paragraph 0.10 of Auditing Standard (AS) 1015 (PCAOB 2002) states that exercising due professional care allows auditors "to obtain reasonable assurance about whether the financial statements are free of material misstatement . . . Absolute assurance is not attainable because of the nature of audit evidence and the characteristics of fraud." Therefore, financial reporting quality may vary, even while following the Audit Risk Model, although it is unlikely to vary materially.

of this research suggests that auditors primarily waive material audit adjustments that are subjective (Wright and Wright 1997; Braun 2001; Church et al. 2008) or when auditors lack independence (Lennox, Wu, and Zhang 2014; Messier and Schmidt 2018). However, despite the finding that auditors are more likely to discuss audit adjustments with management when control risk is higher (Cannon and Bedard 2017), studies fail to find the expected positive association between internal control strength and the decision to waive versus correct misstatements (Joe et al. 2011; Cannon and Bedard 2017). We extend this research by examining material weaknesses in ICFR, which represent a severe form of control risk, and by proposing another reason for auditors' failure to correct misstatements: subconscious effects of psychological licensing.

Unintended Consequences of Reporting Material Weaknesses in ICFR

In an effort to improve audit quality, regulators have communicated that auditors often fail to identify and report material weaknesses in ICFR (Croteau 2013; PCAOB 2013). However, psychological licensing theory suggests that more diligent reporting of material weaknesses in ICFR could, in fact, reduce audit quality as it relates to the substantive audit of financial statements. Licensing research finds that individuals who display moral attitudes or behavior on an initial task are more likely to display moral lapses on a subsequent task, without concern of feeling or looking immoral (Monin and Miller 2001; Sachdeva, Iliev, and Medin 2009; Effron and Monin 2010).[2] In other words, past good deeds give license to individuals to subsequently act inconsistently and do bad, because the former acts excuse the latter act or reconstrue the latter act as done with legitimate good intentions (Merritt, Effron, and Monin 2010). Analogously, once auditors dutifully perform an initial action—reporting a material weakness in ICFR—they may (even unknowingly) grant themselves a mental release or license to be less dutiful in a subsequent action—detecting and requiring management to correct aggressive reporting.

Licensing has been demonstrated with students or the general public in various generic settings, including political correctness (e.g., Monin and Miller 2001; Effron, Cameron, and Monin 2009), prosocial behavior (e.g., Sachdeva et al. 2009; Mazar and Zhong 2010), and conflicts of interest (e.g., Cain et al. 2005, 2011). For example, individuals are less likely to suggest that a woman is qualified for a stereotypically male job or to donate to a charity, when first given the opportunity (versus not) to select a woman for a gender-neutral job or to describe themselves with positive words, respectively (Monin and Miller 2001; Sachdeva et al. 2009).

Licensing studies involving conflict of interest most closely capture features of the audit setting, as individuals are often placed in a verification/advisor role (Cain et al. 2005, 2011; Koch and Schmidt 2010; Jamal, Marshall, and Tan 2016). Results of these studies show that when advisors first disclose their conflict of interest (a dutiful act), they subsequently give advice in their own best interest, regardless of accuracy (a less dutiful act); advisors feel that they possess a license or excuse to be less accurate since users of their advice have been warned. Advisors rationalize that they earned the right to give personally beneficial advice even if it is inaccurate, or that their advice is not actually inaccurate because they typically act dutifully, particularly when what constitutes "accurate" is ambiguous (Cain et al. 2011; Effron and Monin 2010).[3]

However, licensing effects may not generalize to the audit setting because auditors, unlike students who participated in most of the prior studies, are required to perform tasks while adhering to a strong professional and moral code. Auditors could be less prone to psychological effects that manifest in more general settings when they make decisions in context-rich settings that reflect the audit environment (Smith and Kida 1991).[4] Two studies, Jamal et al. (2016) and Griffin (2014), provide important insight in professional settings.

Jamal et al. (2016) use professional valuation specialists as participants, and they find results consistent with psychological licensing; participants' estimates are more biased toward their conflicted interest when they disclose such conflict to users of their opinion. But like other conflict of interest studies (e.g., Cain et al. 2005; Church and Kuang 2009; Koch and Schmidt 2010), participants in Jamal et al. (2016) received direct financial compensation for misleading users.[5] In our experimental setting, as in the natural audit setting, neither reporting a material weakness nor opining on the financial statements *directly*

[2] Moral license, self-license, and psychological license are used relatively interchangeably in the literature, with moral license applying to issues of morality, the most common setting for research in this area. Feeling licensed does not mean that an individual has to act a certain way, it just grants them the ability to do so without fear of discrediting oneself (Miller and Effron 2010).

[3] Licensing can occur when individuals (1) believe their actions balance out; i.e., they know a subsequent action is inappropriate, but believe it is excusable given the *credit* earned from a prior appropriate action, or (2) perceive (or believe others will perceive) that their subsequent action is done for appropriate reasons; i.e., their prior action provides *credentials* that they are someone who acts appropriately (Merritt et al. 2010; Effron and Monin 2010). The two mechanisms are "two independent routes to licensing . . . [that] may often operate simultaneously to produce moral licensing" (Miller and Effron 2010, 128). In the audit context, the less conscious credentials mechanism is more likely to occur, as auditors are unlikely to consciously bias behavior (Kadous, Kennedy, and Peecher 2003).

[4] Consistent with the notion that licensing effects may be less likely for participants with more real-world experience, Cain et al. (2011) fail to find licensing effects in one study with college graduates as participants.

[5] As such, results could be due to licensing or strategic exaggeration, whereby advisors bias in their favor to offset expected advisee discounting (Cain et al. 2011); Jamal et al. (2016) focus on strategic exaggeration, not licensing.

links to auditor monetary incentives; moreover, auditor reporting of a material weakness does not represent a disclosure of conflict of interest. Thus, if bad deeds follow good deeds in our setting, then it will more likely result from subtle changes in cognition related to licensing than from explicit auditor self-interest.

Griffin (2014) uses professional auditors as participants and, similar to our study, does not include direct monetary incentives. Griffin (2014) finds that auditors are less likely to require their client to adjust uncertain fair value estimates when the client's footnote disclosure contains additional details about the uncertainty. As in reporting a material weakness, such disclosure gives financial statement users a warning about the potential for inaccuracy (i.e., misstatements) to exist. This can give auditors a license to permit management to report more aggressively. But as noted by Griffin (2014, 1184), adjusting an estimate and disclosing its details are explicitly related actions and they *substitute* for each other. In our setting, ICFR can directly or indirectly impact financial statements, but a material weakness opinion does not necessarily contain a direct discussion of financial statement accounts or provide financial information that users could substitute for a reported financial statement amount.

Thus, results in Griffin (2014) could have arisen from licensing, although he notes that he does not specifically test for licensing. Alternatively, the results could have arisen from auditors feeling that users had more information when disclosure was present. The additional details described in the footnote about the estimate can permit users to assess the uncertainty of the estimate and adjust it themselves. Regardless, we extend his work and test licensing more broadly by examining a setting where the initial and subsequent action are less directly connected. While a material weakness and the area of potential material misstatement that auditors need to evaluate could overlap directly, they can also overlap indirectly or not at all.[6] Importantly, psychological licensing theory suggests that licensing can occur regardless of whether two decisions are seen as related or unrelated, provided that the subsequent decision is conceived as part of a sequence of decisions that define one's morals/duty (Klotz and Bolino 2013; Mullen and Monin 2016).[7]

A recent study by Luippold, Kida, Piercey, and Smith (2015) suggests that auditors may not be affected by licensing when making decisions in unrelated areas. When auditors in their study are diverted by management to an audit area with (versus without) errors, they are more likely to identify an unusual fluctuation in another audit area as earnings management. In our setting, this would suggest that when auditors first encounter an issue in ICFR (versus not), they will be *more* likely to identify a financial reporting issue as aggressive. However, a key difference between our setting and that of Luippold et al. (2015) could lead to different auditor behavior, consistent with licensing. In Luippold et al. (2015), participants *identify* a potential error in the financial statements, but do not *disclose* that potential error to the public. Signaling to potential observers a good deed (e.g., through disclosure) is an important step in licensing theory for individuals to feel a release that licenses them to later act inappropriately (Miller and Effron 2010). Thus, licensing effects are more likely to be observed in our setting than that of Luippold et al. (2015) because auditors who decide to report a material weakness (versus not) can more sufficiently signal a good deed and/or warn financial statement users of the potential for misstatement.

In sum, we expect that when auditors intend to report a material weakness in ICFR, it will subconsciously grant them license to less dutifully perform their professional duty to constrain aggressive, ambiguous reporting by client management (when performing subsequent substantive procedures). Accordingly, we predict the following:

H1: Auditors will exhibit psychological licensing such that they accept more aggressive client reporting when concurrently reporting a weakness in ICFR than when reporting no such weakness.

Mitigating the Unintended Consequence: Prompting an Audit Quality Goal

Auditors who exhibit licensing effects as described above are acting in an undesirable manner and acting inconsistent with required audit standards. It would, therefore, be useful to identify mechanisms to reduce the unintended licensing consequences. The crux of the issue is that auditors will perceive that past good deeds excuse or reconstrue (as legitimate) subsequent transgressions. Breaking or reframing this mental sequencing, such that past good behavior is less likely to justify future bad behavior, could be fruitful. There is limited evidence on such moderators of psychological licensing effects (Merritt et al. 2010; Mullen and Monin 2016), but see Conway and Peetz (2012) and Susewind and Hoelzl (2014) for notable exceptions.

[6] As discussed in Section III, we examine a potential material misstatement that arises in one area of the audit, while the material weakness arises in a directly related, indirectly related, and unrelated area.

[7] Auditors may perceive ICFR and financial reporting decisions to be unrelated due to structural features in how ICFR and financial statements are audited. For example, audits are performed in teams, and the individual auditor who decides whether to correct a client's financial reporting may have little to no involvement in deciding to report a material weakness. Kouchaki (2011), however, provides evidence that licensing effects can still occur even if an actor's group (e.g., other members of the auditor's team) performs the initial action.

