Financial rewards: The Trojan horse of frivolous whistleblowing?

Dwight Waeye *
Ghent University, Department of Accounting, Corporate Finance and Taxation
Sint-Pietersplein 7, 9000 Ghent, Belgium
Dwight.Waeye@UGent.be

Sophie Maussen
Ghent University, Department of Accounting, Corporate Finance and Taxation
FlandersMake@UGent – Corelab CVAMO
Sint-Pietersplein 7, 9000 Ghent, Belgium
Sophie.Maussen@UGent.be

Acknowledgements: We would like to thank participants at the MAS Midyear Meeting (Orlando, 2024), the EAA 45th Annual Congress (Helsinki, 2023), the MCA conference (Durham, 2023), the Accounting Research Day (Ghent, 2023) and the FEB Research Day (Ghent, 2023) for useful comments and suggestions. We are also grateful for the helpful comments on earlier drafts provided by Leslie Berger, Martine Cools, Ignace De Beelde, Chris Ford, Sophie Hoozée, Evelien Opdecam, Andreas Ostermaier, Tom Vanacker, and Berend van der Kolk.

Funding: The first author acknowledges financial support provided by the Research Foundation – Flanders (FWO, grant number: 1117223N).

Data availability: The data and research instrument are available from the authors upon request.

*** Corresponding author:** Dwight.Waeye@UGent.be, +32 9 264 35 80

Financial rewards: The Trojan horse of frivolous whistleblowing?

ABSTRACT

While financial whistleblowing rewards effectively increase the number of submitted tips, there are concerns with regard to the quality of these tips. Responding to these concerns, we study whether financial rewards incentivize frivolous (i.e., inaccurate and opportunistic) whistleblowing and how organizations can address this undesirable side effect. Relying upon the revised Fraud Triangle, we identify rationalization as an important confounding variable which may align mixed results in prior literature. Results from an online experiment focusing on a participative budgeting scenario show that rewards for accurate whistleblowing incentivize frivolous whistleblowing when rationalization is present, but not when rationalization is absent. Second, we also find evidence that adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing disincentivizes frivolous whistleblowing. Process evidence and supplemental analyses confirm that financial rewards do not necessarily incentivize frivolous whistleblowing, and when they do, that this is driven by individuals' opportunism. Our results show that organizations can include financial incentives into their internal whistleblowing policy without having to fear an increase in frivolous whistleblowing, which has important implications for research and practice.

Keywords Financial incentive schemes; Fraud Triangle; Frivolous whistleblowing; Rationalization; Whistleblowing behavior

Data availability: The data and research instruments are available from the authors upon request.

1 | INTRODUCTION

Due to major accounting scandals in the last decades, firms are under increasing pressure to implement effective internal whistleblowing policies (Chen, Nichol and Zhou 2017). Existing literature has primarily focused on increasing the number of submitted tips, revealing that financial rewards incentivize whistleblowing (e.g., Butler, Serra and Spagnolo 2020). However, concerns have been raised with regard to the quality of these tips (e.g., Kuang, Lee and Qin 2021; Andon et al. 2018), emphasizing the importance of maintaining a balance between quantity and quality. For instance, while the Securities and Exchange Commission's (SEC) whistleblower award program (Dodd-Frank Act 2010) received 14,116 whistleblowing tips between August 2011 and September 2015, only 22 whistleblowers received a financial reward (U.S. Securities and Exchange Commission 2015). As a result, both academics and practitioners warn that a significant part of these tips may be frivolous by nature. We define frivolous tips as whistleblowing tips which are (1) inaccurate and (2) opportunistic. However, evidence on whether or not financial rewards can act as a 'Trojan horse', facilitating frivolous whistleblowing by incentivizing opportunistic individuals to take undue advantage of the whistleblowing system is, at best, mixed. Specifically, whereas some authors conclude that there is no evidence that stronger monetary incentives result in more frivolous whistleblowing (e.g., Dyck, Morse and Zingales 2010), others highlight frivolous whistleblowing as a potential side effect of financial rewards (e.g., Andon et al. 2018). Responding to the call to conduct more research on this ambiguity (Lee and Xiao 2018; Berger and Lee 2022), we investigate whether, and if so, when, a financial reward incentivizes frivolous internal whistleblowing. Moreover, this study explores whether expanding the incentive system with a financial penalty for inaccurate whistleblowing can be an effective solution for frivolous internal whistleblowing (as proposed by Andon et al. 2018).

We rely upon a revised version of the original Fraud Triangle (Cressey 1973; Dorminey et al. 2012) to align mixed results in prior literature on whether or not financial rewards incentivize frivolous whistleblowing. This theoretical framework is well-suited to study opportunistic behavior of solo offenders (Morales, Gendron and Guénin-Paracini 2014; Free and Murphy 2015). The revised Fraud Triangle identifies the four drivers of opportunistic behavior: opportunity, pressure/motivation (e.g., money), personal integrity, and rationalization. While the first three drivers have been regularly accounted for in prior literature, rationalization has been consistently omitted. More specifically, when individuals act opportunistically, they experience psychological discomfort. However, when individuals can rationalize their opportunistic act, they are able to frame the behavior in such a way that it seems acceptable to themselves (Murphy and Dacin 2011; Sloane 1944). According to the revised Fraud Triangle, individuals will only pursue the opportunistic act if they can rationalize it. In line with this, we argue that financial rewards can incentivize frivolous whistleblowing when rationalization is present, but not when rationalization is absent. Importantly, we identify a potential solution and argue that adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing will deter frivolous whistleblowing when rationalization is present. By introducing a penalty for inaccurate whistleblowing, which triggers feelings of loss aversion (Kahneman and Tversky 1984), individuals will become more reluctant to submit frivolous tips. In particular, these feelings of loss aversion will lead individuals to be more skeptic about the (lack of) evidence and, hence, the likelihood of receiving a reward or penalty. In turn, this will disincentivize frivolous whistleblowing.

To test our hypotheses, we performed a scenario-based online experiment with a 3 (financial incentive) × 2 (rationalization) between-subjects factorial design. Participants were informed that they are financial consultants and that it is their job to draft a budget proposal for a new project

together with a hypothetical colleague. Participants are informed about company guidelines to prevent deliberate budget misstatements. Participants are asked to review the entire budget and, if they believe this to be necessary, blow the whistle on their colleague. Importantly, we did not provide participants any indication or evidence at all that a budget would be misstated. We argue that the absence of indicative evidence in our experiment makes any whistleblowing allegation by definition frivolous (i.e., (1) inaccurate and (2) opportunistic). As hypothesized, we find that a financial reward incentivizes frivolous whistleblowing when rationalization is present, but not when rationalization is absent. Second, as predicted, we find that adding a penalty for inaccurate whistleblowing acts as an effective deterrent of frivolous whistleblowing. Furthermore, we find process evidence for the theoretical argumentation of both hypotheses.

By examining whether financial incentives impact frivolous internal whistleblowing, we do not only contribute to prior literature focusing on the quantity of submitted whistleblowing tips (e.g., Chen et al. 2017; Brink, Lowe and Victoravich 2013), but also add a new dimension to the discussion by investigating the quality of these tips. Given that audit fees go up significantly regardless of the substance of a whistleblowing tip (Kuang et al. 2021), it is crucial for organizations to strike an effective balance between the quantity and quality of submitted tips. The necessity for more research on this causal relationship between financial incentives and frivolous whistleblowing has also been highlighted through recent calls (e.g., Smaili and Arroyo 2019; Lee et al. 2018). Relying upon the revised Fraud Triangle, which is well-suited to study solo offenders' opportunistic behavior (Morales et al. 2014; Free et al. 2015), we identify rationalization as an important, but consistently omitted confounding variable. Specifically, our process evidence shows that the increase in frivolous tips is driven by opportunistic individuals who can frame an act of frivolous whistleblowing as acceptable to themselves. This empirical test contributes to whistleblowing literature by aligning opposite views on whether or not financial rewards

incentivize frivolous whistleblowing. In particular, we show that the question should not be 'does it or does it not happen', but rather 'when does it happen'.

Second, we add to literature at the intersection of budgeting, as one of the core areas of management accounting and control (e.g., Libby and Lindsay 2010; Hannan, Rankin and Towry 2010; Maussen, Cardinaels and Hoozée 2024), and literature on how management control system design influences (un-)ethical decision-making (e.g., Towry 2003; Zhang 2008). While employees may misstate their budgets, which results in unfair advantages due to the misallocation of resources (e.g., Brunner and Ostermaier 2019), prior literature has confirmed that this issue can be addressed by introducing financial whistleblowing rewards (e.g., Zhang 2008). However, the existence of undesirable side effects in terms of frivolous whistleblowing has been neglected (e.g., Andon et al. 2018; Lee et al. 2018; Berger et al. 2022). We close this gap and offer the first empirical test of the impact of financial rewards and penalties on frivolous whistleblowing.

Third and finally, our results have important practical implications. Given that the increase in frivolous whistleblowing is driven by individuals' opportunism, rather than their intention to accurately blow the whistle, we are able to propose an effective solution. Specifically, adding a penalty for inaccurate whistleblowing disincentivizes frivolous whistleblowing by making individuals more skeptic with regard to the (lack of) gathered evidence. While previous literature highlighted that organizations already adopt penalties in order to increase the number of submitted whistleblowing tips (Chen et al. 2017), our results show that a dual focus on both the quantity and quality of these tips is pressing. Yet, we are confident that, when financial whistleblowing incentives are carefully designed, concerns for frivolous whistleblowing can be adequately addressed.

2 | BACKGROUND AND HYPOTHESES DEVELOPMENT

2.1 | Basic setting

We focus on a participative budgeting setting where financial consultants (subordinates) are responsible for co-developing budget proposals for projects, which they submit to their shared superior. Since budget proposals are susceptible to subtle forms of rent extraction, which represent managerial dishonesty (e.g., Brown, Evans and Moser 2009; Evans et al. 2001; Zhang 2008; Maussen et al. 2024), this is a relevant setting to study internal whistleblowing behavior. More specifically, when co-developing budget proposals, employees observe each other's reporting behavior (i.e., mutual monitoring), and hence, it is likely that budget misstatements can be observed and reported (i.e., internal whistleblowing). We are particularly interested in how financial incentives, as a management control practice, impact opportunistic behavior in terms of frivolous whistleblowing. Importantly, in our setting, there is no indicative evidence to believe that the budget is deliberately misstated. Rather, the setting requires the estimation of overhead costs and as such, inherently allows for the existence of deliberate budget misstatements (e.g., budget slack[1]).