Research on self-regulation suggests that thinking about one's commitment to a goal can trigger consistency between initial and subsequent actions. Fishbach and Dhar (2005) show, for example, that student participants prompted to focus on their commitment to (versus progress on) academic goals were less likely to shirk their studies and socialize with friends, after first studying hard. Said differently, they were less likely to exhibit licensing behavior. Fishbach et al. (2006) and Susewind and Hoelzl (2014) find similar results in various settings, although they find that thinking about commitment to a goal can be a double-edged sword. Individuals prompted to focus on commitment to a goal follow initial goal-congruent behavior with subsequent goal-congruent behavior, but also follow initial goal-incongruent behavior with more goal-incongruent behavior. That is, actions congruent with the goal reinforce commitment to the goal (and lead people to persist on the goal); however, actions incongruent with the goal reduce commitment to the goal (and lead people to give up on the goal) (Fishbach et al. 2006).

Prior accounting research has examined both goal commitment and prompts. Kadous et al. (2003), for example, find that auditors with higher versus lower commitment to a directional goal of accepting client-preferred accounting methods are more likely to judge the client's method as appropriate. Recent studies have found that a wide variety of prompts are successful at inducing certain auditor behavior (e.g., Hoffman and Zimbelman 2009; Griffith, Hammersley, Kadous, and Young 2015; Rasso 2015; Backof, Carpenter, and Thayer 2018; Kadous and Zhou 2019). Most related to our study, Ricci (2020) finds that auditors who are prompted to focus on commitment to a goal (of client satisfaction) are more likely to act in a goal-congruent manner (agree with the client) when they perceive that they have performed higher versus lower on this goal already.

Overall, research in different domains suggests that focusing individuals on their commitment to goals, or prompting goals more generally, increases the likelihood that they will make a series of consistent decisions. This suggests that a prompt containing a goal is likely to mitigate licensing effects that otherwise generate decisions inconsistent with each other. In the context of our study, prompting auditors with an audit quality goal should mitigate licensing effects; they should follow initial goal-congruent behavior (reporting a material weakness) with further goal-congruent behavior (limiting aggressive client reporting). Whether the benefits observed in prior psychology and accounting studies will occur in our integrated audit setting is an empirical question because regulators indicate that auditors have difficulty integrating decisions related to ICFR and financial reporting (PCAOB 2009). Auditors' difficulties may be particularly persistent or resistant to mechanisms meant to mitigate issues arising in the integrated audit context. However, supported by prior psychology research, we expect that prompting a goal when licensing effects are likely to occur—that is, when auditors report a material weakness—will mitigate such effects. We predict the following:

H2: Prompting an audit quality goal will mitigate the effects of psychological licensing on auditors' acceptance of aggressive client reporting.

III. EXPERIMENT ONE

Method: Participants and Design

To test H1, we conduct an experiment where we adapt an ICFR issue from Hammersley et al. (2011) and an inventory obsolescence issue from Bauer (2015). Participants in Experiment One are 89 audit seniors from two Big 4 audit firms.[8] Their audit experience ranges from four to 96 months, with a mean experience of 36 months.[9] Seniors are appropriate for the experimental task because they routinely perform both internal control and substantive testing, and provide input to their superiors on both material weakness (or internal control deficiency) reporting and resolving potentially aggressive financial reporting. Additional demographic data reported in Table 1 suggests that participants have sufficient experience with ICFR and inventory to meaningfully complete the experiment. We randomly assign participants to one of four between-subjects conditions, where we manipulate the material weakness reported as: none (*No MW*), in the same account as the aggressive financial reporting (*Same Account MW*), in a different account than the aggressive reporting (*Different Account MW*), or at the entity level (*Entity-Level MW*). We expect H1 to hold regardless of material weakness type, but since Doyle et al. (2007) find variation in the strength of the relation between reported material weaknesses and lower financial reporting quality across material weakness types, we include three types of material weaknesses for robustness. Refer to Figure 1, Panel A for a graphical depiction of H1 across conditions.

[8] We obtained approval from the Institutional Review Board (IRB) of the university where the experiment took place and all participants consented to participate, per IRB requirements. Participants from the first (second) firm completed the experimental task at a national training program (sessions held on the same day in three different offices). There are no significant main or interactive effects of firm; thus, we combine observations across firms in all analyses. We received 93 completed instruments from auditors meeting our criteria of seniors who audit public companies with ICFR. We eliminate data from four auditors who omitted responses to variables used in our model.

[9] Participants' audit experience does not vary significantly across conditions ($F_{85,3} = 0.83$, $p = 0.48$, two-tailed).

TABLE 1

Both Experiments
Participant Demographics: Mean, (Median), [Standard Error]

Task-Specific Experience[a]	Experiment One Seniors n = 89	Experiment Two Managers and Above n = 118
Material Weakness (MW) Experience		
Number of audits encountering an MW in ICFR	0.49 (0.00) [0.08]	1.72 (1.00) [0.32]
ICFR-Related Experience		
Experience with controls related to ICFR, relative to peers	5.67 (6.00) [0.31]	6.98 (7.35) [0.23]
Similar Industry Experience[b]		
Percent of time auditing clients in industry similar to case	29.6% (15.0%) [0.04]	34.4% (30.0%) [0.03]
Inventory Experience		
Experience with inventory/obsolescence, relative to peers	4.41 (5.00) [0.32]	5.90 (6.90) [0.26]

[a] No measure of experience varies significantly across conditions; as such, we do not report cell means or statistics.
[b] In Experiment Two, only 107 participants provided a response to this question.
Measures of experience include: experience with material weaknesses (MW) in ICFR (number of audit engagements where an MW was encountered); experience concluding on internal controls related to the ICFR opinion (self-assessed relative to peers, on a scale from "no experience" [0] to "extensive experience" [10]); experience with clients in a similar industry to the hypothetical client (self-reported percentage of time auditing similar clients); and experience dealing with inventory or inventory obsolescence issues (self-assessed relative to peers on a scale from "no experience" [0] to "extensive experience" [10]).

Method: Procedures and Variables

Participants are asked to assume that they are the audit senior on the audit of a furniture manufacturing client. After receiving general engagement information, participants are provided details about testing of ICFR that contains the material weakness manipulations. Those in the *No MW* condition are informed that they did not identify any internal control deficiencies that require a material weakness classification, the partner agrees with this conclusion, and their audit team has informed the client that no material weakness will be reported in the audit opinion on ICFR. In the other three conditions, participants are told that they identified one material weakness in inventory valuation (*Same Account MW*), derivatives valuation (*Different Account MW*), or the financial statement close process (*Entity-Level MW*). Again, in all cases, the partner agrees with the conclusion and the audit team has informed the client that the particular material weakness will be reported in the audit opinion on ICFR. To reinforce the manipulation, all participants complete a question asking them how involved they were in the ICFR conclusion.[10] While, implicitly, the risk of material misstatement is higher in the three MW conditions, if auditors merely respond to perceived risk differences, they will be less likely to allow aggressive reporting for a riskier client, which biases against finding results consistent with H1.

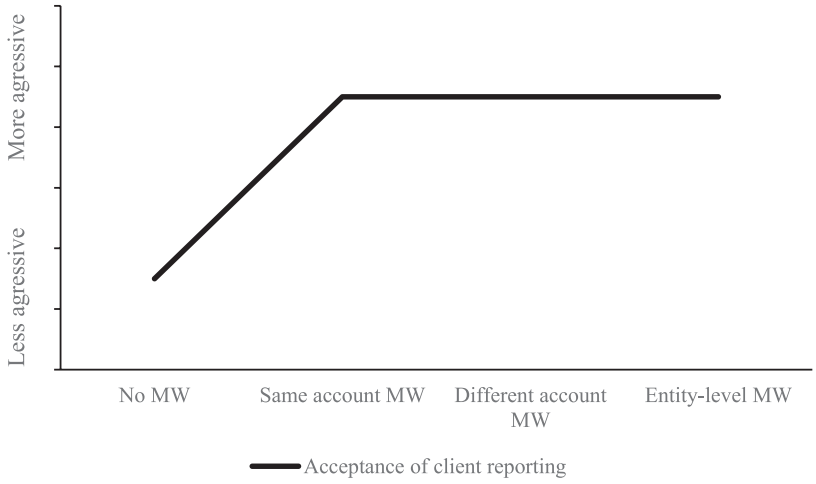
Participants then receive details about an inventory obsolescence issue subsequently discovered by the audit team.[11] Participants are told that the client has inventory that may require a write-down of value due to expiring sales contracts and specifications that do not fit those of its new retailers. Working papers provided to participants indicate that the audit team

[10] The mean of participant responses is 7.91 on an 11-point scale from "definitely not involved" (0) to "definitely involved" (10), and is significantly higher than the scale midpoint ($t_{88} = 15.3$, $p < 0.01$, two-tailed). Responses do not differ across conditions ($F_{85,3} = 0.62$, $p = 0.60$, two-tailed). As noted earlier, individuals need not be directly involved in the initial action for licensing effects to occur (Kouchaki 2011); however, such a requirement would bias against finding results consistent with H1.

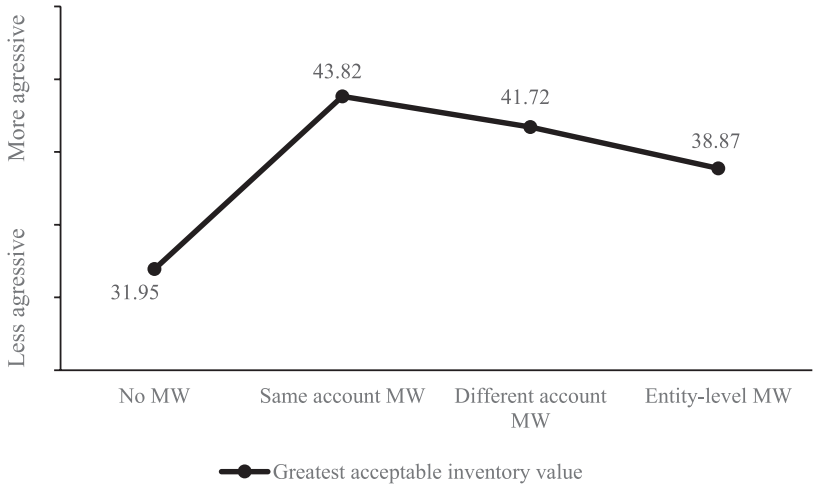
[11] Material weaknesses may be identified at any point in the audit and remain part of the dialog until final reporting is complete. Our design takes advantage of that by having our participants consider both ICFR and financial statement conclusions as they would naturally, while evaluating the results of the overall audit.