2.2 | Rewards for whistleblowing

Employees are often reluctant to raise their voice when observing misconduct by colleagues (European Commission 2017). Prior literature has responded to this and consistently confirmed that financial rewards are a more effective solution to address non-reporting behavior as compared to non-financial incentives (e.g., code of conduct; Pope and Lee 2013) or no incentives at all (e.g., Berger, Perreault and Wainberg 2017; Brink et al. 2013; Xu and Ziegenfuss 2008; Butler et al. 2020). Berger et al. (2022) provide evidence from practice on the effectiveness of financial rewards

[1] Budget slack, or more generally rent extraction, is a more subtle form of misreporting and refers to a situation where employees can join forces to game budgeting processes (Maas and Yin 2022). Budget slack is considered an euphemistic term for lying in budget requests (Endenich and Trapp 2020; Brunner et al. 2019). Although budget slack might be beneficial in some settings (Davila and Wouters 2005), we focus on a setting in which it is dysfunctional and, therefore, explicitly state in our experiment that it is harmful for the organization in that it allocates resources inefficiently.

by showing that the SEC's whistleblower program (Dodd-Frank Act 2010) has received an increasing number of whistleblowing tips over the past decade (i.e., from 3,001 in 2012 up to 12,210 in 2021). Given the effectiveness of the program in collecting whistleblowing tips, it is remarkable that between August 2011 and September 2015, although 14,116 tips were received, only 22 whistleblowers received a financial reward (totaling 54 million USD; U.S. Securities and Exchange Commission 2015). More recently, the SEC has barred two individuals from their whistleblower award program because each of them submitted hundreds of frivolous award applications (U.S. Securities and Exchange Commission 2021). While some argue that there is no reason for skepticism (Dyck et al. 2010), others have raised concerns with regard to the quality of submitted tips. In particular, recent literature has argued that financial rewards may incentivize frivolous whistleblowing (e.g., Andon et al. 2018; Lee et al. 2018; Berger et al. 2022).

Frivolous whistleblowing tips are undesirable both from a public policy as well as organizational point of view. From a policy perspective, handling frivolous claims is costly since they consume considerable staff time and resources, hinder the efficient operation of the program and do not contribute to any successful enforcement action (U.S. Securities and Exchange Commission 2021). From an organizational point of view, Kuang et al. (2021) show that these frivolous tips are costly in that audit fees go up with, on average, $156,159 compared to the median. These concerns may have resulted in an organizational reluctance to implement financial whistleblowing rewards. In particular, the Association of Certified Fraud Examiners (ACFE, 2008, 2022) shows that although organizations' overall use of financial whistleblowing rewards has tripled over the last decade (from 5.4% in 2008 to 15% in 2022), it is still very limited. We respond to this ambiguity between, on the one hand, academic evidence on the effectiveness of financial rewards in increasing the number of submitted whistleblowing tips and, on the other hand, concerns by academics and practitioners with regard to the quality of these tips.

2.3 | Theoretical framework for frivolous whistleblowing

In this study, we investigate whether financial rewards can act as a 'Trojan horse', facilitating frivolous whistleblowing by incentivizing opportunistic individuals to take undue advantage of the whistleblowing system. In order to understand what drives this opportunistic behavior, and to find an effective solution to address the problem, we rely upon a revised version of the Fraud Triangle (Dorminey et al. 2012).[2]

Since its inception, Cressey's Fraud Triangle (1973) has been criticized (e.g., Kassem and Higson 2012; Murphy 2012; Carcello and Hermanson 2008; Hogan et al. 2008; Wells 2004; Wolfe and Hermanson 2004; Albrecht, Howe and Romney 1984). Therefore, several authors have revised the original Fraud Triangle and further refined the model. Combining these insights, Dorminey et al. (2012) have introduced the 'Fully Ascribed Meta-Model of White-Collar Crime' (cf. Appendix S1[3]) which consists of three pillars: characteristics of the individual (i.e., revised Fraud Triangle), characteristics of the fraudulent act, and contextual factors. In this study, we rely upon the revised Fraud Triangle given its focus on the psychology of solo offenders (Free et al. 2015; Morales et al. 2014). The revised Fraud Triangle identifies the fundamental drivers of opportunistic behavior: opportunity (including an individual's capabilities), pressure/motivation, rationalization, and personal integrity.

First, a wrongdoer (i.e., frivolous whistleblower) needs to have the perception that there is an opportunity to engage in the opportunistic act without getting detected (Wolfe et al. 2004). Stated differently, there has to be a control weakness in the system which the wrongdoer could exploit. Second, the perceived pressure creates the motive for the opportunistic act. According to the

[2] We acknowledge the existence of the Whistleblowing Triangle (e.g., Latan, Chiappetta Jabbour and Lopes de Sousa Jabbour 2019, 2021; Smaili et al. 2019). However, as we argue throughout the study, we aim to investigate whether financial rewards incentivize individuals to act opportunistically and to exploit the whistleblowing system by submitting frivolous tips solely to gain financial benefits. In order to test this, we need insights on what drives opportunistic behavior, rather than what drives people's good intentions to submit a whistleblowing tip which happens to be inaccurate, but not opportunistically.

[3] See Appendix S1 in the Supporting Information.

M.I.C.E. model, a set of motivations including money, ideology, coercion and ego or entitlement might introduce this pressure. Third, the wrongdoer should be able to rationalize the opportunistic behavior. In particular, a wrongdoer looks for a morally defensible excuse prior to acting opportunistically (Murphy et al. 2011).[4] When wrongdoers find this morally defensible excuse, the opportunistic act seems more acceptable to themselves. Murphy et al. (2011) identify seven categories of rationalization: (1) moral justification, (2) advantageous comparison, (3) euphemistic labeling, (4) minimize, ignore, or misconstrue consequences of the act, (5) denial of the victim, (6) displacing responsibility, and (7) diffuse responsibility. When the context allows for at least one of these rationalization categories, individuals can morally disengage from their opportunistic behavior (Moore 2015). Fourth and finally, personal integrity reflects the extent to which an individual is able to rationalize opportunistic behavior in general (Dorminey et al. 2012).

Applying the revised Fraud Triangle to our basic setting, we identify the different drivers of frivolous whistleblowing. First, opportunity is provided by the ambiguity around the existence of a deliberate budget misstatement (i.e., there is moral wiggle room). Furthermore, opportunity is also provided by accounting for recent developments in anti-retaliation legislation (e.g., Dodd-Frank Act 2010; Directive (EU) 2019/1937 2019). Anti-retaliation legislation finds its origin in the response of legislators to one of the major reasons to not blow the whistle, i.e., fear of retaliation (e.g., Curtis et al. 2021; Spoelma, Chawla and Ellis 2021). While these directives are intended to protect whistleblowers, they also inherently allow opportunistic individuals to avoid disciplinary sanctions from the organization when submitting frivolous tips.

[4] Importantly, rationalization as part of the Whistleblowing Triangle is defined as a prosocial behavior (Latan et al. 2019, 2021). Given that we argue that frivolous whistleblowing is an opportunistic act, we argue that rationalizing such an act is more closely related to the definition provided in the Fraud Triangle (i.e., a morally defensible excuse). Therefore, we are confident that the revised Fraud Triangle is most appropriate to study frivolous whistleblowing.

Second, monetary motivation (i.e., pressure or incentive) is present through the manipulation of financial whistleblowing incentive.[5] Contrary to the first two pillars, the third construct, rationalization, has been largely neglected in prior research. In this study however, we explicitly manipulate rationalization and test its moderating effect. Fourth and finally, our setting allows personal integrity to impact frivolous whistleblowing for which we control through the measurement of personality traits (i.e., the Dark Triad of Personalities; Jonason and Webster 2010) as well as social value orientation (Van Lange et al. 2007).

2.4 | Hypotheses development

Relying upon the revised Fraud Triangle (Dorminey et al. 2012), we argue that simply providing a reward for accurate whistleblowing will not incentivize frivolous whistleblowing. This is because an important driver of opportunistic behavior is missing. In particular, when individuals pursue an opportunistic act, they need to be able to frame the act as acceptable to themselves. Only when people find this morally defensible excuse (i.e., rationalization), and hence, can frame the act of frivolous whistleblowing as acceptable to themselves, the revised Fraud Triangle implies that the introduction of a reward for accurate whistleblowing can incentivize frivolous whistleblowing. Hence, we argue that financial rewards will incentivize frivolous whistleblowing when rationalization is present, but not when rationalization is absent.[6, 7]

[5] It is important to admit that other types of motivation might also induce frivolous whistleblowing according to the M.I.C.E. model. Brüggen and Luft (2011) for example show that an increase in competitiveness can stimulate employees to report dishonestly. Although interesting, the core objective of this study is to look at whether financial rewards incentivize frivolous whistleblowing. In order to answer this question unambiguously, it is necessary to isolate monetary motivation from other types of motivation in line with the M.I.C.E. motivations. Therefore, the baseline scenario used in this study will not induce any motivation which could be categorized as ideology, coercion, and ego or entitlement.

[6] An alternative way to approach our research question would be to rely upon the Psychological Pathways to Fraud model (Murphy et al. 2011). They argue that individuals will use reasoning in order to analyze whether the potential benefits outweigh the costs of acting opportunistically. Only if this cost-benefit analysis is in their advantage, individuals will behave opportunistically, experience negative affect and try to reduce this negative affect by, for example, rationalizing their behavior. A similar reasoning can be applied to the development of hypothesis 2.

[7] There is one exception to this reasoning: the fraud predator (Wolfe et al. 2004). Looking at the Fraud Triangle from a fraud predator's point of view, pressure and rationalization are replaced by a criminal mindset and arrogance. In other words, the predator only needs opportunity. We expect these fraud predators to be equally present in all experimental conditions, such that they do not impact the direction of our effects.

Hypothesis 1 (H1). Financial rewards for accurate whistleblowing incentivize frivolous whistleblowing when rationalization is present, but not when rationalization is absent.

In hypothesis 1, we argue that financial rewards can unintentionally incentivize frivolous whistleblowing when all fundamental drivers of opportunistic behavior are present. Responding to this undesirable side effect, we are interested in how organizations can limit frivolous whistleblowing when offering employees a reward for accurate whistleblowing. In line with Andon et al. (2018), we argue that adding a penalty for inaccurate whistleblowing may be an effective control mechanism to disincentivize frivolous whistleblowing.[8]

In particular, when introducing a penalty for inaccurate whistleblowing, the financial motive to act opportunistically disappears. Since we are studying a setting in which there is no indicative evidence of a deliberate budget misstatement, the outcome of whistleblowing in terms of financial benefits and costs is highly uncertain. More specifically, in terms of benefits, the reward will only be obtained when accurately blowing the whistle, and, in terms of costs, the penalty will be allocated when inaccurately blowing the whistle. We argue that adding a penalty for inaccurate whistleblowing will increase individuals' feelings of loss aversion (Kahneman et al. 1984) and thereby, will increase their skepticism related to the (lack of) evidence. In turn, this will make individuals reluctant to submit frivolous whistleblowing tips. Importantly, we only expect feelings of loss aversion to influence evidence skepticism when rationalization is present, since we only expect frivolous whistleblowing to occur in that situation. In summary, adding a penalty for inaccurate whistleblowing will disincentivize frivolous whistleblowing when rationalization is present, but not when rationalization is absent.

[8] Based on this theoretical argumentation, we choose to only study a financial incentive which includes a penalty for inaccurate whistleblowing when there is also a reward for accurate whistleblowing, since we only expect an increase in frivolous whistleblowing when a financial reward is present.

Hypothesis 2 (H2). Adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing decreases the number of frivolous whistleblowing tips when rationalization is present, but not when rationalization is absent.

Jointly, our two hypotheses predict that although a financial reward incentivizes frivolous whistleblowing when rationalization is present, importantly, adding a penalty for inaccurate whistleblowing will offset this undesirable side effect. Figure 1 provides a graphical display of our predictions.

[Insert Figure 1]

3 | METHODOLOGY

We employ a 3 (financial incentive) × 2 (rationalization) between-subjects scenario-based online experiment using Qualtrics to investigate the effects of *financial incentive* and *rationalization* on *frivolous whistleblowing* by employees in a participative budgeting setting. Within this context, participants may blow the whistle when they suspect a deliberate misstatement (i.e., budget slack) to be present in the proposed budget of their (hypothetical) colleague. Participants were randomly assigned to an experimental condition. Prior to running the experiment, we executed several quality-assuring activities to increase the internal validity of our experiment, as well as to assure mundane realism.[9]

3.1 | Experimental procedures

Before starting the experiment, participants received general information on the experimental task, minimum payoff, experimental currency exchange rate, and duration of the experiment. After

[9] First, two academic colleagues and 52 graduate business students at a large European university pretested the full experimental instrument. Based on this, we changed two features. First, our operationalization of rationalization was changed to get a better fit with the definitions as provided by Murphy et al. (2011). Second, the scenario of the pre-test included a degree of competitiveness with regard to budget acceptance, providing all participants a motivation to frivolously blow the whistle. For the purpose of testing a clean relation between our motivation of interest (i.e., financial motivation) and frivolous whistleblowing, we excluded the competitiveness. This new scenario was tested with 11 academic colleagues and two practitioners, with multiple years of experience in either the industrial sales sector or the financial audit sector, who checked the scenario's mundane realism. This final experimental instrument was proofread by a native English speaker.

this general introduction, participants had to indicate whether or not they agreed with the prescribed informed consent form. Only if they agreed, they could start the experiment.

Next, participants read the scenario that consists of five main parts (cf. Appendix S2[10]). Part one informs participants that they take the role of a financial consultant and are instructed to draft a two-piece budget proposal for a new project together with a hypothetical colleague. Participants themselves take responsibility for the part on direct costs and their hypothetical colleague estimates overhead costs. Since participants have to give their opinion on the entire budget before submitting, they also receive information on company guidelines which are in place to prevent deliberate budget misstatements. In particular, we inform participants that there is a formal guideline which indicates that overhead costs should be estimated at approximately 25% of direct costs, but never exceed 50%.

In the second part of the scenario, participants are informed about their fixed salary (45,000 Lira[11]) and their performance-based bonus (30,000 Lira). This bonus is an essential design choice in our experimental scenario which ensures external validity and enables us to introduce a financial penalty for inaccurate whistleblowing. In particular, employees can only be punished financially when part of their compensation is performance contingent (cf. Directive (EU) 2019/1937 2019). Hence, participants in our experimental study can only lose their performance-based bonus, but are assured to receive a fixed minimum wage which is independent of their performance.

In part three of the scenario, we provide participants with information on their hypothetical colleague's past working performance. We tell them explicitly that their colleague is in a bad financial situation. Since we also tell participants that whistleblowing might result in dismissal of the hypothetical colleague, we argue that this introduces a psychological cost. Prior literature has

[10] See Appendix S2 in the Supporting Information.
[11] We used Lira as an experimental currency that was converted to GBP (£) at the rate of 10,000 Lira = £1.00 (GBP).

emphasized that whistleblowers take into account such emotional relations. This ensures both internal and external validity.[12]

In part four and part five of the scenario, we implemented the manipulations as described in the next section. Part four informs participants about their organization's whistleblowing policy, including the manipulation of *financial incentive*. In part five, participants attend a meeting with their team leader in which the presence of *rationalization* is manipulated.

Once participants completed a quiz to make sure they understood everything correctly, they proceeded to the experimental task. In this task, participants are provided with the estimate for overhead costs by their hypothetical colleague, as well as the formal company guideline. The company guideline states that overhead costs should be estimated at around 25% of direct costs, but never exceed 50%. Based upon this, participants are asked to review their colleague's estimate for the overhead costs and indicate one of two choices: (1) agree with the budget or (2) submit a whistleblowing tip to the project leader indicating that they believe their colleague misstated the budget. This experimental task is repeated five times. Each round consisted of the same task, but provided participants with different estimates for the amount of direct costs and estimated amount of overhead costs.[13] In particular, overhead costs were estimated between 27.50% and 35% of direct costs across the five independent rounds. We randomized the order in which participants received the different budgets. Importantly, since the highest estimate of overhead costs as compared to direct costs equaled 35%, which is far below the maximum threshold of 50%, we are confident that, when participants' whistleblowing behavior differs across experimental conditions,

[12] Although this was specifically introduced with the aim of introducing a psychological cost which discourages whistleblowing, we acknowledge that this might also have provided participants with an indication that they are working with an unreliable person. We thank an anonymous conference reviewer for this comment. We did, however, not anticipate such reasoning. We do argue that, if this would have a discouraging or encouraging effect, it would be randomized across conditions (i.e., it is not part of a manipulation) and, therefore, not influence the direction of our effects.

[13] (Overhead costs ; direct costs) per round were respectively: round one (206,250 ; 750,000), round two (600,000 ; 2,000,000), round three (568,750 ; 1,750,000), round four (812,500 ; 2,500,000) and round five (420,000 ; 1,200,000).

this should be interpreted as opportunistic (i.e., frivolous) behavior driven by our manipulations. We also empirically test this assumption by checking whether whistleblowers in our sample were aware that the evidence was not indicative of misreporting. In particular, we compare whistleblowers' (i.e., participants who blew the whistle at least once across the five rounds) answer to a question regarding *evidence skepticism*[14] to the answer of those who never blew the whistle. Whistleblowers were significantly less concerned with regard to the quality of the evidence as compared to those who did not blow the whistle ($F = 14.49$, $p < 0.001$). This confirms the idea that whistleblowing in our experiment is driven by people's opportunism, and not by their intention to accurately blow the whistle.[15]

As a final step in this experiment, participants responded to attention checks, manipulation checks, questions related to process and control variables, and demographical items in the post-experimental questionnaire (PEQ) and learned their final payoff. Consistent with prior studies, one whistleblowing round was randomly selected as payment round and only the payoffs from this round were paid out (e.g., Hannan, Rankin and Towry 2006; Hannan et al. 2010; Maussen et al. 2024).

3.2 | Experimental factors

The first manipulated factor is *financial incentive*. In the 'no incentive' condition, participants only received information on the organization's anti-retaliation policy which states that they are assured a fixed minimum wage and that there will be no reputational damage after blowing the

[14] The question with regard to *evidence skepticism* that was included in the PEQ: "I would only blow the whistle if I am 100% sure that the budget includes budgetary slack." (1 = *completely disagree*, 7 = *completely agree*).

[15] Adding to this empirical test, the theoretical argumentation for hypothesis 1 and 2 implies that our manipulation of *rationalization* should not influence people who truly believe their whistleblowing tip is accurate. In particular, rationalization is only necessary when people are considering to act opportunistically. When people believe that their whistleblowing tip is accurate, there should be no impact of our manipulation of *rationalization* on their behavior (i.e., this is what differentiates the Fraud Triangle from the Whistleblowing Triangle). As such, when people believe their whistleblowing tip is accurate, even in this setting where there is no indicative evidence, we should see a similar pattern of behavior regardless of whether *rationalization* is present or absent. When this is not the case, we are confident that this difference in behavior is driven by people who are acting opportunistically.

whistle, which is in line with the specific forms of protection included in Directive 2019/1937 (Chapter VI, Article 19, p. 305/43). Participants in the 'financial reward' condition learned that, in addition to the anti-retaliation policy, there is a whistleblower award program present at their company. This means that whenever a whistleblower's tip results in the successful discovery of budget slack, the whistleblower will be granted a fixed reward (30,000 Lira). Participants in the 'financial reward and penalty' condition received the same information as those participants in the 'financial reward' condition, but were briefed that on top of a reward for accurate whistleblowing, their whistleblower award program also includes a penalty for inaccurate whistleblowing. In particular, if participants submit an inaccurate whistleblowing tip, they will receive a fixed penalty (30,000 Lira). Both the reward and penalty are identical in absolute amounts (i.e., 40% of participants' base pay[16]). Importantly all participants could obtain the same financial payoff.

With regard to our second manipulation, half of the participants attended a meeting with their supervisor in which five elements were presented which they could use to make an act of frivolous whistleblowing seem more acceptable to themselves (i.e., morally disengage). We prefer to operationalize *rationalization* as a multidimensional construct rather than relying upon one specific element because this gives participants a limited degree of discretion as to how they personally want to rationalize an act of frivolous whistleblowing. We focus on the five most frequently used categories of rationalization (Murphy 2012) out of the seven categories as defined by Murphy et al. (2011): (1) moral justification, (2) advantageous comparison, (3) euphemistic labeling, (4) denial of the victim, and (5) displacing responsibility. Table 1 shows how these categories were applied to our setting in the description of the experimental scenario.

[Insert Table 1]

[16] $\frac{30{,}000\ Lira\ (fixed\ bonus\ or\ penalty)}{45{,}000\ Lira\ (fixed\ wage) + 30{,}000\ Lira\ (bonus\ for\ good\ performance)} = 40\%$.

3.3 | Dependent Variable

Our dependent variable is the participant's internal whistleblowing decision, which is coded as one if a participant decides to submit a whistleblowing tip and zero otherwise. Different from several prior studies relying on continuous scales to measure whistleblowing intentions (e.g., Boo, Ng and Shankar 2021; Andon et al. 2018; Barr-Pulliam 2019; Berger et al. 2017; Robinson, Robertson and Curtis 2012), we chose for a dichotomous dependent variable, similar to Chen et al. (2017), to measure whistleblowing behavior. Importantly, in practice, whistleblowing requires observers of misconduct to take action. Therefore, it is impossible to determine from which point on a continuous scale participants would transfer from being a non-reporting observer to being a whistleblower. As mentioned before, we measure whistleblowing behavior in five independent rounds. As such, if a participant blew the whistle in all five rounds, the dependent variable takes the value five, while if the participant never blew the whistle, the dependent variable equals zero. We further refer to our dependent variable as *frivolous whistleblowing*.

3.4 | Participants

We recruited participants through the online panel Prolific Academic.[17,18] Online participants have proven to make up a reliable sample when studying ethical decision-making (Farrell, Grenier and Leiby 2017). To further improve the reliability of our sample, we follow the best practices when recruiting participants online outlined by Bentley (2021). First, Prolific allows researchers to predefine criteria for sample selection based on demographic details about the participant pool. We restricted participation in our experiment to the following conditions: (1) English as a first language, (2) residence in U.K. or U.S., (3) age between 18 and 65, (4) full-time or part-time

[17] Prolific is a readily accessible platform that connects researchers to participants in return for a fee. Data collected on Prolific Academic is of higher reliability and quality when compared to other online sources and/or institutional participant pools (Peer et al. 2017).
[18] The experiment was approved by the ethics board of the authors' university.

employment status and (5) employment sector is either (a) Business Management & Administration, (b) Finance or (c) Marketing & Sales. Requirements one and two were important since this is a scenario-based experiment which requires a substantial amount of reading. Furthermore, all requirements together make sure that participants have the adequate upfront knowledge necessary to effectively complete the experiment to ensure internal validity (Libby, Bloomfield and Nelson 2002).