FIGURE 1
Experiment One
H1 Graphs

Panel A: H1 Prediction



Panel B: H1 Results



Refer to Table 2 for independent and dependent variable descriptions.

estimates a write-down of up to $75.6 million could be supported based on unrecoverable value, but the client's CFO believes a write-down of $39.8 million is reasonable because the president will likely approve investments to renew sales contracts or transform old inventory. Audit materiality is $2 million, so the difference between the audit team's estimate and management's estimate is clearly material. Participants assess the lowest amount of inventory write-down that they would be willing to accept being recorded, between zero and the total value of old inventory ($100.8 million), where lower assessments indicate greater agreement with the client's proposed write-down (Bauer 2015). We subtracted this assessment from the total inventory value ($100.8M) to determine the greatest inventory value they would accept.[12] This serves as our dependent variable. A higher inventory value indicates acceptance of more aggressive client reporting.

[12] This adjustment aligns the direction of the dependent variable with our construct of interest, higher acceptance of management aggressive reporting. Statistical results are identical when using the assessed lowest acceptable inventory write-down as the dependent variable.

Participants then complete a series of additional questions, including process and demographic measures. For licensing effects to occur, auditors should view their ICFR reporting decision as more dutifully performed when they report a material weakness versus not. As such, we ask participants the extent to which they agree that their ICFR decision (1) shows to outsiders that they are objective in evaluating financial statements, (2) forewarns investors about the risk of material misstatement, and (3) has value to investors, with each rated on an 11-point scale from "strongly disagree" (−5) to "strongly agree" (+5). If licensing occurs, then when auditors report a material weakness in ICFR and feel they acted dutifully, they should be more likely to perceive the client-proposed write-down as legitimate. Thus, we ask participants to assess, on an 11-point scale, the following regarding the client's proposed write-down: (1) its appropriateness, ranging from "not at all appropriate" (0) to "completely appropriate" (+10); (2) its aggressiveness, ranging from "very conservative" (−5) to "very aggressive" (+5); (3) their likelihood of requesting a different write-down, ranging from "not at all likely" (0) to "very likely" (+10); and (4) their agreement that the financial statements will be materially misstated if they allow the proposed write-down, ranging from "no agreement" (0) to "complete agreement" (+10).

Results: Tests of H1

We predict that auditors will accept more aggressive client reporting when concurrently reporting a material weakness in ICFR than when reporting no such weakness (H1). To test H1, we include one covariate in a 1×4 analysis of covariance (ANCOVA), with the auditor's greatest acceptable inventory value (in $millions) as the dependent variable. The covariate is self-assessed perceptions of experience with ICFR relative to peers, rated on an 11-point scale from "no experience" (0) to "extensive experience" (+10). It is not significantly correlated with our independent factor ($F_{85,3} = 0.31$, $p > 0.81$, two-tailed).[13] Figure 1, Panel B graphically depicts results of the 1×4 ANCOVA.

Results are consistent with H1. Table 2, Panels A–D show results for the greatest acceptable inventory value. The planned contrast with weights of $\{-3, +1, +1, +1\}$ corresponding to *No MW, Same Account MW, Different Account MW,* and *Entity-level MW,* respectively, is significant ($p = 0.02$, one-tailed). Further, the insignificant residual between-cells variance ($p = 0.67$, two-tailed) and contrast variance residual of 16 percent indicate that the hypothesized contrast describes the data well. The mean assessment is significantly lower in *No MW* ($M = 32.0$) compared to the three MW conditions combined ($M = 41.3$), and compared to *Same Account MW* ($M = 43.8$, $p = 0.02$, one-tailed) and *Different Account MW* ($M = 41.7$, $p = 0.05$, one-tailed), but not compared to *Entity-Level MW* ($M = 38.9$, $p = 0.11$, one-tailed). Overall, these results largely support H1; auditors are willing to accept a higher inventory value (i.e., allow more aggressive client reporting) when reporting a weakness in ICFR than when reporting no such weakness.

Supplemental Analysis: Evidence of Psychological Licensing Process

We next examine the cognitive process through which H1 unfolds. We expect that auditors will first feel licensed, by viewing their ICFR reporting decision as a good deed, to a greater extent when they report a material weakness versus not. This "dutiful" feeling will then lead to greater perceptions of legitimacy of the client-proposed write-down; that is, in feeling dutiful, auditors will more strongly perceive that the client's proposal, and allowing it, is legitimate. Our measure of perceptions of acting dutifully is a factor of three assessments described earlier—whether participants felt their ICFR conclusion shows objectivity to, forewarns, or provides value to investors. Our measure of legitimacy is a factor of another four assessments described earlier: participants' ratings of the appropriateness and aggressiveness of the write-down, likelihood of requesting a different write-down amount, and agreement that the proposed write-down would result in a material misstatement (the latter three measures are reverse-coded). See Appendix A for means, standard errors, and other supporting details on factor scores.

To test for licensing effects predicted to underlie H1, we performed bias-corrected and accelerated bootstrapping for serial mediation, as suggested in Preacher and Hayes (2008, model 6), separately analyzing each material weakness type against no material weakness. See Figure 2 for a graphical depiction of this model. Participants' judgment of whether they acted dutifully is included as the first of two mediating variables. Next, participants' judgment of the legitimacy of management's proposed write-down is included as the second mediating variable. Our analyses strongly support psychological licensing for *Same* and *Different Account MW*.

For *Same Account MW, Different Account MW,* and *Entity-level MW,* we find the 90 percent confidence interval for the conditional indirect effect of the path depicted in Figure 2 is (0.01, 5.08), (0.04, 3.30), and (−0.04, 6.21), respectively. Since this interval does not contain zero for the first two, those tests provide evidence that the effect of account-level material

[13] Results hold when excluding the covariate from the model.

TABLE 2

Experiment One
Tests of H1

Panel A: Greatest Acceptable Inventory Value

Condition	n	Greatest Acceptable Inventory Value	
		Adj. Mean	Std. Error
No MW	22	31.95	(4.05)
Any Material Weakness	67	41.27	(2.31)
Same Account MW	23	43.82	(3.95)
Different Account MW	21	41.72	(4.15)
Entity-Level MW	23	38.87	(3.95)

Panel B: ANCOVA

Source	SS	df	MS	F	p-value[c]
Material Weakness	1777.28	3	592.43	1.65	0.184
ICFR-Related Experience	2010.85	1	2010.85	5.60	0.020
Error	30157.42	84	359.02		

Panel C: Planned Tests

Contrast and Residual Between Cells Variance	SS	df	MS	F	p-value
No MW lower than other three MW conditions [Contrast weights: −3, 1, 1, 1]	1492.79	1	1492.79	4.16	0.022[b]
Residual between-cells variance	284.49	2	142.25	0.40	0.672[c]
Total between-cells variance	1777.28	3	592.43		
Contrast Variance Residual[a]	16.01%				

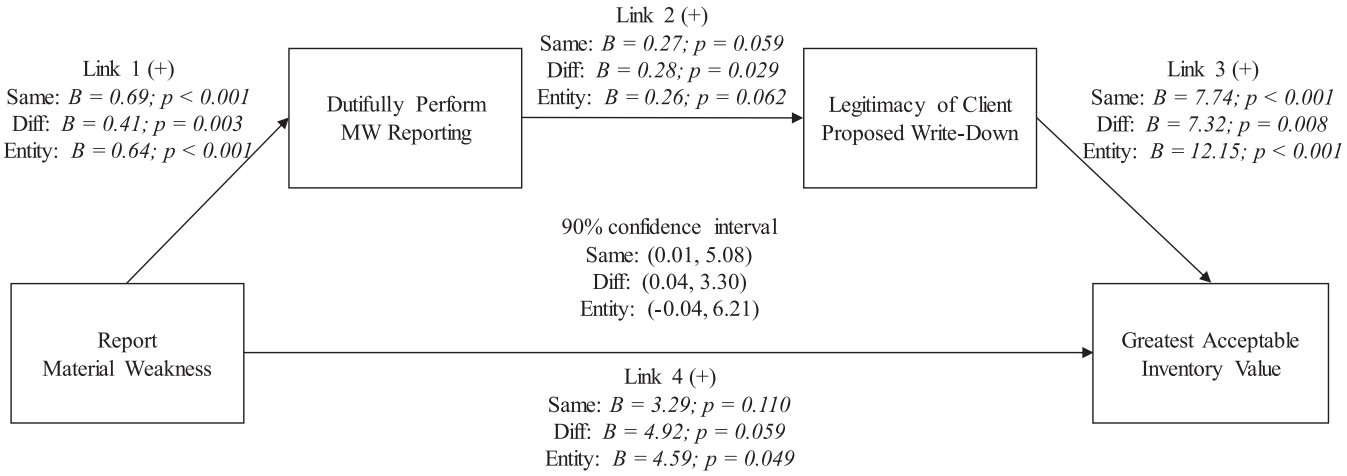
Panel D: Planned Tests, continued from Panel C

Simple Effects Tests	t	p-value
No MW versus *Same Account MW*	−2.10	0.020[b]
No MW versus *Different Account MW*	−1.68	0.048[b]
No MW versus *Entity-Level MW*	−1.22	0.113[b]
Same Account MW versus *Different Account MW*	0.37	0.714[c]
Same Account MW versus *Entity-Level MW*	0.89	0.378[c]
Different Account MW versus *Entity-Level MW*	0.50	0.620[c]