In total, we recruited 201 participants through Prolific. Following Bentley (2021), we screened our sample for the number of total approvals on Prolific. Based upon a boxplot, we excluded 15 outliers who had substantially more approvals on Prolific compared to the remaining sample (i.e., 2,115 or more).[19] This is because highly productive participants may be less naïve with respect to research (Bentley 2021), harming the internal validity of the study. Third, to ensure that participants read the instructions carefully, we included a Natural Language Attention-Check Question. Only participants who correctly indicated that whistleblowers are protected from retaliatory activities (e.g., lay-off, withholding of promotion or negative performance assessments), were included in the final sample. This resulted in the exclusion of 16 participants. Importantly, this attention check question also served a second purpose. A crucial element of our theory, and basic setting, is that participants are aware of the fact that there is an opportunity to submit frivolous whistleblowing tips while simultaneously avoid disciplinary sanctions by the organization. The combination of these checks makes sure that we end up with a high-quality sample (cf. Peer, Vosgerau and Acquisti 2014).

[19] This intervention increases researcher degrees of freedom (Bentley 2021). We come back to this decision in the Robustness checks section.

The remaining sample includes 170 observations.[20] 47% of our participants are female. 149 participants (88%) are British while the other 21 (12%) are American. Participants are on average 36 years old (ranging from 19 up to 64 years old). 74% of our participants have at least a bachelor's degree. Mean work experience among participants is 16 years (ranging from one up to 49 years). 97 participants (57%) work in Finance, 39 participants (23%) work in Business Management & Administration and 34 participants (20%) work in Marketing & Sales. The majority (83%) works full-time. Finally, while 52% of the participants have participated in at least five experiments, 74% has taken at least one experiment in the past. Participants received a bench fee of £7.50 (GBP) for participation. However, depending upon their whistleblowing behavior, this fee could be decreased to £4.50. On average, participants earned £7.13. More specifically, 149 participants received £7.50, while 21 participants received £4.50.

Randomization was largely successful. No significant differences across experimental conditions were found for total approvals, gender, nationality, academic degree, employment sector, and employment status (all $p \geq 0.150$). Furthermore, also participants' score on validated scales such as people's ability to morally disengage from (general) undesirable behavior (Moore et al. 2012), Dark Triad of personality (Jonason et al. 2010) and social value orientation (Van Lange et al. 2007) did not differ significantly across experimental conditions (all $p \geq 0.281$). However, we did find significant differences across conditions for age ($p = 0.011$), work experience[21] ($p = 0.050$), and the level of endowment effect[22] ($p = 0.069$). Hence, we will include these variables as covariates in a robustness check.

[20] Participants who were excluded did not differ significantly (all $p > 0.446$) from included participants based on *moral disengagement* (Moore et al. 2012) and *Dark Triad of personality* (Jonason et al. 2010). Participants who were excluded scored, however, significantly lower on the scale of *social value orientation* ($p = 0.025$; Van Lange et al. 2007).

[21] Age and work experience are highly correlated with one another ($r = 0.94$, $p < 0.001$). When including covariates into our analyses (see the Robustness checks section), we choose to only include work experience to avoid multicollinearity. Including age instead of work experience as a covariate does not change our reported results inferentially.

[22] The endowment effect is an alternative construct to measure individuals' general disutility from experiencing a loss, which is closely related to loss aversion. In particular, the endowment effect implies that individuals attach a higher value to what they

To test whether our manipulations were successful, we included three manipulation check questions in our post-experimental questionnaire (Table 2). In total, 37 participants (22%) failed at least one manipulation check. Our results remain qualitatively and inferentially the same when excluding these participants (cf. Robustness checks section).

[Insert Table 2]

4 | RESULTS

4.1 | Descriptives

Table 3 reports the descriptive statistics of the dependent variable. Across all experimental conditions, participants have submitted on average 2.22 (*S.D.* = 1.74) whistleblowing tips over the five independent rounds. When rationalization is present, a reward for accurate whistleblowing (*M* = 3.24, *S.D.* = 1.53) seems to increase the number of submitted whistleblowing tips as compared to a situation with no incentive (*M* = 2.33, *S.D.* = 1.71) or both a reward for accurate and penalty for inaccurate whistleblowing (*M* = 2.38, *S.D.* = 1.69). However, this trend disappears almost completely when rationalization is absent. These findings are in line with our predictions.

[Insert Table 3]

4.2 | Hypotheses testing

4.2.1 | Hypothesis 1

Hypothesis 1 predicts that rewarding whistleblowers incentivizes frivolous whistleblowing when rationalization is present, but not when rationalization is absent. To test hypothesis 1, we only select participants in the following four experimental conditions: (1) no incentive, rationalization absent, (2) no incentive, rationalization present, (3) reward for accurate whistleblowing, rationalization absent, and (4) reward for accurate whistleblowing, rationalization

possess as compared to the same object if they would have to acquire it (Thaler 1980). It is measured as: $\frac{Willingness\ to\ sell - Willingness\ to\ pay}{Willingness\ to\ pay}$

present. We conducted a two-way ANOVA ($F_{(3, 114)} = 4.47$, $p = 0.005$) with *financial incentive* and *rationalization* as independent variables and *frivolous whistleblowing* as dependent variable (Table 4, Panel A). Results show a significant main effect of both *financial incentive* ($F = 3.23$, $p = 0.074$) and *rationalization* ($F = 8.96$, $p = 0.003$) and a non-significant interaction effect ($F = 1.33$, $p = 0.252$). Since we predict an ordinal interaction effect in H1, we perform contrast analysis and follow up with simple effect tests to formally test the hypothesis (Buckless and Ravenscroft 1990; Bobko 1986; Strube and Bobko 1989). We follow the three-step-procedure recommended by Guggenmos, Piercey and Agoglia (2018) to evaluate evidence of an ordinal interaction. First, we show visual evidence of fit by comparing Figure 2, which shows the observed marginal means for hypothesis 1 and 2, with the predicted pattern in Figure 1. Visually, the observed data follow the predicted pattern. Second, we perform a one-way ANOVA omnibus test with *experimental condition* (representing the four experimental conditions) as independent variable and *frivolous whistleblowing* as dependent variable. Panel B of Table 4 shows a significant one-way ANOVA ($F = 4.47$, $p = 0.005$), indicating that the number of frivolous whistleblowing tips is significantly different across the four experimental conditions. Given our theoretical prediction, we use contrast weights of -1, -1, -1, and +3 for condition (1), (2), (3), and (4) respectively to test the ordinal interaction effect (Table 4, Panel C). As shown in Panel D of Table 4, the hypothesized contrast is significant ($F = 11.65$, $p < 0.001$) and the residual between-cells variance test is non-significant ($F = 0.59$, $p = 0.624$). To provide further evidence for our hypothesis, we follow up on this significant ordinal interaction effect with simple effect tests in Panel E of Table 4. We can conclude that, as predicted, providing a reward for accurate whistleblowing does not incentivize *frivolous whistleblowing* when rationalization is absent ($F = 0.21$, $p = 0.649$), but does incentivize *frivolous whistleblowing* when rationalization is present ($F = 4.36$, $p = 0.039$). Furthermore, the impact of *rationalization* is not significant when there is no incentive ($F = 1.73$, $p = 0.192$), but is significant

when there is a reward for accurate whistleblowing (F = 8.45, p = 0.004). Third, we quantitatively evaluate the contrast variance residual (q^2), which is calculated as $1 - r^2$ where r represents the correlation between the predicted contrast and the corresponding observed cell means (Guggenmos et al. 2018). We can conclude that 12.67% (q^2) of the between-cells variance is not explained by the contrast. In sum, the contrast analysis and simple effect tests suggest that the observed data follow the predicted pattern of the ordinal interaction in hypothesis 1. Specifically, financial rewards for accurate whistleblowing increase the number of frivolous tips when rationalization is present, but not when rationalization is absent.

[Insert Table 4]

[Insert Figure 2]

4.2.2 | Hypothesis 2

Hypothesis 2 predicts that adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing disincentivizes frivolous whistleblowing when rationalization is present, but not when rationalization is absent. To test hypothesis 2, we only select participants in the following four experimental conditions: (1) reward for accurate whistleblowing, rationalization absent, (2) reward for accurate whistleblowing, rationalization present, (3) reward for accurate and penalty for inaccurate whistleblowing, rationalization absent, and (4) reward for accurate and penalty for inaccurate whistleblowing, rationalization present. The two-way ANOVA ($F_{(3, 106)}$ = 4.66, p = 0.004) in Table 5 Panel A shows significant main effects of *financial incentive* (F = 3.19, p = 0.077) and *rationalization* (F = 9.33, p = 0.003) and a non-significant interaction effect (F = 0.81, p = 0.372). Since we predict an ordinal interaction effect, we will focus on the planned contrast test and follow the same procedure as for testing H1. First, comparing Figure 1 and Figure 2 shows visual evidence of fit. Second, panel B of Table 5 shows a significant one-way ANOVA (F = 4.66, p = 0.004). Given our theoretical prediction, we use contrast weights of -1, +3, -1, -1 for

condition (1), (2), (3), and (4) respectively to test the ordinal interaction effect (see Table 5, Panel C). As shown in Panel D of Table 5, the hypothesized contrast is significant ($F = 11.41$, $p = 0.001$) and the residual between-cells variance test is non-significant ($F = 0.86$, $p = 0.465$). To provide further evidence for hypothesis 2, we follow up with simple effect tests in Panel E of Table 5. We conclude that, while adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing does not influence *frivolous whistleblowing* when rationalization is absent ($F = 0.41$, $p = 0.522$), it disincentivizes *frivolous whistleblowing* significantly when rationalization is present ($F = 3.46$, $p = 0.066$). Third, identical as to what was reported under the first hypothesis, *rationalization* has a strong effect when there is a reward for accurate whistleblowing ($F = 8.28$, $p = 0.005$). Fourth and finally, although unpredicted, there is an indication that *rationalization* influences *frivolous whistleblowing* even when there is both a reward for accurate and a penalty for inaccurate whistleblowing ($F = 2.20$, $p = 0.141$). We come back to this finding in our conclusion. Furthermore, we conclude that 17.90% (q^2) of the between-cells variance is not explained by the contrast. To summarize, the contrast analysis and simple effect tests suggest that the observed data follow the predicted pattern of the ordinal interaction for hypothesis 2. Specifically, adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing decreases the number of frivolous whistleblowing tips when rationalization is present, but not when rationalization is absent.

[Insert Table 5]

4.3 | Process tests

4.3.1 | Hypothesis 1

In the development of H1, we argued that financial rewards will only incentivize individuals to submit frivolous whistleblowing tips when they can frame this act as acceptable to themselves. To provide process evidence of H1, we test the model described in Figure 3 and selected participants

in the following four experimental conditions: (1) no incentive, rationalization absent, (2) no incentive, rationalization present, (3) reward for accurate whistleblowing, rationalization absent, and (4) reward for accurate whistleblowing, rationalization present. We used structural equations-based path analysis to estimate the causal model with R (Lavaan).

To measure the extent to which participants believed frivolous whistleblowing is acceptable (i.e., *moral disengagement*), participants reported in the PEQ to what extent they agreed (1 = *completely disagree*, 7 = *completely agree*) with statements included in Table 6. An exploratory factor analysis using the five Likert-type scale responses reveals that one factor (Eigenvalue = 2.25) explains 45.00% of the variance. Based on a confirmatory factor analysis, we allowed an error correlation between items 3 and 4. This measurement model results in a good fit (χ^2 = 6.27 and p = 0.180; CFI = 0.99; TLI = 0.98; RMSEA = 0.06; SRMR = 0.03) and its Cronbach's alpha amounts to 0.80.

Overall, the structural model shown in Figure 3 is a good fit for the data (χ^2 = 36.26, p = 0.166; CFI = 0.96; TLI = 0.94; RMSEA = 0.07; SRMR = 0.07). The model confirms that a financial reward for accurate whistleblowing incentivizes an individual to morally disengage, and hence to perceive frivolous whistleblowing to be more acceptable, when rationalization is present (β = 0.29, p = 0.035), but not when rationalization is absent (β = -0.08, p = 0.572). Furthermore, the indirect effect of *financial incentive* on *frivolous whistleblowing* through *moral disengagement* is (marginally) significant when rationalization is present (β = 0.06, p = 0.103), but not when rationalization is absent (β = -0.02, p = 0.581; Table 7, Panel A). We assessed this interaction following Rigdon, Schumacker and Wothke (1998). We compared an unrestricted model, where the paths from *financial incentive* to both *frivolous whistleblowing* and *moral disengagement* could vary across the different levels of *rationalization*, with a restricted model, where these paths were restricted to be the same under both levels of *rationalization*. The unrestricted model provides a

significantly better fit for the data than the restricted model (χ^2-difference = 4.67, p = 0.097), implying the interaction is significant. Overall, the model thus confirms a moderated mediation with *moral disengagement* as mediator and *rationalization* as moderator of the effect of *financial incentive* on *frivolous whistleblowing*.

[Insert Figure 3]

[Insert Table 6]

[Insert Table 7]

4.3.2 | Hypothesis 2

For hypothesis 2, we argued that adding a penalty for inaccurate whistleblowing will increase individuals' feelings of loss aversion. In turn, this will increase individuals' skepticism regarding the (lack of) evidence when rationalization is present, but not when rationalization is absent. Finally, the increase in evidence skepticism will decrease the number of frivolous tips. To test this underlying process, we built the causal model depicted in Figure 4.

To capture *loss aversion*, we adapted the scale developed by Li et al. (2021) to our specific setting. Participants reported in the PEQ whether they agreed (1 = *completely disagree*, 7 = *completely agree*) with statements included in Table 8. An exploratory factor analysis showed negative factor loadings for items four and seven after recoding them, which indicates participants did not pay attention to the reversed character of the items. Therefore, we had to exclude these items. The exploratory factor analysis using the remaining six items reveals that one factor (Eigenvalue = 2.06) explains 34.30% of the variance. This measurement model results in a good fit (χ^2 = 14.73 and p = 0.099; CFI = 0.97; TLI = 0.95; RMSEA = 0.06; SRMR = 0.05) and its Cronbach's alpha amounts to 0.74. Second, to measure participants' degree of *evidence skepticism*, they indicated to what extent they agreed with the following statement: 'I would only blow the

whistle if I am 100% sure that the budget includes budgetary slack.' (1 = *completely disagree*, 7 = *completely agree*).

Overall, the structural model depicted in Figure 4 shows a reasonable fit for the data (χ^2 = 86.75 and p = 0. 011; CFI = 0.85; TLI = 0.82; RMSEA = 0.09; SRMR = 0.11). The model confirms that the indirect effect (Table 9, Panel A) of *financial incentive* (i.e., adding a penalty for inaccurate whistleblowing) through *loss aversion* and *evidence skepticism* is significant when rationalization is present (β = -0.06, p = 0.035), but not when rationalization is absent (β = -0.01, p = 0.524). Total effects (Table 9, Panel B) confirm that the total effect of *financial incentive* on *frivolous whistleblowing* is significant when rationalization is present (β = -0.26, p = 0.038), but not when rationalization is absent (β = -0.07, p = 0.569). We assessed this interaction following Rigdon et al. (1998). We compared an unrestricted model, where the path from *financial incentive* to *frivolous whistleblowing* and from *loss aversion* to *evidence skepticism* could vary across the different levels of *rationalization*, with a restricted model, where the paths were restricted to be the same under both levels of *rationalization*. The unrestricted model provides a (marginally) significant better fit for the data than the restricted model (χ^2-difference = 4.45; p = 0.106), implying the interaction is significant and providing support for our underlying theory.

[Insert Figure 4]

[Insert Table 8]

[Insert Table 9]

4.4 | Robustness checks

In addition to the main analyses for our hypotheses, we perform four robustness checks (Table 10). First, as indicated before, personal integrity is a fourth pillar which, according to the revised Fraud Triangle, might influence individuals' degree to which they are willing to rationalize opportunistic behavior. Hence, the first model in Table 10 controls for Dark Triad of personalities

(i.e., *machiavellianism*, *psychopathy* and *narcissism*; Jonason et al. 2010) and *social value orientation* (Van Lange et al. 2007). Second, in our total sample (n = 201), we identified 15 outliers with substantially more *total approvals* on Prolific (i.e., ≥ 2,115). When including these 15 outliers in our sample (n = 185), randomization across conditions with regard to *total approvals* is unsuccessful (p = 0.097). Therefore, we include these 15 participants, but control for *total approvals* as a covariate in model two. Third, out of the 170 observations included in our sample, 37 participants failed at least one manipulation check. In model three, we exclude these participants and remain with 133 observations. Fourth and finally, since randomization was not completely successful in our initial sample of 170 observations, we include level of *endowment effect* and *work experience* as covariates in model four. All four models support our first hypothesis that a reward for accurate whistleblowing incentivizes *frivolous whistleblowing* when rationalization is present (all $p \leq 0.084$), but not when rationalization is absent (all $p \geq 0.538$). Similarly, all four models support our second hypothesis that adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing disincentivizes *frivolous whistleblowing* when rationalization is present (all $p \leq 0.079$), but not when rationalization is absent (all $p \geq 0.460$). Unexpectedly, we find mixed results with regard to the effect of *rationalization* given no incentive, which is significant in model two, and with regard to the effect of *rationalization* given a reward for accurate and penalty for inaccurate whistleblowing, which is (marginally) significant in model one and two. We address these unexpected findings in our conclusion.

[Insert Table 10]

5 | CONCLUSION

In search to improve the effectiveness of internal whistleblowing policies, there has been an increasing interest into the effectiveness of financial rewards for whistleblowers. Despite strong evidence that financial rewards effectively increase the number of submitted tips, both practitioners

and academics have concerns with regard to the quality of these tips (e.g., U.S. Securities and Exchange Commission 2021; Berger et al. 2022). We respond to this ambiguity by investigating when financial rewards can act as a 'Trojan horse', facilitating frivolous whistleblowing, and explore how organizations can adequately address these concerns.

Using a scenario-based online experiment in a participative budgeting setting, we find strong support for our hypothesis that rewarding whistleblowers increases the number of frivolous tips when rationalization is present, but not when rationalization is absent. Hence, we identify rationalization as an important moderator which may explain the inconsistent results in prior literature. Importantly, we also find strong support for our second hypothesis that adding a penalty for inaccurate whistleblowing decreases the number of frivolous tips when rationalization is present. Hence, our study introduces a penalty for inaccurate whistleblowing as an effective solution for the undesirable side effect introduced by rewarding accurate whistleblowing. Our findings have important theoretical contributions to the accounting literature and more specifically, ethical decision-making. We find that financial rewards do not necessarily incentivize frivolous whistleblowing and, when they do, that this negative side effect is driven by individuals' opportunism. Importantly, the introduction of a financial penalty for inaccurate whistleblowing will limit this opportunistic behavior by making whistleblowers more skeptic about the (lack of) gathered evidence, and hence, the likelihood of accurately blowing the whistle.

Our study is subject to some limitations which can be addressed in future research. First, since we conducted a scenario-based experiment, the whistleblower and the person on whom the whistle is blown do not directly interact with each other. As such, we were only able to indirectly induce psychological costs, through written text, for blowing the whistle. Related to this, while employees are protected from retaliatory activities from the organization in line with Directive (EU) 2019/1937 (2019; e.g., firing and/or demotion), it is still possible that frivolous whistleblowers will

encounter retaliatory activities from peers (e.g., Sommerfeldt 2023). An interactive lab experiment could solve both issues by allowing colleagues to interact with each other. Second, some analyses showed that rationalization can have a significant effect on frivolous whistleblowing given no incentive or given reward for accurate and penalty for inaccurate whistleblowing. Although we did not formulate a formal hypothesis on this, the revised Fraud Triangle cannot explain these results. Our post-hoc argumentation for these effects is that our operationalization of rationalization might have induced other types of motivation according to the M.I.C.E. model of motivations for certain participants, which was not part of our objective. Future research might look at how these other types of motivation (ideology, coercion, and ego or entitlement) can incentivize frivolous whistleblowing. It is especially interesting to study whether a penalty for inaccurate whistleblowing is still an effective tool to disincentivize frivolous whistleblowing in these situations.

Furthermore, we identify three more fruitful avenues for future research based upon our results. First, while the Fraud Triangle is an extremely useful tool to study solo offenders' opportunistic behavior (Free et al. 2015; Morales et al. 2014), its emphasis on individualistic explanations of fraud is to the detriment of the wider organizational and social context in which individuals operate in practice (Morales et al. 2014). A sociological approach to the problem of frivolous whistleblowing is also well-placed to investigate how colleagues' reactions to frivolous whistleblowing in the past influence (frivolous) whistleblowing behavior in the future. Related to this, participants in our experiment had to decide on whether or not to blow the whistle in five independent rounds. While it is less likely that employees are confronted with misconduct multiple times on a short period of time (i.e., situations where whistleblowing would be accurate), it is unclear to what extent frivolous whistleblowers will repeatedly act opportunistically. Third, whistleblowing tips in our experiment are by definition frivolous. However, for organizations, it might be more difficult to differentiate between frivolous whistleblowing tips (i.e., inaccurate and

opportunistic) and inaccurate whistleblowing tips submitted by individuals with good intentions. Future research on how organizations' internal whistleblowing policies can increase (1) the quality of whistleblowing tips, as well as (2) organizations' effectiveness in evaluating the received tips are therefore two important topics on the research agenda.

Despite the limitations, this study provides new practical insights for all people involved in designing effective internal whistleblowing policies and, in particular, management accountants. Management accountants are often responsible to (co-)develop management control systems (MCS) of which an incentive system is part. Despite academic evidence on the effectiveness of financial rewards in stimulating whistleblowing, organizations remain reluctant to include them in their MCS (i.e., only 15% of the organizations has a financial reward for whistleblowers; ACFE 2022). One of the reasons for this may be the concern that financial rewards incentivize frivolous whistleblowing, and handling frivolous whistleblowing tips is costly for organizations (Kuang et al. 2021). Responding to this, we show that introducing a penalty for inaccurate whistleblowing can solve this unintended side effect of financial rewards. By doing so, we add to an extensive body of research promoting the use of financial whistleblowing rewards. While we, on the one hand, highlight an undesirable side effect of financial rewards, we, on the other hand, also offer an easy-to-implement solution to organizations. This way, organizations should feel more comfortable to reward accurate whistleblowing, without fearing an increase in frivolous whistleblowing.