[a] Following Guggenmos, Piercey, and Agoglia (2018), we calculated the contrast variance residual as the residual between-cells variance sum of squares divided by total between-cells variance sum of squares. Results indicate that approximately 16 percent of the between-cells variance in our model is explained by factors other than the contrast.
[b] Directional prediction; reported p-values are one-tailed.
[c] No directional prediction; reported p-values are two-tailed.
Participants assessed (in $M) the lowest inventory write-down they would accept, and we subtracted this assessed value from the total inventory value ($100.8M) to determine the greatest inventory value they would accept. This assessment is covariate-adjusted for experience concluding on internal controls related to the ICFR opinion (self-assessed relative to peers on a scale from "no experience" [0] to "extensive experience" [10]). The four between-subjects conditions are: no material weakness (*No MW*), material weakness in inventory (*Same Account MW*), material weakness in derivatives valuation (*Different Account MW*), and material weakness in the financial statement close process (*Entity-Level MW*).

weaknesses (versus *No MW*) on judgments of the greatest acceptable inventory value is mediated through assessments of dutiful performance and the legitimacy of the write-down. We also find significant or marginally significant effects (in the direction expected) for each of Links 1–3 in the path in Figure 2. Participants perceive themselves as performing more dutifully in any MW condition versus *No MW* (Link 1: all $p < 0.01$, one-tailed); higher perceptions of performing dutifully result in

FIGURE 2
Experiment One
Psychological Licensing Process for No MW versus MW Conditions



This figure captures the psychological licensing process for when a material weakness (*MW*) is reported (in one of three areas) versus not at all, using the Preacher and Hayes (2008) method for testing indirect effects. We examine the links in the model for each of three MWs reported: in the same account as the aggressive client reporting (*Same*), in a different account than the aggressive client reporting (*Diff*), or at the entity level (*Entity*). The model includes the same covariate as the model in Table 2, and the mediators (*Dutifulness* and *Legitimacy*) are measured as defined in Appendix A. Each individual link of Links 1 through 3 is significant or marginally significant in the expected direction (all p-values are one-tailed). We expect that reporting a material weakness, compared to reporting no material weakness, will increase the perception that the auditor performed dutifully (Link 1); these higher perceptions will, in turn, increase the auditor's assessment that the client proposed write-down is legitimate (Link 2); and these higher assessments will, in turn, increase the auditor's assessment of the greatest acceptable inventory value (Link 3). The 90 percent bias-corrected and bootstrapped confidence interval for the indirect effect of Links 1 through 3 is reported and does not contain zero for *Same* and *Diff*, supporting mediation of the results of H1 (but not for *Entity*). Link 4 is the direct effect of reporting a material weakness versus not on the greatest acceptable inventory value, when including both mediators in the model.

higher assessments that the write-down is legitimate (Link 2: all $p < 0.10$, one-tailed); and higher assessments of legitimacy result in higher assessments of the greatest acceptable inventory value (Link 3: all $p < 0.01$, one-tailed).[14]

Supplemental Analysis: Evidence that Results are Not Due to Perceived Negotiation

One potential alternative explanation for our results is that auditors who intend to report material weaknesses identified during ICFR testing allow more aggressive reporting by management because they view the two decisions as a form of negotiation. That is, if an auditor will report a material weakness—a negative outcome for management—then they could concede on a financial reporting issue and allow more aggressive reporting in order to give management a positive outcome in return. Such a result would not eliminate the worrisome unintended consequence we find, but it would suggest that another process is potentially at work. Additional analyses (untabulated) do not support this alternative explanation. Participants believed that the inventory obsolescence issue felt somewhat like a negotiation with management ("strongly disagree" [−5], "strongly agree" [+5]; M = 1.76). However, in the only significant difference between conditions (all other $t < 1.42$, all other $p > 0.16$, two-tailed), auditors in the *Entity-Level MW* condition (M = 0.91) less strongly believed that the issue felt like a negotiation than those in the *No MW* condition (M = 2.59, $t_{44} = 2.20$, $p = 0.03$, two-tailed). Such a result would mean that auditors in this *MW* (versus *No MW*) condition would be less, not more, willing to permit aggressive reporting in order to trade off their ICFR decision, which is opposite to the results reported earlier and in Table 2.

IV. EXPERIMENT TWO

Method: Participants and Design

In Experiment Two, we replicate results for H1 and test H2 to investigate whether licensing effects can be mitigated by prompting an audit quality goal. Participants in Experiment Two are 118 auditors at the manager level or above, who were

[14] We ran our analysis collapsed across the material weakness types and found similar results (untabulated): Links 1 and 3 were significant (both $p < 0.01$, one-tailed) and Link 2 was marginally significant ($p = 0.09$, one-tailed), while the confidence interval contained zero (−0.09, 2.53).

recruited through the Center for Audit Quality and American Accounting Association Access to Audit Personnel program; 76 are Big 4 auditors and 42 are non-Big 4 auditors.[15] Their audit experience ranges from 4.5 to 20 years, with a mean experience of eight years, and additional demographic data reported in Table 1 indicate that participants have sufficient experience to complete the study.[16] We desired more experienced participants for Experiment Two to examine the robustness of psychological licensing at higher levels or ranks within audit firms. We randomly assign participants in a 2 × 3 between-subjects design. We manipulate whether a material weakness is reported (*No MW* versus *MW*). We also manipulate the presence versus absence of a prompt related to the goal of audit quality. We include two versions of the prompt, as the framing of the goal may influence the effectiveness of mitigating the licensing effects observed in Experiment One. Prior research indicates that framing that emphasizes abstract (i.e., the why) versus concrete (i.e., the how) characteristics of the goal may be more effective, as the former can induce thoughts of goal commitment (Fishbach et al. 2006, 238). Thus, the prompt conditions are: *No Prompt* (to provide a baseline control condition), *Why Prompt,* and *How Prompt*.

Method: Procedures and Variables

The main experimental task, including assessments of inventory as the dependent variable, was identical to Experiment One with one exception: as participants in Experiment Two are at the manager level or above, they were assigned the role of "a member of the engagement team" (rather than the audit senior on the engagement as in Experiment One). For the material weakness manipulation, we use two of the conditions from Experiment One: *No MW* and *Same Account MW* (hereafter referred to simply as *MW*). We chose to include the *Same Account MW* condition in Experiment Two (rather than *Entity-Level MW* or *Different Account MW*) because we want to test whether our proposed intervention will mitigate psychological licensing for the most problematic scenario, and Experiment One results were strongest for this condition.

Participants in the *No Prompt* condition began with the main experimental task, whereas participants in the *Why Prompt* or *How Prompt* conditions began by completing the goal prompt manipulation before proceeding to the main experimental task. We developed the goal prompts by shortening and modifying the manipulations in Freitas, Gollwitzer, and Trope (2004) for the audit quality context. As illustrated in the following paragraph, participants given the *Why [How] Prompt* considered why [how] they improve and maintain audit quality:

> **Why [How] do I maintain audit quality?**
>
> For everything we do, there is always a reason why [process of how] we do it. Moreover, we can often trace the causes of our behavior back to broad goals that we have [follow our broad goals down to our very specific behaviors]. Research suggests that engaging in a thought exercise in which one thinks about how one's actions relate to one's ultimate goals [ultimate goals can be expressed through specific actions], can improve people's satisfaction. This thought exercise is intended to focus your attention on why [how] you do the things you do as an auditor. For this thought exercise, please consider the following activity: "improving and maintaining audit quality."

Following the respective paragraphs, participants listed "three ways in which improving and maintaining audit quality can help you meet important goals" in the *Why Prompt,* and "three means by which you can improve and maintain audit quality" in the *How Prompt*. To reinforce each prompt, on the next screen and for each of the three items listed, participants responded on an 11-point scale from "a little" (0) to "a lot" (+10) to the question "How much will improving and maintaining audit quality help you meet this important goal?" in the *Why Prompt,* and "How much will engaging in this activity improve and maintain audit quality?" in the *How Prompt*.

After the main task, as in Experiment One, all participants complete a series of additional questions, including process and demographic measures. We collect the same process measures as in Experiment One, related to dutifulness and legitimacy, to allow us to test whether our prompt intervention mitigates the underlying psychological licensing process.

---

[15] We obtained approval from the Institutional Review Board of the university where the experiment took place and all participants consented to participate, per IRB requirements. Three Big 4 firms and four non-Big 4 firms participated in the study. There are no significant main or interactive effects of firm type; thus, we combine observations across all firms in all analyses. While 145 auditors completed the study, via Qualtrics, we identified and excluded data from 26 participants who did not complete it in one uninterrupted sitting as instructed, evident by a long pause on any of the task screens (median time = 67 minutes) and the time taken to complete the study (median time = 148 minutes). Of the 119 participants who completed the study uninterrupted (median time = 24 minutes), we exclude one participant who did not respond to our main dependent variable. The number of excluded participants does not differ across conditions ($p = 0.16$, two-tailed).

[16] Participants' audit experience does not vary significantly across *MW* conditions ($F_{112,1} = 0.45$, $p = 0.50$, two-tailed), but does so across *Prompt* conditions ($F_{112,2} = 2.90$, $p = 0.06$, two-tailed); participants in *No Prompt* (M = 9.0 years) have more experience than participants in either *Why* (M = 8.1 years) or *How Prompt* (M = 7.3 years). However, participants' audit experience is unrelated to their assessments of inventory, and including experience as a covariate does not change results.

Results: Preliminary Tests

To validate that auditors completed the prompts as instructed, we analyze participant responses to the prompt manipulation. Prior research suggests that responses in the *Why Prompt* should fit a structure of "I get [response] (i.e., outcomes) by maintaining audit quality," whereas responses in the *How Prompt* should fit a structure of "I maintain audit quality by doing [response] (i.e., means)" (Fujita, Trope, Liberman, and Levin-Sagi 2006). Two independent coders, who were blind to hypotheses and experimental conditions, coded each response to indicate whether it fit the first structure (coded +1), the second structure (coded −1), or neither (coded 0). Initial inter-rater agreement was 79.6 percent and Cohen's kappa was 0.62. For each participant, their "structure" score is the average of their individual response scores; higher (lower) scores indicate responses that comply with the why (how) prompt. Auditors' structure scores were significantly more positive in the *Why Prompt* (M = 0.52) versus the *How Prompt* (M = −0.98, $F_{75,1} = 265.30$, $p < 0.01$, two-tailed); we found no significant effect of *MW* ($F_{75,1} = 1.01$, $p = 0.32$, two-tailed) or the *MW* and *Prompt* interaction ($F_{75,1} = 0.45$, $p = 0.51$, two-tailed).