REFERENCES

Albrecht, W. S., Howe, K. R., and Romney, M. B. (1984). *Deterring fraud: The internal auditor's perspective.* Institute of Internal Auditors Research Foundation.

Andon, P., Free, C., Jidin, R., Monroe, G. S., and Turner, M. J. (2018). The impact of financial incentives and perceptions of seriousness on whistleblowing intention. *Journal of Business Ethics, 151*, 165-178.

Association of Certified Fraud Examiners. (2008). 2008 Report to the nation on occupational fraud & abuse, https://www.acfe.com/-/media/files/acfe/pdfs/2008-rttn.pdf

Association of Certified Fraud Examiners. (2022). Report to the nations: 2022 global study on occupational fraud and abuse, https://acfepublic.s3.us-west-2.amazonaws.com/2022+Report+to+the+Nations.pdf

Barr-Pulliam, D. (2019). The effect of continuous auditing and role duality on the incidence and likelihood of reporting management opportunism. *Management Accounting Research, 44*, 44-56.

Bentley, J. W. (2021). Improving the statistical power and reliability of research using Amazon Mechanical Turk. *Accounting Horizons, 35*(4), 45-62.

Berger, L., Perreault, S., and Wainberg, J. (2017). Hijacking the moral imperative: How financial incentives can discourage whistleblower reporting. *Auditing: A Journal of Practice & Theory, 36*(3), 1-14.

Berger, P. G., and Lee, H. (2022). Did the Dodd–Frank whistleblower provision deter accounting fraud? *Journal of Accounting Research, 60*(4), 1337-1378.

Bobko, P. (1986). A solution to some dilemmas when testing hypotheses about ordinal interactions. *Journal of Applied Psychology, 71*(2), 323.

Boo, E. f., Ng, T., and Shankar, P. G. (2021). Effects of advice on auditor whistleblowing propensity: Do advice source and advisor reassurance matter? *Journal of Business Ethics, 174*(2), 387-402.

Brink, A. G., Lowe, D. J., and Victoravich, L. M. (2013). The effect of evidence strength and internal rewards on intentions to report fraud in the Dodd-Frank regulatory environment. *Auditing: A Journal of Practice & Theory, 32*(3), 87-104.

Brown, J. L., Evans, J. H., and Moser, D. V. (2009). Agency theory and participative budgeting experiments. *Journal of Management Accounting Research, 21*(1), 317-345.

Brüggen, A., and Luft, J. (2011). Capital rationing, competition, and misrepresentation in budget forecasts. *Accounting, Organizations and Society, 36*(7), 399-411.

Brunner, M., and Ostermaier, A. (2019). Peer influence on managerial honesty: The role of transparency and expectations. *Journal of Business Ethics, 154*, 127-145.

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

Butler, J. V., Serra, D., and Spagnolo, G. (2020). Motivating whistleblowers. *Management Science, 66*(2), 605-621.

Carcello, J. V., and Hermanson, D. R. (2008). Fraudulent financial reporting: How do we close the knowledge gap. *Research Studies (White Papers) of Institute for Fraud Prevention (IFP)*.

Chen, C. X., Nichol, J. E., and Zhou, F. H. (2017). The effect of incentive framing and descriptive norms on internal whistleblowing. *Contemporary Accounting Research, 34*(4), 1757-1778.

Cressey, D. R. (1973). *Other people's money: a study of the social psychology of embezzlement.* Patterson Smith Publishing Corporation.

Curtis, M. B., Robertson, J. C., Cockrell, R. C., and Fayard, L. (2021). Peer ostracism as a sanction against wrongdoers and whistleblowers. *Journal of Business Ethics, 174*(2), 333-354.

Davila, T., and Wouters, M. (2005). Managing budget emphasis through the explicit design of conditional budgetary slack. *Accounting, Organizations and Society, 30*(7), 587-608.

Directive (EU) 2019/1937 of the European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law (2019). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L1937

Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010, Pub. L. No. 111-203, 124 Stat. 1375 (2010). https://www.congress.gov/bill/111th-congress/house-bill/4173

Dorminey, J., Fleming, A. S., Kranacher, M.-J., and Riley, R. A. (2012). The evolution of fraud theory. *Issues in Accounting Education, 27*(2), 555-579.

Dyck, A., Morse, A., and Zingales, L. (2010). Who blows the whistle on corporate fraud? *The Journal of Finance, 65*(6), 2213-2253.

Endenich, C., and Trapp, R. (2020). Ethical implications of management accounting and control: A systematic review of the contributions from the Journal of Business Ethics. *Journal of Business Ethics, 163*, 309-328.

European Commission. (2017). Summary results of the public consultation on whistleblower protection, https://ec.europa.eu/newsroom/just/items/54254/en

Evans, J. H., Hannan, R. L., Krishnan, R., and Moser, D. V. (2001). Honesty in managerial reporting. *The Accounting Review, 76*(4), 537-559.

Farrell, A. M., Grenier, J. H., and Leiby, J. (2017). Scoundrels or stars? Theory and evidence on the quality of workers in online labor markets. *The Accounting Review, 92*(1), 93-114.

Free, C., and Murphy, P. R. (2015). The ties that bind: The decision to co-offend in fraud. *Contemporary Accounting Research, 32*(1), 18-54.

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

Hannan, R. L., Rankin, F. W., and Towry, K. L. (2006). The effect of information systems on honesty in managerial reporting: A behavioral perspective. *Contemporary Accounting Research, 23*(4), 885-918.

Hannan, R. L., Rankin, F. W., and Towry, K. L. (2010). Flattening the organization: The effect of organizational reporting structure on budgeting effectiveness. *Review of Accounting Studies, 15*(3), 503-536.

Hogan, C. E., Rezaee, Z., Riley Jr, R. A., and Velury, U. K. (2008). Financial statement fraud: Insights from the academic literature. *Auditing: A Journal of Practice & Theory, 27*(2), 231-252.

Jonason, P. K., and Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment, 22*(2), 420.

Kahneman, D., and Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341.

Kassem, R., and Higson, A. (2012). The new fraud triangle model. *Journal of Emerging Trends in Economics and Management Sciences, 3*(3), 191-195.

Kuang, Y. F., Lee, G., and Qin, B. (2021). Whistleblowing allegations, audit fees, and internal control deficiencies. *Contemporary Accounting Research, 38*(1), 32-62.

Latan, H., Chiappetta Jabbour, C. J., and Lopes de Sousa Jabbour, A. B. (2019). 'Whistleblowing triangle': Framework and empirical evidence. *Journal of Business Ethics, 160*(1), 189-204.

Latan, H., Chiappetta Jabbour, C. J., and Lopes de Sousa Jabbour, A. B. (2021). To blow or not to blow the whistle: The role of rationalization in the perceived seriousness of threats and wrongdoing. *Journal of Business Ethics, 169*(3), 517-535.

Lee, G., and Xiao, X. (2018). Whistleblowing on accountingy-related misconduct: A synthesis of the literature. *Journal of Accounting Literature, 41*(1), 22-46.

Li, J., Chai, L., Nordstrom, O., Tangpong, C., and Hung, K.-T. (2021). Development of a Loss Aversion Scale. *Journal of Managerial Issues, 33*(1), 69-89.

Libby, R., Bloomfield, R., and Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society, 27*(8), 775-810.

Libby, T., and Lindsay, R. M. (2010). Beyond budgeting or budgeting reconsidered? A survey of North-American budgeting practice. *Management Accounting Research, 21*(1), 56-75.

Maas, V. S., and Yin, H. (2022). Finding partners in crime? How transparency about managers' behavior affects employee collusion. *Accounting, Organizations and Society, 96*, Article 101293.

Maussen, S., Cardinaels, E., and Hoozée, S. (2024). Costing system design and honesty in managerial reporting: An experimental examination of multi-agent budget and capacity reporting. *Accounting, Organizations and Society, 112*, Article 101541.

Moore, C. (2015). Moral disengagement. *Current Opinion in Psychology, 6*, 199-204.

Moore, C., Detert, J. R., Klebe Treviño, L., Baker, V. L., and Mayer, D. M. (2012). Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel Psychology, 65*(1), 1-48.

Morales, J., Gendron, Y., and Guénin-Paracini, H. (2014). The construction of the risky individual and vigilant organization: A genealogy of the fraud triangle. *Accounting, Organizations and Society, 39*(3), 170-194.

Murphy, P. R. (2012). Attitude, Machiavellianism and the rationalization of misreporting. *Accounting, Organizations and Society, 37*(4), 242-259.

Murphy, P. R., and Dacin, M. T. (2011). Psychological pathways to fraud: Understanding and preventing fraud in organizations. *Journal of Business Ethics, 101*(4), 601-618.

Peer, E., Brandimarte, L., Samat, S., and Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153-163.

Peer, E., Vosgerau, J., and Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior research methods, 46*, 1023-1031.

Pope, K. R., and Lee, C.-C. (2013). Could the Dodd–Frank Wall Street Reform and Consumer Protection Act of 2010 be helpful in reforming corporate America? An investigation on financial bounties and whistle-blowing behaviors in the private sector. *Journal of Business Ethics, 112*(4), 597-607.

Rigdon, E. E., Schumacker, R. E., and Wothke, W. (1998). A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*. Erlbaum Associates.

Robinson, S. N., Robertson, J. C., and Curtis, M. B. (2012). The effects of contextual and wrongdoing attributes on organizational employees' whistleblowing intentions following fraud. *Journal of Business Ethics, 106*(2), 213-227.

Sloane, E. H. (1944). Rationalization. *The Journal of Philosophy, 41*(1), 12-21.

Smaili, N., and Arroyo, P. (2019). Categorization of whistleblowers using the whistleblowing triangle. *Journal of Business Ethics, 157*(1), 95-117.

Sommerfeldt, R. D. (2023). Who Gets Stitches? The Effects of Rewarding Whistleblowers and Protecting Their Identity on Subsequent Cooperation.

Spoelma, T. M., Chawla, N., and Ellis, A. P. (2021). If you can't join 'em, report 'em: A model of ostracism and whistleblowing in teams. *Journal of Business Ethics*, *173*(2), 345-363.

Strube, M. J., and Bobko, P. (1989). Testing hypotheses about ordinal interactions: Simulations and further comments. *Journal of Applied Psychology*, *74*(2), 247.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39-60.

Towry, K. L. (2003). Control in a teamwork environment—The impact of social ties on the effectiveness of mutual monitoring contracts. *The Accounting Review*, *78*(4), 1069-1095.

U.S. Securities and Exchange Commission. (2015). 2015 Annual report to congress on the Dodd-Frank Whistleblower Program, https://www.sec.gov/files/owb-annual-report-2015.pdf

U.S. Securities and Exchange Commission. (2021). SEC Bars Two Individuals from Whistleblower Award Program, https://www.sec.gov/news/press-release/2021-199

Van Lange, P. A., Bekkers, R., Schuyt, T. N., and Vugt, M. V. (2007). From games to giving: Social value orientation predicts donations to noble causes. *Basic and Applied Social Psychology*, *29*(4), 375-384.

Wells, J. T. (2004). New approaches to fraud deterrence. *Journal of Accountancy*, *197*(2), 72-76.

Wolfe, D. T., and Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. *CPA Journal*, 38-42.

Xu, Y., and Ziegenfuss, D. E. (2008). Reward systems, moral reasoning, and internal auditors' reporting wrongdoing. *Journal of Business and Psychology*, *22*(4), 323-331.

Zhang, Y. (2008). The effects of perceived fairness and communication on honesty and collusion in a multi-agent setting. *The Accounting Review*, *83*(4), 1125-1146.

FIGURES & TABLES

Figures

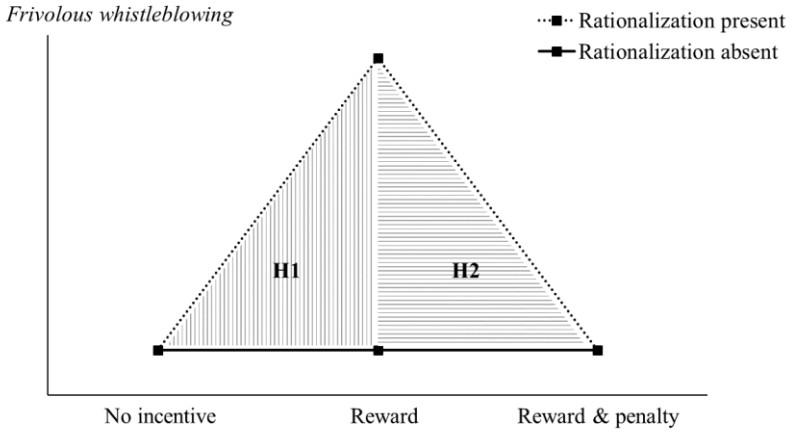

FIGURE 1 Predicted pattern of results for hypothesis 1 (H1) and 2 (H2).

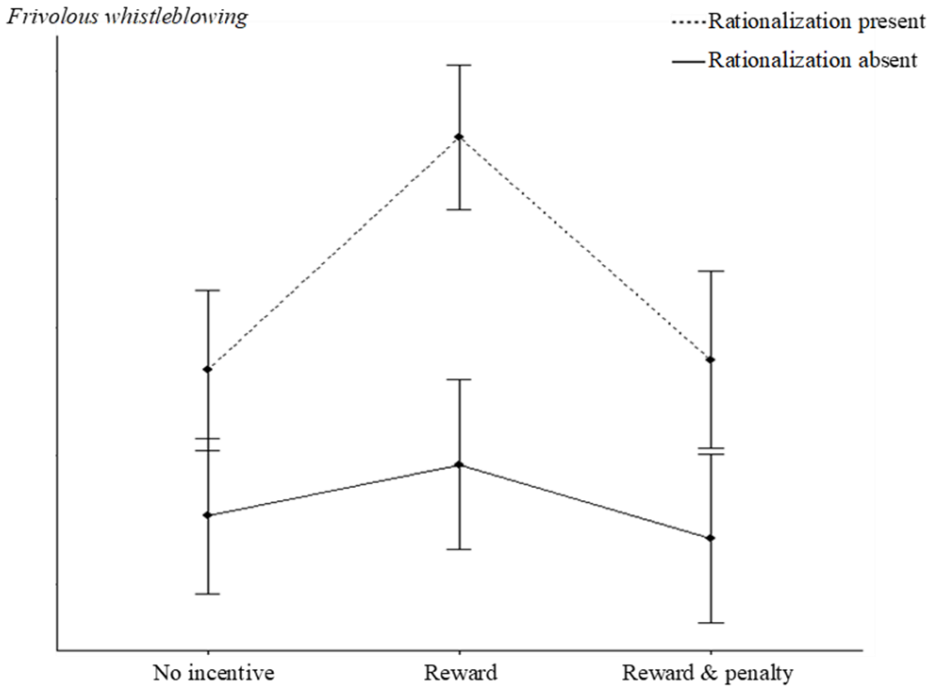

FIGURE 2 Observed data for hypothesis 1 and 2.

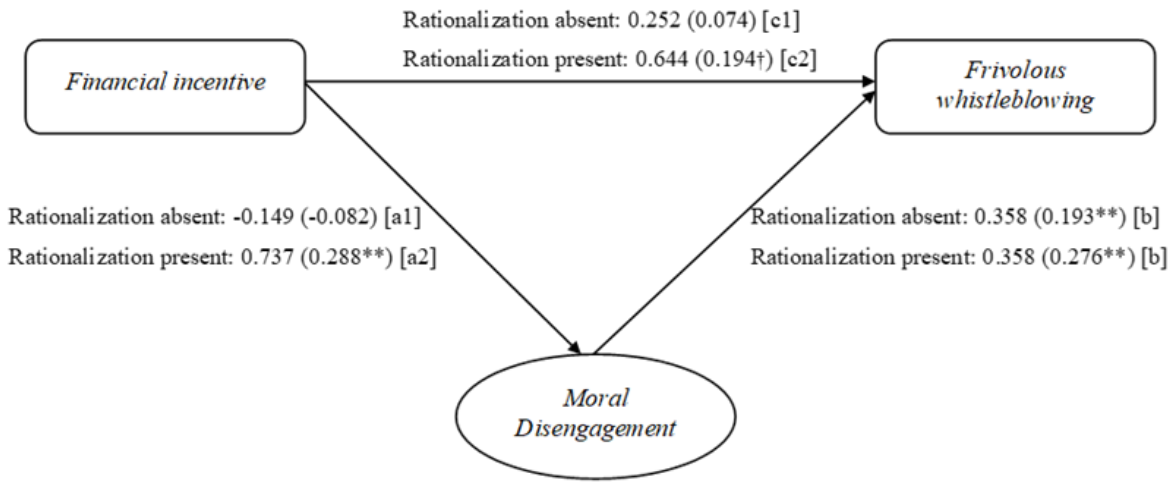

FIGURE 3 Causal model H1.[a,b,c, d]

Notes:

[a] The unconstrained model (where path 'a' and 'c' are free) results in a significantly better model (p = 0.097) as compared to the constrained model where path 'a' and 'c' are restricted to be the same across different levels of *rationalization*.

[b] This figure reports unstandardized (standardized) path coefficients resulting from structural equation modeling in R using Lavaan.

[c] *, **, *** indicates significance at the 10%, 5% and 1% level respectively (two-tailed). † indicates a marginally significant p-value between 0.10 and 0.15.

[d] Overall goodness of fit: χ^2 = 36.261; p = 0.166; CFI = 0.960; TLI = 0.942; RMSEA = 0.065; SRMR = 0.065.

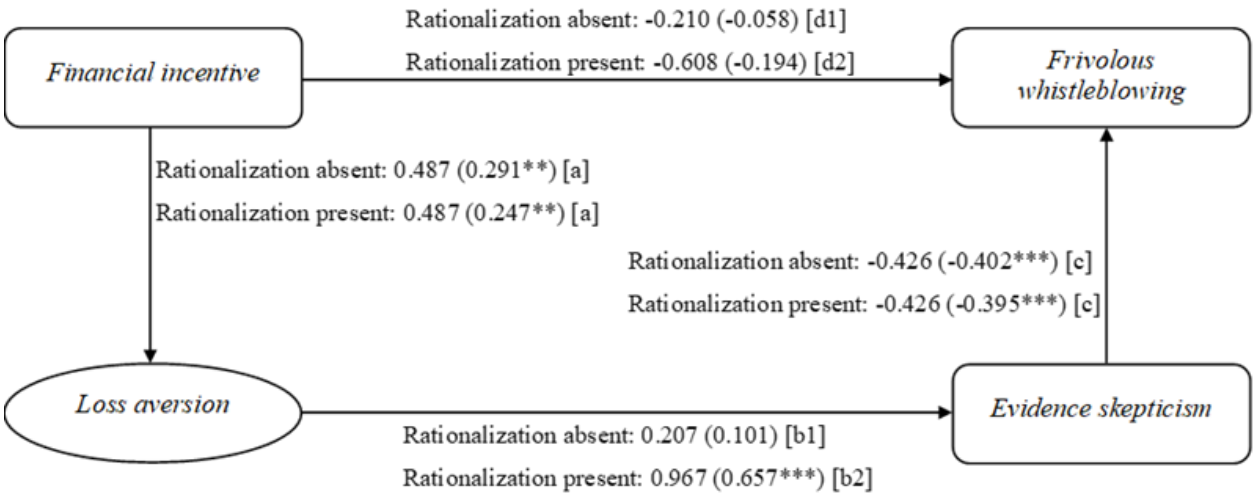

FIGURE 4 Causal model H2.[a,b,c, d]

Notes:

[a] The unconstrained model (where path 'b' and 'd' are free) results in a marginally significantly better model ($p = 0.106$) as compared to the constrained model where path 'b' and 'd' are restricted to be the same across different levels of *rationalization*.

[b] This figure reports unstandardized (standardized) path coefficients resulting from structural equation modeling in R using Lavaan.

[c] *, **, *** indicates significance at the 10%, 5% and 1% level respectively (two-tailed).

[d] Overall goodness of fit: $\chi^2 = 86.747$; $p = 0.011$; CFI = 0.851; TLI = 0.818; RMSEA = 0.092, SRMR = 0.105.

Tables

TABLE 1 *Rationalization* as operationalized in our experimental instrument.

Murphy and Dacin (2011)	Experimental scenario
Moral justification (appeal to higher loyalties)	…the '4D-Theatre' project is vital for long term company performance. As such, I sincerely hope we can find a way to both receive the governmental grant in the near future, as well as to push through with the '4D-Theatre' project on the long term.
Advantageous comparison	Indeed, delaying a budget in order to be sure that it is entirely correct is much better than having the budget rejected by the CFO in the short term because of some office politics.
Euphemistic labeling	Whistleblowing = asking for a quality-assuring audit.
Denial of the victim	Second, no colleagues are wrongfully hurt at all. Only if your colleague responsible for estimating overhead costs actually included budgetary slack, your colleague will be penalized.
Displacing responsibility	First, it is your project leader who is responsible for conducting the audit and hence, who is responsible for punishing wrongdoers.

TABLE 2 Manipulation check questions.

Question	# incorrect answers
Truthful whistleblowing is rewarded: if my whistleblowing tip results in the successful detection of budgetary slack, I will receive a reward equal to 30,000 Lira.	8
Wrongful whistleblowing is penalized: if my whistleblowing tip is unsubstantiated/frivolous, I will receive a penalty equal to 30,000 Lira.	11
The CFO will not accept any new projects which focus on the development of tangible assets in the near future.	23

TABLE 3 Descriptives.[a, b, c]

	Rationalization absent	Rationalization present	Total
No incentive	53 (1.77 | 1.65) *N* = 30	70 (2.33 | 1.71) *N* = 30	123 (2.05 | 1.69) *N* = 60
Reward for accurate whistleblowing	57 (1.97 | 1.78) *N* = 29	94 (3.24 | 1.53) *N* = 29	151 (2.60 | 1.77) *N* = 58
Reward for accurate & penalty for inaccurate whistleblowing	47 (1.68 | 1.74) *N* = 28	57 (2.38 | 1.69) *N* = 24	104 (2.00 | 1.74) *N* = 52
Total	157 (1.80 | 1.71) *N* = 87	221 (2.66 | 1.68) *N* = 83	378 (2.22 | 1.74) *N* = 170

Notes:

[a] Each cell first reports the number of whistleblowing tips. The numbers between parentheses are the corresponding average (ranging from 0 to 5) and their standard deviations.