In Experiment Two, we replicate tests of H1 and test H2 by including the same covariate (perceived ICFR experience) in both a 2 × 2 ANCOVA (Table 3) and a 2 × 3 ANCOVA (Table 4 and Figure 3), with the auditor's greatest acceptable inventory value as the dependent variable.[17] In Table 3, we collapse across the two prompt conditions.

Results: Tests Replicating H1 and the Psychological Licensing Process

In Experiment One, we find evidence supporting that auditors accept more aggressive client reporting when concurrently reporting a material weakness in ICFR than when reporting no such weakness (H1), and psychological licensing underlies this result. We find consistent corroborating evidence in Experiment Two. First, examining only the *No Prompt* condition in Table 4, the mean assessment is marginally significantly lower in *No MW* (M = 36.0) versus *MW* (M = 47.8, p = 0.06, one-tailed). We describe results from Table 4 only since we find nearly identical results across both Tables 3 and 4 in Panels A, C, and D.

Second, examining only the *No Prompt* condition in Figure 4, we report a mediation model with results similar to those of Figure 2. Specifically, we find the 90 percent confidence interval for the conditional indirect effect of the path depicted in Figure 4 does not contain zero (1.63, 8.20), replicating evidence that the effect of a material weakness on judgments of the greatest acceptable inventory value is mediated through assessments of dutiful performance and the legitimacy of the write-down. See Appendix A for means, standard errors, and other supporting details on factor scores of the two mediators. We also find significant effects for each of Links 1–3 in Figure 4 (all $p < 0.01$, one-tailed). Participants perceive themselves as performing more dutifully in *MW* versus *No MW* (Link 1); higher perceptions of performing dutifully result in higher assessments that the write-down is legitimate (Link 2); and higher assessments of legitimacy result in higher assessments of the greatest acceptable inventory value (Link 3).

Results: Tests of H2

We next test our prediction that licensing effects will be mitigated when auditors receive an audit quality goal prompt (H2). We first examine results in Table 3, where we collapse across the two prompts. We find a significant *MW* × *Prompt* interaction in Panel B ($p = 0.01$, two-tailed). Further, consistent with H2, in Panel C, we find that licensing is mitigated in *MW,* as mean inventory assessments are significantly lower when a prompt is provided (M = 36.0) versus not (M = 47.8, $p = 0.04$, one-tailed). Also consistent with H2, we find that when a prompt is provided, mean inventory assessments are significantly lower when auditors will report an MW versus no MW (M = 47.9, $p = 0.03$, two-tailed). Auditors who feel they acted dutifully by virtue of reporting an MW continue to act dutifully by permitting less aggressive accounting. They behave consistently across ICFR and financial reporting decisions after considering an audit quality goal.

Interestingly, auditors who consider an audit quality goal, but report no material weakness may also behave in a manner that is consistent across ICFR and financial reporting decisions. In *No MW,* mean inventory assessments are marginally significantly higher when a prompt is provided versus not (M = 36.0, $p = 0.07$, two-tailed). One potential explanation for this result is that prompted auditors feel they did not identify audit concerns by virtue of reporting no MW, and they subsequently, and consistently, identify fewer audit concerns by virtue of allowing more aggressive reporting. While we did not predict this result, it is not inconsistent with results of previous research (e.g., Fishbach and Dhar 2005) in which participants that initially act *not* in accordance with their goal subsequently continue to act *not* in accordance with their goal. It is unclear if our prompted auditors believe that reporting no MW is incongruent, congruent, or simply not incongruent with the audit quality goal, but to the extent that they believe it is a less goal-congruent action than reporting an MW, then it may have caused them to be less goal-congruent in their reporting decision.

[17] ICFR experience is not significantly correlated with either independent factor (both $p > 0.44$) and results also hold when excluding it from the model.

TABLE 3

Experiment Two
Tests of Hypotheses (2 × 2 ANCOVA)

Panel A: Greatest Acceptable Inventory Value: Adjusted Mean, (Standard Error), n

	Prompt	
	No Prompt	**Either Prompt**
Material Weakness		
No MW	36.04	47.93
	(5.21)	(3.69)
	20	40
MW	47.76	35.96
	(5.21)	(3.79)
	20	38

Panel B: ANCOVA

Source	df	MS	F	p-value[b]
Material Weakness	1	0.41	<0.01	0.978
Prompt	1	0.06	<0.01	0.992
MW × Prompt	1	3682.35	6.78	0.011
ICFR-Related Experience	1	4.72	0.01	0.926
Error	113	543.28		

Panel C: Simple Effects Tests

	t	p-value
No Prompt: No MW versus *MW* (H1)	−1.59	0.058[a]
MW: Prompt versus *No Prompt* (H2)	−1.83	0.035[a]
No MW: Prompt versus *No Prompt*	1.86	0.065[b]
Prompt: No MW versus *MW*	2.26	0.026[b]

[a] Directional prediction; reported p-values are one-tailed.
[b] No directional prediction; reported p-values are two-tailed.
Participants assessed (in $M) the lowest inventory write-down they would accept, and we subtracted this assessed value from the total inventory value ($100.8M) to determine the greatest inventory value they would accept. This assessment is covariate-adjusted for experience concluding on internal controls related to the ICFR opinion (self-assessed relative to peers on a scale from "no experience" [0] to "extensive experience" [10]). The two between-subjects Material Weakness (MW) conditions are: *No MW* and *MW*. The two between-subjects prompt conditions are: *No Prompt* and *Prompt* (collapsed across *Why Prompt* and *How Prompt*).

Results in Table 4, where we do not collapse across the two prompts, are similar to results in Table 3 and show that the two prompts had nearly identical effects. We find the *MW* × *Prompt* (i.e., 2 × 3) interaction in Panel B is significant ($p = 0.04$, two-tailed), as is each *MW* × *Prompt* (i.e., 2 × 2) interaction in Panel C (both $p < 0.05$, two-tailed). Further, there is no significant difference ($p = 0.74$, two-tailed) between the *MW* × *Prompt* interactions of *No Prompt* and *Why Prompt* (only) and *No Prompt* and *How Prompt* (only). Additional support for H2 and for licensing being mitigated when an MW is reported is shown in Panel D. Mean inventory assessments are significantly lower for *Why Prompt* ($M = 33.8$) versus *No Prompt* ($M = 47.8$, $p = 0.04$, one-tailed) and marginally significantly lower for *How Prompt* ($M = 37.7$) versus *No Prompt* ($p = 0.09$, one-tailed). We also find that for *No MW,* mean assessments are higher in *Why Prompt* ($M = 47.8$) or *How Prompt* ($M = 48.1$) compared to *No Prompt* ($M = 36.0$), albeit not quite at a significant level (both $p = 0.11$, two-tailed). Finally, we find that when either prompt is provided, mean assessments are higher when auditors will report *No MW* versus an *MW,* at a marginally significant level for *Why Prompt* ($p = 0.07$, two-tailed), but not at a significant level for *How Prompt* ($p = 0.16$, two-tailed).

Overall, we generally find support for H2 and find similar effects of prompting a goal across the two prompts. Prompting a goal (versus not) leads auditors to permit less aggressive financial reporting judgments when they will report an MW, but it also results in auditors allowing more aggressive reporting when no MW will be reported. This latter result suggests that consideration should be given to the situations in which prompts are provided.

TABLE 4

Experiment Two
Tests of Hypotheses (2 × 3 ANCOVA)

Panel A: Greatest Acceptable Inventory Value: Adjusted Mean, (Standard Error), n

	Prompt		
	None	**Why**	**How**
Material Weakness			
No MW	36.04	47.77	48.12
	(5.25)	(5.13)	(5.39)
	20	21	19
MW	47.76	33.84	37.67
	(5.26)	(5.72)	(5.13)
	20	17	21

Panel B: ANCOVA

Source	**df**	**MS**	**F**	**p-value[b]**
Material Weakness	1	520.14	0.94	0.334
Prompt	2	42.13	0.08	0.927
MW × Prompt	2	1892.65	3.43	0.036
ICFR-Related Experience	1	2.40	<0.01	0.948
Error	111	551.82		

Panel C: Planned Tests

	F	**p-value[b]**
MW × Prompt for *No Prompt* and *Why Prompt* only	5.73	0.018
MW × Prompt for *No Prompt* and *How Prompt* only	4.43	0.038
Test of difference in significance between two interactions above	0.11	0.744

Panel D: Planned Tests, continued from Panel C

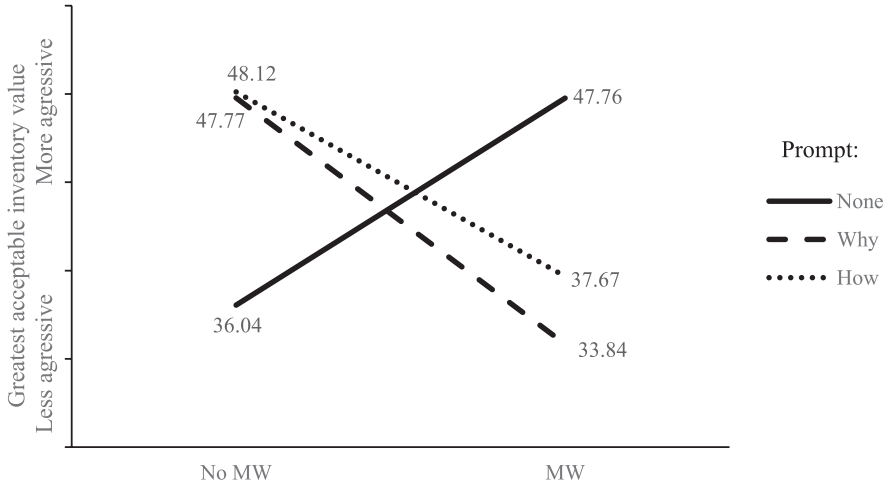
Simple Effects Tests	**t**	**p-value**
No Prompt: No MW versus *MW* (H1)	−1.58	0.059[a]
MW: Why Prompt versus *No Prompt* (H2)	−1.79	0.038[a]
No MW: Why Prompt versus *No Prompt*	1.60	0.113[b]
Why Prompt: No MW versus *MW*	1.81	0.073[b]
MW: How Prompt versus *No Prompt* (H2)	−1.37	0.086[a]
No MW: How Prompt versus *No Prompt*	1.60	0.112[b]
How Prompt: No MW versus *MW*	1.40	0.164[b]

[a] Directional prediction; reported p-values are one-tailed.
[b] No directional prediction; reported p-values are two-tailed.
Participants assessed (in $M) the lowest inventory write-down they would accept, and we subtracted this assessed value from the total inventory value ($100.8M) to determine the greatest inventory value they would accept. This assessment is covariate-adjusted for experience concluding on internal controls related to the ICFR opinion (self-assessed relative to peers on a scale from "no experience" [0] to "extensive experience" [10]). The two between-subjects Material Weakness (MW) conditions are: *No MW* and *MW*. The three between-subjects prompt conditions are: *No Prompt, Why Prompt,* and *How Prompt*.

Supplemental Analysis: Evidence of Mitigation of Psychological Licensing

As noted earlier, results from both experiments provide robust evidence for psychological licensing being the process through which auditors accept more aggressive client reporting when reporting a material weakness in ICFR versus not, in conditions where a prompt is not provided. To the extent that prompting a goal helps mitigate licensing effects, then we expect certain links in our licensing model (e.g., Figure 4) to be broken. Prior research suggests that goal-focused individuals can still feel they have performed dutifully given a good first deed, but this action can reinforce their commitment to their goal such that

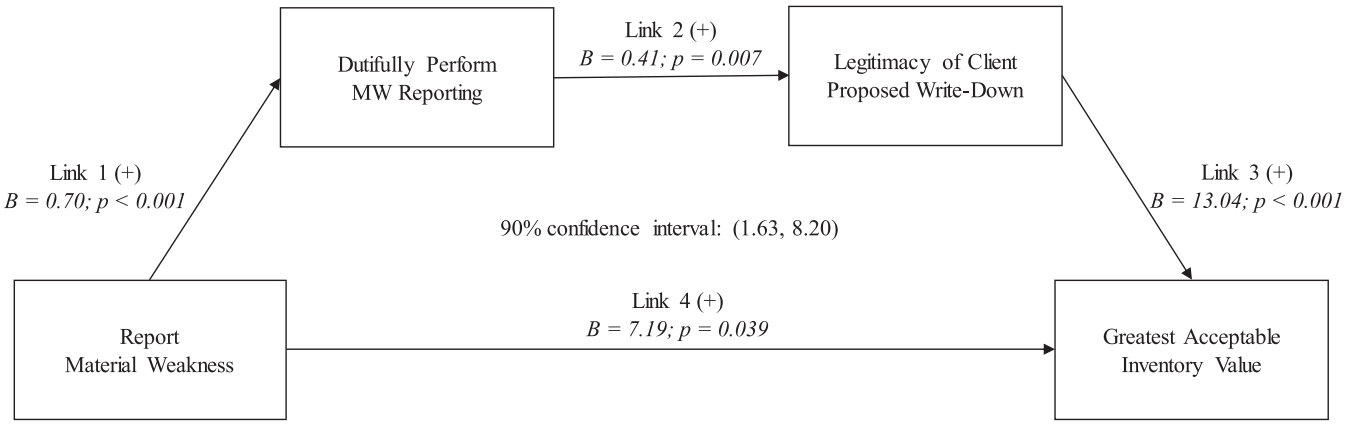
FIGURE 3
Experiment Two
Results Graph



Refer to Table 4 for independent and dependent variable descriptions.

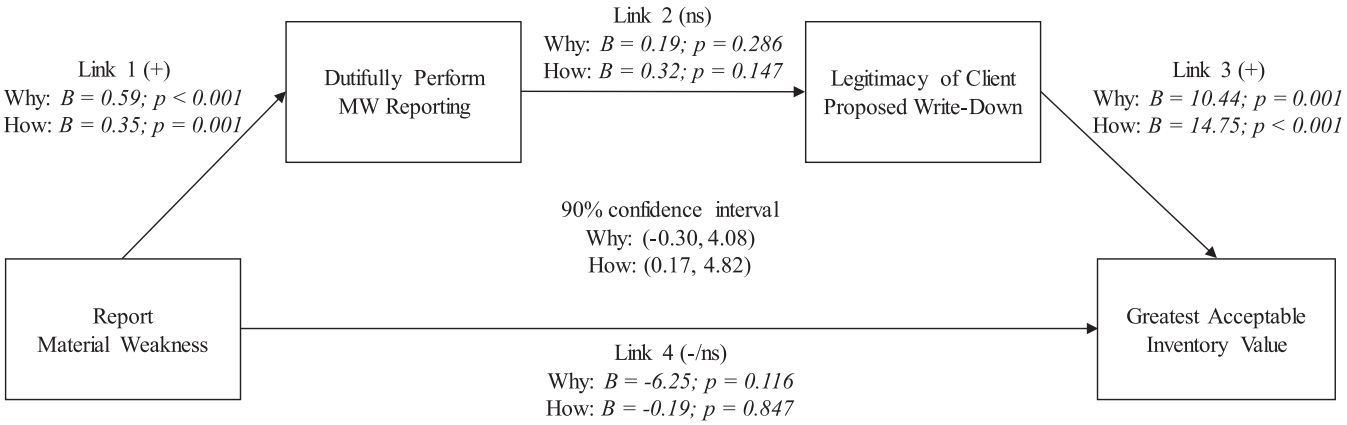
subsequent actions will be consistent with, not contrary to, the first action (Fishbach et al. 2006; Susewind and Hoelzl 2014). We expect Link 1 (positive association between reporting an MW and dutiful perceptions) to be intact, while Link 2 (positive association between dutiful perceptions and assessments of write-down legitimacy) will be broken. Examining *Why Prompt* and *How Prompt* separately in Figure 5, we find evidence to support our expectations. Link 1 is significant (both $p < 0.01$, two-

FIGURE 4
Experiment Two
Psychological Licensing Process for No MW versus MW
(No Prompt Condition Only)



This figure captures the psychological licensing process in the *No Prompt* condition for when a material weakness is reported versus not at all, using the Preacher and Hayes (2008) method for testing indirect effects. The model includes the same covariate as the model in Tables 3 and 4, and the mediators (*Dutifulness* and *Legitimacy*) are measured as defined in Appendix A. Each individual link of Links 1 through 3 is significant in the expected direction (all p-values are one-tailed). We expect reporting a material weakness, compared to reporting no material weakness, will increase the perception that the auditor performed dutifully (Link 1); these higher perceptions will, in turn, increase the auditor's assessment that the client-proposed write-down is legitimate (Link 2); and these higher assessments will, in turn, increase the auditor's assessment of the greatest acceptable inventory value (Link 3). The 90 percent bias-corrected and bootstrapped confidence interval for the indirect effect of Links 1 through 3 is reported and does not contain zero, supporting mediation of the results of H1. Link 4 is the direct effect of reporting a material weakness versus not on the greatest acceptable inventory value, when including both mediators in the model.

FIGURE 5
Experiment Two
Mitigation of Psychological Licensing Process for No MW versus MW
(Both Prompt Conditions)



This figure captures the mitigation of the psychological licensing process in either *Prompt* condition (*Why Prompt* and *How Prompt*), for when a material weakness is reported versus not at all, using the Preacher and Hayes (2008) method for testing indirect effects. The model includes the same covariate as the model in Tables 3 and 4, and the mediators (*Dutifulness* and *Legitimacy*) are measured as defined in Appendix A. Links 1 and 3 are significant for both prompts, as expected, while Link 2 is not significant for both prompts, as expected (all p-values are two-tailed). We expect that reporting a material weakness, compared to reporting no material weakness, will increase the perception that the auditor performed dutifully (Link 1). However, due to the prompt making salient the goal of audit quality, we do not expect these higher perceptions to, in turn, increase the auditor's assessment that the client-proposed write-down is legitimate (Link 2). Higher assessments can still increase the auditor's assessment of the greatest acceptable inventory value (Link 3). The 90 percent bias-corrected and bootstrapped confidence interval for the indirect effect of Links 1 through 3 is reported: for *Why,* it contains zero, indicating a lack of support for mediation, but for *How,* it does not contain zero, indicating some support for mediation. Link 4 is the direct effect of reporting a material weakness versus not on the greatest acceptable inventory value, when including both mediators in the model.

tailed), but Link 2 is not (both $p > 0.14$, two-tailed); Link 3 is also significant (both $p < 0.01$, two-tailed). Finally, the 90 percent confidence interval for the indirect effect contains zero for *Why Prompt* (−0.30, 4.08), but not for *How Prompt* (0.17, 4.82), suggesting that the effect of reporting a material weakness on inventory assessments is not (somewhat) mediated by these measures for the former (latter) prompt.

V. CONCLUSIONS

We report the results of two experiments designed to investigate how auditors respond to aggressive and potentially materially misstated reporting by management when they do or do not concurrently report a material weakness in ICFR. We find results consistent with theory on psychological licensing in that auditors accept more aggressive client reporting when they report a material weakness in ICFR versus when they report no material weakness. These results hold regardless of whether the material weakness is in the same or a different account as the potential material misstatement. Further, these results are mediated by auditors' beliefs that they acted dutifully (to a greater extent when reporting a material weakness), which, in turn, gave them license to view the aggressive client reporting as more legitimate and accept it more readily. We also provide evidence of a potential remedy to mitigate the licensing effects: auditors reporting a material weakness in ICFR accept less aggressive reporting when receiving a prompt related to the goal of audit quality versus not. Caution should be exercised in deciding when to use this remedy, however, as we also find that auditors who report no material weakness may accept more aggressive reporting when receiving the prompt versus not.