[b] '*N*' represents the number of observations in each cell.

[c] Whistleblowing is defined as 'disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to effect action' (Near and Miceli (1985, p.4). It is coded as 1 if the participant decides to submit a whistleblowing tip and 0 otherwise. The outcome of all five whistleblowing rounds are summed up to one another. As such, the average number of whistleblowing tips between brackets is the average number of whistleblowing tips submitted across five rounds.

TABLE 4 Hypothesis 1.

Panel A: Two-Way ANOVA Hypothesis 1[a]: $F_{(3, 114)} = 4.47$, $p = 0.005$				
Factor	SS	*F*-value	df	*p*-value[b]
Financial incentive	9.03	3.23	1	0.074*
Rationalization	25.03	8.96	1	0.003***
Financial incentive × *Rationalization*	3.71	1.33	1	0.252
Error	318.31		114	
Panel B: One-Way ANOVA Hypothesis 1[a] (planned contrast test)				
Factor	SS	*F*-value	df	*p*-value[b]
Experimental condition	37.45	4.47	3	0.005***
Error	318.31		114	
Panel C: Contrast weights Hypothesis 1[a] (planned contrast test)				
Contrast	1[c]	2[c]	3[c]	4[c]
Contrast weights	-1	-1	-1	+3
Panel D: Contrast test Hypothesis 1[a] (planned contrast test)				
Source of variation	SS	*F*-value	df	*p*-value[b]
Hypothesized contrast	32.53	11.65	114	<0.001***
Residual[d]	4.92	0.59	3	0.624
Panel E: Follow-up simple effect tests Hypothesis 1[a]		*F*-value	df	*p*-value[b]
Effect of *financial incentive* given rationalization absent		0.21	1	0.649
Effect of *financial incentive* given rationalization present		4.36	1	0.039**
Effect of *rationalization* given no incentive		1.73	1	0.192
Effect of *rationalization* given reward for accurate whistleblowing		8.45	1	0.004***

Notes:

[a] Hypothesis 1: Financial rewards for accurate whistleblowing incentivize frivolous whistleblowing when rationalization is present, but not when rationalization is absent.

[b] *, **, *** indicates significance at the 10%, 5% and 1% level respectively (two-tailed).

[c] Experimental condition 1: no incentive & rationalization absent; Experimental condition 2: no incentive & rationalization present; Experimental condition 3: reward for accurate whistleblowing & rationalization absent; Experimental condition 4: reward for accurate whistleblowing & rationalization present.

[d] The residual sum of squares represents the between-group variance unexplained by the hypothesized contrast. A non-significant *p*-value indicates that the hypothesized contrast explains all of the between-cells variance in the data.

TABLE 5 Hypothesis 2.

Panel A: Two-Way ANOVA Hypothesis 2[a]: $F_{(3, 106)} = 4.66$, $p = 0.004$				
Factor	SS	*F*-value	df	*p*-value[b]
Financial incentive	9.09	3.19	1	0.077*
Rationalization	26.58	9.33	1	0.003***
Financial incentive × *Rationalization*	2.29	0.81	1	0.372
Error	302.01		106	
Panel B: One-Way ANOVA Hypothesis 2[a] (planned contrast test)				
Factor	SS	*F*-value	df	*p*-value[b]
Experimental condition	39.86	4.66	3	0.004***
Error	302.01		106	
Panel C: Contrast weights Hypothesis 2[a] (planned contrast test)				
Contrast	1[c]	2[c]	3[c]	4[c]
Contrast weights	-1	3	-1	-1
Panel D: Contrast test Hypothesis 2[a] (planned contrast test)				
Source of variation	SS	*F*-value	df	*p*-value[b]
Hypothesized contrast	32.51	11.41	106	0.001***
Residual[d]	7.34	0.86	3	0.465
Panel E: Follow-up simple effect tests Hypothesis 2[a]		*F*-value	df	*p*-value[b]
Effect of *financial incentive* given rationalization absent		0.41	1	0.522
Effect of *financial incentive* given rationalization present		3.46	1	0.066**
Effect of *rationalization* given reward for accurate whistleblowing		8.28	1	0.005***
Effect of *rationalization* given reward for accurate and penalty for inaccurate whistleblowing		2.20	1	0.141†

Notes:

[a] Hypothesis 2: Adding a penalty for inaccurate whistleblowing to the reward for accurate whistleblowing decreases the number of frivolous whistleblowing tips when rationalization is present, but not when rationalization is absent.

[b] *, **, *** indicates significance at the 10%, 5% and 1% level respectively (two-tailed). † indicates a marginally significant *p*-value between 0.10 and 0.15.

[c] Experimental condition 1: reward for accurate whistleblowing & rationalization absent; Experimental condition 2: reward for accurate whistleblowing & rationalization present; Experimental condition 3: reward for accurate and penalty for inaccurate whistleblowing & rationalization absent; Experimental condition 4: reward for accurate and penalty for inaccurate whistleblowing & rationalization present.

[d] The residual sum of squares represents the between-group variance unexplained by the hypothesized contrast. A non-significant *p*-value indicates that the hypothesized contrast explains all of the between-cells variance in the data.

TABLE 6 Measurement model *moral disengagement*.

Item	Question	Std. factor loading
1	Wrongfully blowing the whistle is morally justified as long as it is in the best interest of the company as a whole.	0.692
2	There is no such a thing as wrongful whistleblowing.	0.691
3	Whistleblowing is just requesting an additional audit.	0.565
4	Compared to lying to external parties, wrongfully blowing the whistle is not bad.	0.811
5	Wrongfully accusing someone is not that bad, since the accused person is only punished if they actually did something wrong.	0.630

TABLE 7 Indirect and total effects causal model H1.

Panel A: Indirect effects			
	Estimate	Std. Estimate	*p*-value[a]
Rationalization absent: a1×b	-0.053	-0.016	0.581
Rationalization present: a2×b	0.264	0.056	0.103†
Panel B: Total effects			
	Estimate	Std. Estimate	*p*-value[a]
Rationalization absent: c1 + (a1×b)	0.199	0.059	0.651
Rationalization present: c2 + (a2×b)	0.908	0.250	0.029**

Notes:
[a] *, **, ***, indicate significance at the 10%, 5%, and 1% levels respectively (two-tailed). † indicates a marginally significant *p*-value between 0.10 and 0.15.

TABLE 8 Measurement model *loss aversion*.

Item	Question	Std. factor loading
1	When deciding on whether or not to blow the whistle, I thought more about what might be lost due to a wrongful whistleblowing tip than what might be gained from submitting a truthful whistleblowing tip.	0.437
2	I felt nervous when I needed to decide on whether or not to blow the whistle.	0.524
3	The pain from losing something (e.g., reputation, money) by submitting a wrongful whistleblowing tip mattered much more to me than the pleasure from gaining the same by submitting a truthful whistleblowing tip.	0.569
4 (reverse coded)	Avoiding failure by submitting a wrongful whistleblowing tip is less important than seeking success by submitting a truthful whistleblowing tip.	Excluded due to negative factor loading.
5	Experiencing a major loss by submitting a wrongful whistleblowing tip will stay longer in my mind than experiencing a major gain from submitting a truthful whistleblowing tip.	0.548
6	A potential failure by submitting a wrongful whistleblowing tip scared me more than a potential success by submitting a truthful whistleblowing tip encouraged me.	0.827
7 (reverse coded)	The suffering that came with submitting a wrongful whistleblowing tip could be fully offset by the pleasure that came from submitting a truthful whistleblowing tip.	Excluded due to negative factor loading.
8	The fear of losing money by submitting a wrongful whistleblowing tip would matter more to me than gaining the same amount of money by submitting a truthful whistleblowing tip.	0.532

TABLE 9 Indirect and total effects causal model H2.

Panel A: Indirect effects			
	Estimate	Std. Estimate	*p*-value[a]
Rationalization absent: a×b1×c	-0.043	-0.010	0.524
Rationalization present: a×b2×c	-0.201	-0.064	0.035**
Panel B: Total effects			
	Estimate	Std. Estimate	*p*-value[a]
Rationalization absent: d1 + (a×b1×c)	-0.253	-0.068	0.569
Rationalization present: d2 + (a×b2×c)	-0.809	-0.259	0.038**

Notes:
[a] *, **, *** indicates significance at the 10%, 5% and 1% level respectively (two-tailed).

TABLE 10 Robustness checks.

Model[a, b]	1		2		3		4	
n in each model	170		185		133		170	
Panel A: ANCOVA	H1	H2	H1	H2	H1	H2	H1	H2
Financial incentive	0.078*	0.077*	0.097*	0.082*	0.213	0.114†	0.731	0.482
Rationalization	0.005*	0.004***	<0.001***	0.001***	<0.001***	0.011**	0195	0.004***
Financial incentive × *Rationalization*	0.270	0.448	0.423	0.423	0.213	0.097*	0.223	0.302
Machiavellianism	0.807	0.083*	-	-	-	-	-	-
Psychopathy	0.523	0.180	-	-	-	-	-	-
Narcissism	0.769	0.451	-	-	-	-	-	-
Social value orientation	0.927	0.838	-	-	-	-	-	-
Total approvals	-	-	0.107†	0.830	-	-	-	-
Endowment effect	-	-	-	-	-	-	0.738	0.122†
Work experience	-	-	-	-	-	-	0.126†	0.815
Panel B: Simple effect tests	H1	H2	H1	H2	H1	H2	H1	H2
Effect of *financial incentive* given rationalization absent[c]	0.619	0.472	0.538	0.460	>.999	0.954	0.731	0.482
Effect of *financial incentive* given rationalization present[c]	0.040**	0.078*	0.079*	0.079*	0.084*	0.021**	0.041**	0.041**
Effect of *rationalization* given no incentive	0.197	-	0.055*	-	0.168	-	0.195	-
Effect of *rationalization* given reward for accurate whistleblowing	0.007***	0.009***	0.004***	0.005***	0.002***	0.002***	0.003***	0.004***

Effect of rationalization given reward for accurate and penalty for inaccurate whistleblowing	-	0.135†	-	0.071*	-	0.538	-	0.200
Panel C: Hypotheses								
Hypothesis 1	Supported		Supported		Supported		Supported	
Hypothesis 2	Supported		Supported		Supported		Supported	

Notes:
[a] Each cell contains the two-sided *p*-value.
[b] *, **, *** indicates significance at the 10%, 5% and 1% level respectively (two-tailed). † indicates a marginally significant *p*-value between 0.10 and 0.15.
[c] While the *financial incentive* for H1 entails either no incentive or a financial reward for accurate whistleblowing, the *financial incentive* for H2 entails either a reward for accurate whistleblowing or a reward for accurate and penalty for inaccurate whistleblowing.