Our findings provide an explanation for why empirical evidence shows that reporting of a material weakness in ICFR is associated with lower accruals quality (Doyle et al. 2007). While archival research cannot decipher whether such a result is due to auditor failure to detect/correct misstatements that potentially result from weak ICFR, or due to higher risk that manifests despite auditor efforts to increase substantive testing in light of weak ICFR (Hogan and Wilkins 2008), our results indicate that the former is at least one reason. By using an experiment, we are able to hold constant the amount of substantive testing and the presence and type of aggressive reporting. Under such conditions, auditors should be just as likely to require the client to correct aggressive reporting when reporting a material weakness as when not. Again, we find that this is not the case and that for different types of material weakness across two experiments, auditors accept more aggressive reporting than when no

material weakness is present. Our study also extends prior research on psychological licensing that focuses on evidence related to outcomes, but rarely evaluates the underlying processes (e.g., Griffin 2014). We provide evidence of the psychological licensing process underlying the auditors' behavior.

This study has implications for academic research, audit professionals seeking to provide value to investors through audit procedures, and regulators seeking to increase audit quality and the reporting of material weaknesses in ICFR. To be clear, we are not proposing that regulators stop encouraging the reporting of material weaknesses, as material weaknesses provide valuable information to capital markets on financial reporting quality. But as regulators continue to press auditors to increase reporting of material weaknesses in ICFR (e.g., Croteau 2013; PCAOB 2013), they should recognize the potential for unintended consequences, such as those which we find, and exercise caution in how to achieve improved ICFR reporting without sacrificing financial reporting quality. While the audit profession obviously recognizes audit quality as a critical goal, our study shows evidence that auditor judgments can suffer from subconscious licensing effects when making a series of explicit audit conclusions.

Our study also provides evidence that a simple and timely prompt to consider why or how auditors improve and maintain audit quality can moderate licensing effects, but future research can consider other prompts or mechanisms that could be as or more effective than ours. Further, in our study, the prompt occurs minutes before the material weakness and material misstatement decisions. In practice, prompts likely occur during firm training and continuing education, which can occur at different time intervals. Future research can further explore the importance of the timing of such prompts. For example, providing a prompt while an audit is ongoing, particularly after a material weakness is encountered, could have the most value. Additionally, concerns may exist around the effectiveness of a prompt over time or after repeated exposure. Future research could consider these factors, although such concerns are reduced by evidence in prior research that repeated prompts have long-term and even increased effectiveness (e.g., Cave 1997; Garland, Gaylord, and Fredrickson 2011). Finally, while we investigate licensing effects in the integrated audit setting, due to noted auditor struggles in this area (e.g., Bhaskar et al. 2019) and puzzling archival findings about material weaknesses (e.g., Hogan and Wilkins 2008), auditors are likely susceptible to licensing in other audit areas (e.g., Griffin 2014). Future research can investigate licensing, and ways to successfully mitigate licensing, in other audit contexts.

A potential limitation of our study is that we do not examine whether auditors would change substantive testing in response to the ICFR conclusions, prior to dealing with the aggressive financial reporting issue. However, prior audit research shows that auditors insufficiently adjust their audit plans in response to material weaknesses in ICFR (Hammersley et al. 2011; Mauldin and Wolfe 2014), so it is unlikely that the increased leniency displayed by our participants would diminish with changes, if any, they might have made to audit procedures. Moreover, increases in audit procedures help to *detect* material misstatements, whereas participants in our study were provided with a potential misstatement that was already detected and, thus, were tasked with deciding whether to require management to *correct* a misstatement. Therefore, we extend prior research by focusing on the *outcomes* of overall audit procedures, and demonstrating that auditor judgments systematically differ depending upon factors other than actual procedures performed.

REFERENCES

Asare, S. K., B. C. Fitzgerald, L. E. Graham, J. R. Joe, E. M. Negangard, and C. J. Wolfe. 2013. Auditors' internal control over financial reporting decisions: Analysis, synthesis, and research directions. *Auditing: A Journal of Practice & Theory* 32 (Supplement 1): 131–166. https://doi.org/10.2308/ajpt-50345

Ashbaugh-Skaife, H., D. W. Collins, W. R. Kinney, and R. LaFond. 2008. The effect of SOX internal control deficiencies and their remediation on accrual quality. *The Accounting Review* 83 (1): 217–250. https://doi.org/10.2308/accr.2008.83.1.217

Backof, A. G., T. D. Carpenter, and J. Thayer. 2018. Auditing complex estimates: How do construal level and evidence formatting impact auditors' consideration of inconsistent evidence? *Contemporary Accounting Research* 35 (4): 1798–1815. https://doi.org/10.1111/1911-3846.12368

Bauer, T. 2015. The effects of client identity strength and professional identity salience on auditor judgments. *The Accounting Review* 90 (1): 95–114. https://doi.org/10.2308/accr-50863

Bedard, J. 1989. An archival investigation of audit program planning. *Auditing: A Journal of Practice & Theory* 9 (1): 57–71.

Bhaskar, L. S., J. H. Schroeder, and M. L. Shepardson. 2019. Integration of internal control and financial statement audits: Are two audits better than one? *The Accounting Review* 94 (2): 53–81. https://doi.org/10.2308/accr-52197

Braun, K. 2001. The disposition of audit-detected misstatements: An examination of risk and reward factors and aggregation effects. *Contemporary Accounting Research* 18 (1): 71–99.

Cain, D. M., G. Loewenstein, and D. A. Moore. 2005. The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *Journal of Legal Studies* 34 (1): 1–25. https://doi.org/10.1086/426699

Cain, D. M., G. Loewenstein, and D. A. Moore. 2011. When sunlight fails to disinfect: Understanding the perverse effects of disclosing conflicts of interest. *Journal of Consumer Research* 37 (5): 836–857. https://doi.org/10.1086/656252

Cannon, N., and J. C. Bedard. 2017. Auditing challenging fair value measurements: Evidence from the field. *The Accounting Review* 92 (4): 81–114. https://doi.org/10.2308/accr-51569

Cave, C. 1997. Very long-lasting priming in picture naming. *Psychological Science* 8 (4): 322–325. https://doi.org/10.1111/j.1467-9280.1997.tb00446.x

Center for Audit Quality (CAQ). 2018. *Request for Proposals for Academic Research in Auditing*. Washington, DC: CAQ.

Church, B., and X. Kuang. 2009. Conflicts of interest, disclosure, and (costly) sanctions: Experimental evidence. *Journal of Legal Studies* 38 (2): 505–532. https://doi.org/10.1086/596117

Church, B., S. M. Davis, and S. A. McCracken. 2008. The auditor's reporting model: A literature overview and research synthesis. *Accounting Horizons* 22 (1): 69–90. https://doi.org/10.2308/acch.2008.22.1.69

Conway, P., and J. Peetz. 2012. When does feeling moral actually make you a better person? Conceptual abstraction moderates whether past moral deeds motivate consistency or compensatory behavior. *Personality and Social Psychology Bulletin* 38 (7): 907–919. https://doi.org/10.1177/0146167212442394

Croteau, B. T. 2013. Audit policy and current auditing and internal control matters. Remarks before the 2013 National Conference on Current SEC and PCAOB Developments, December 9. Available at: http://www.sec.gov/News/Speech/Detail/Speech/1370540472057#

Doyle, J., W. Ge, and S. McVay. 2007. Accruals quality and internal control over financial reporting. *The Accounting Review* 82 (5): 1141–1170. https://doi.org/10.2308/accr.2007.82.5.1141

Effron, D. A., and B. Monin. 2010. Letting people off the hook: When do good deeds excuse transgressions? *Personality and Social Psychology Bulletin* 36 (12): 1618–1634. https://doi.org/10.1177/0146167210385922

Effron, D. A., J. S. Cameron, and B. Monin. 2009. Endorsing Obama licenses favoring whites. *Journal of Experimental Social Psychology* 45 (3): 590–593. https://doi.org/10.1016/j.jesp.2009.02.001

Fishbach, A., and R. Dhar. 2005. Goals as excuses or guides: The liberating effect of perceived goal progress on choice. *Journal of Consumer Research* 32 (3): 370–377. https://doi.org/10.1086/497548

Fishbach, A., R. Dhar, and Y. Zhang. 2006. Subgoals as substitutes or complements: The role of goal accessibility. *Journal of Personality and Social Psychology* 91 (2): 232–242. https://doi.org/10.1037/0022-3514.91.2.232

Franzel, J. M. 2014. Effective audits of internal control in the current "perfect storm." Speech presented at the Institute of Internal Auditors 2014 General Audit Management Conference, March 26. Available at: http://pcaobus.org/News/Speech/Pages/03262014_IIA.aspx

Freitas, A. L., P. Gollwitzer, and Y. Trope. 2004. The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology* 40 (6): 739–752. https://doi.org/10.1016/j.jesp.2004.04.003

Fujita, K., Y. Trope, N. Liberman, and M. Levin-Sagi. 2006. Construal levels and self-control. *Journal of Personality and Social Psychology* 90 (3): 351–367. https://doi.org/10.1037/0022-3514.90.3.351

Garland, E., S. Gaylord, and B. Fredrickson. 2011. Positive reappraisal mediates the stress-reductive effects of mindfulness: An upward spiral process. *Mindfulness* 2 (1): 59–67. https://doi.org/10.1007/s12671-011-0043-8

Griffin, J. 2014. The effects of uncertainty and disclosure on auditors' fair value materiality decisions. *Journal of Accounting Research* 52 (5): 1165–1193. https://doi.org/10.1111/1475-679X.12059

Griffith, E. E., J. S. Hammersley, K. Kadous, and D. Young. 2015. Auditor mindsets and audits of complex estimates. *Journal of Accounting Research* 53 (1): 49–77. https://doi.org/10.1111/1475-679X.12066

Guggenmos, R. D., M. D. Piercey, and C. P. Agoglia. 2018. Custom contrast testing: Current trends and a new approach. *The Accounting Review* 93 (5): 223–244. https://doi.org/10.2308/accr-52005

Hammersley, J. S., K. M. Johnstone, and K. Kadous. 2011. How do audit seniors respond to heightened fraud risk? *Auditing: A Journal of Practice & Theory* 30 (3): 81–101. https://doi.org/10.2308/ajpt-10110

Hoffman, V. B., and M. F. Zimbelman. 2009. Do strategic reasoning and brainstorming help auditors change their standard audit procedures in response to fraud risk? *The Accounting Review* 84 (3): 811–837. https://doi.org/10.2308/accr.2009.84.3.811

Hogan, C. E., and M. S. Wilkins. 2008. Evidence on the audit risk model: Do auditors increase audit fees in the presence of internal control deficiencies? *Contemporary Accounting Research* 25 (1): 219–242. https://doi.org/10.1506/car.25.1.9

Jamal, K., E. Marshall, and H. T. Tan. 2016. Does disclosure of conflict of interest increase or decrease bias? *Auditing: A Journal of Practice & Theory* 35 (1): 89–99. https://doi.org/10.2308/ajpt-51018

Joe, J., A. Wright, and S. Wright. 2011. The impact of client and misstatement characteristics on the disposition of proposed audit adjustments. *Auditing: A Journal of Practice & Theory* 30 (2): 103–124. https://doi.org/10.2308/ajpt-50007

Kadous, K., and Y. Zhou. 2019. How does intrinsic motivation improve auditor judgment in complex audit tasks? *Contemporary Accounting Research* 36 (1): 108–131.

Kadous, K., S. 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. https://doi.org/10.2308/accr.2003.78.3.759

Klotz, A. C., and M. C. Bolino. 2013. Citizenship and counterproductive work behavior: A moral licensing view. *Academy of Management Review* 38 (2): 292–306. https://doi.org/10.5465/amr.2011.0109

Koch, C., and C. Schmidt. 2010. Disclosing conflicts of interest—Do experience and reputation matter? *Accounting, Organizations and Society* 35 (1): 95–107. https://doi.org/10.1016/j.aos.2009.05.001

Kouchaki, M. 2011. Vicarious moral licensing: The influence of others' past moral actions on moral behavior. *Journal of Personality and Social Psychology* 101 (4): 702–715. https://doi.org/10.1037/a0024552

Lennox, C. S., X. Wu, and T. Zhang. 2014. Does mandatory rotation of audit partners improve audit quality? *The Accounting Review* 89 (5): 1775–1803. https://doi.org/10.2308/accr-50800

Luippold, B., T. Kida, M. Piercey, and J. Smith. 2015. Managing audits to manage earnings: The impact of diversions on an auditor's detection of earnings management. *Accounting, Organizations and Society* 41: 39–54. https://doi.org/10.1016/j.aos.2014.07.005

Mauldin, E. G., and C. J. Wolfe. 2014. How do auditors address control deficiencies that bias accounting estimates? *Contemporary Accounting Research* 31 (3): 658–680. https://doi.org/10.1111/1911-3846.12051

Mazar, N., and C. Zhong. 2010. Do green products make us better people? *Psychological Science* 21 (4): 494–498. https://doi.org/10.1177/0956797610363538

Merritt, A. C., D. A. Effron, and B. Monin. 2010. Moral self-licensing: When being good frees us to do bad. *Social and Personality Psychology Compass* 4 (5): 344–357. https://doi.org/10.1111/j.1751-9004.2010.00263.x

Messier, W. F., and M. Schmidt. 2018. Offsetting misstatements: The effect of misstatement distribution, quantitative materiality, and client pressure on auditors' judgments. *The Accounting Review* 93 (4): 335–357. https://doi.org/10.2308/accr-51954

Miller, D., and D. Effron. 2010. Psychological license: When it is needed and how it functions. In *Advances in Experimental Social Psychology*, edited by Zanna, M., and J. Olson, 115–155. Waltham, MA: Academic Press.

Mock, T. J., and A. M. Wright. 1993. An exploratory study of auditor evidential planning. *Auditing: A Journal of Practice & Theory* 12 (2): 39–61.

Mock, T. J., and A. M. Wright. 1999. Are audit programs risk-adjusted? *Auditing: A Journal of Practice & Theory* 18 (1): 55–74. https://doi.org/10.2308/aud.1999.18.1.55

Monin, B., and D. Miller. 2001. Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology* 81 (1): 33–43. https://doi.org/10.1037/0022-3514.81.1.33

Mullen, E., and B. Monin. 2016. Consistency versus licensing effects of past moral behavior. *Annual Review of Psychology* 67 (1): 363–385. https://doi.org/10.1146/annurev-psych-010213-115120

Preacher, K., and A. Hayes. 2008. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods* 40 (3): 879–891. https://doi.org/10.3758/BRM.40.3.879

Public Company Accounting Oversight Board (PCAOB). 2002. *Due Professional Care in the Performance of Work. Auditing Standard No. 1015*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2007. *An Audit of Internal Control over Financial Reporting that is Integrated with an Audit of Financial Statements. Auditing Standard No. 2201*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2009. *Report on the First-Year Implementation of Auditing Standard No. 5: An Audit of Internal Control over Financial Reporting that is Integrated with an Audit of Financial Statements. Release No. 2009-006*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2010. *Audit Risk. Auditing Standard No. 1101*. Washington, DC: PCAOB.

Public Company Accounting Oversight Board (PCAOB). 2013. *Considerations for Audits of Internal Control over Financial Reporting. Staff Audit Practice Alert No. 11*. Washington, DC: PCAOB.

Rasso, J. T. 2015. Construal instructions and professional skepticism in evaluating complex estimates. *Accounting, Organizations and Society* 46: 44–55. https://doi.org/10.1016/j.aos.2015.03.003

Ricci, M. 2020. *How higher client service performance affects auditors' willingness to challenge management's preferred accounting*. Working paper, University of Florida. Available at: https://doi.org/10.2139/ssrn.3098412

Rice, S., and D. Weber. 2012. How effective is internal control reporting under SOX 404? Determinants of the (non-) disclosure of existing material weaknesses. *Journal of Accounting Research* 50 (3): 811–843. https://doi.org/10.1111/j.1475-679X.2011.00434.x

Sachdeva, S., R. Iliev, and D. Medin. 2009. Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science* 20 (4): 523–528. https://doi.org/10.1111/j.1467-9280.2009.02326.x

Smith, J. F., and T. Kida. 1991. Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin* 109 (3): 472–489. https://doi.org/10.1037/0033-2909.109.3.472

Susewind, M., and E. Hoelzl. 2014. A matter of perspective: Why past moral behavior can sometimes encourage and other times discourage future moral striving. *Journal of Applied Social Psychology* 44 (3): 201–209. https://doi.org/10.1111/jasp.12214

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H.R. 3763]. Washington, DC: Government Publishing Office.

Wright, A., and S. Wright. 1997. An examination of factors affecting the decision to waive audit adjustments. *Journal of Accounting, Auditing and Finance* 12 (1): 15–36. https://doi.org/10.1177/0148558X9701200102

APPENDIX A

Both Experiments
Details of Factor-Score Measures

Appendix A provides the details related to the factor-score measures used in our process analyses as shown in Figure 2 (for Experiment One) and Figures 4 and 5 (for Experiment Two).

Both Experiments: Measures

For *Dutifulness,* participants rate the extent to which they agree that their ICFR decision (1) shows to outsiders that they are objective in evaluating financial statements, (2) forewarns investors about the risk of material misstatement, and (3) has value to investors, each rated on an 11-point scale from "strongly disagree" (–5) to "strongly agree" (+5). For *Legitimacy,* participants assess, on an 11-point scale, the following regarding the client's proposed write-down: (1) its appropriateness (from "not at all appropriate" [0] to "completely appropriate" [+10]); (2) its aggressiveness (from "very conservative" [–5] to "very aggressive" [+5]); (3) their likelihood of requesting a different write-down (from "not at all likely" [0] to "very likely" [+10]); and (4) their agreement that the financial statements will be materially misstated if they allow the proposed write-down (from "no agreement" [0] to "complete agreement" [+10]). The latter three measures are reverse-coded (R). See Table 5.

TABLE 5
Both Experiments

Panel A: Experiment One: Factor-Score Measures

		Dutifulness Factor-Score		*Legitimacy* Factor-Score	
Condition	**n**	**Mean**	**Std. Error**	**Mean**	**Std. Error**
No MW	22	−0.85	(0.18)	−0.14	(0.21)
Any Material Weakness	67	0.28	(0.10)	0.05	(0.12)
Same Account MW	23	0.53	(0.18)	0.40	(0.21)
Different Account MW	21	−0.14	(0.19)	−0.19	(0.22)
Entity-Level MW	23	0.41	(0.18)	−0.10	(0.21)

Refer to Table 2 for condition descriptions. Means are standardized between −1 and +1.

Panel B: Experiment One: Factor Patterns and Weights

Item	***Dutifulness***	**Item**	***Legitimacy***
Eigenvalue	2.18	Eigenvalue	2.02
Percent of Variance Explained	72.71	Percent of Variance Explained	50.54
Objective	0.81	Appropriateness	0.72
Forewarns	0.86	Aggressiveness (R)	0.61
Value	0.89	Request Different Amount (R)	0.78
		Materially Misstated (R)	0.72

For both *Dutifulness* and *Legitimacy,* we use the only factor that had an eigenvalue > 1.

Panel C: Experiment Two: Factor-Score Measures

		Dutifulness Factor-Score		*Legitimacy* Factor-Score	
Condition	**n**	**Mean**	**Std. Error**	**Mean**	**Std. Error**
No MW	60	−0.54	(0.11)	0.13	(0.13)
No Prompt	20	−0.75	(0.19)	−0.15	(0.22)
Why Prompt	21	−0.49	(0.18)	0.24	(0.22)
How Prompt	19	−0.36	(0.19)	0.31	(0.23)
MW	58	0.56	(0.11)	−0.14	(0.13)
No Prompt	20	0.64	(0.19)	−0.13	(0.22)
Why Prompt	17	0.67	(0.20)	−0.29	(0.24)
How Prompt	21	0.38	(0.18)	−0.01	(0.22)

Refer to Table 4 for condition descriptions. Means are standardized between −1 and +1.

Panel D: Experiment Two: Factor Patterns and Weights

Item	***Dutifulness***	**Item**	***Legitimacy***
Eigenvalue	2.45	Eigenvalue	2.19
Percent of Variance	81.68	Percent of Variance	54.63
Objective	0.90	Appropriateness	0.70
Forewarns	0.91	Aggressiveness (R)	0.55
Value	0.90	Request different amount (R)	0.82
		Materially misstated (R)	0.85

For both *Dutifulness* and *Legitimacy,* we use the only factor that had an eigenvalue > 1.