The Role of Personality in Product Recalls

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2023

vorgelegt
von

Daniel Frederik Gass, M.Sc.

aus
Schweinfurt

Referent:	Prof. Dr. Andreas Fügener
Koreferent:	Prof. Ulrich W. Thonemann, Ph.D.
Tag der Promotion:	26.04.2024

Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
BFI	Big Five Inventory
CDRH	Center for Devices and Radiological Health
CEO	Chief Executive Officer
COO	Chief Operating Officer
FDA	United States Food and Drug Administration
FE	Fixed Effects
FFM	Five-Factor Model
GEE	Generalized Estimating Equations
ITCV	Impact Threshold of a Confounding Variable
OLS	Ordinary Least Squares (Regression)
RIR	Robustness of Inference to Replacement
S&P 1500	S&P Composite 1500 Index
VIF	Variable Inflation Factor

Chapter 1

Introduction

Personality is to a man what perfume is to a flower.

— *Charles M. Schwab*

1.1 Motivation

Personality is key to our human existence. From an early age, many children grow up to the wisdom of their parents that they are truly "one of a kind". This uniqueness is also — at least thus far — a crucial difference to most machines, algorithms, tools and systems that humans have built. Next to physical differences, humans stand out from one another through our personality.

Extensive evidence and expertise on personality has accumulated within the field of personality psychology. According to a definition by the American Psychological Association (2022), an individuals' personality is made up of "enduring characteristics and behavior that comprise a person's unique adjustment to life, including major traits, interests, drives, values, self-concept, abilities, and emotional patterns". For example, researchers have undertaken considerable efforts to develop classifications and codifications of personality traits, such as the Five-Factor Model (FFM) of personality (McCrae and Costa 1987). Researchers have also studied ways to codify predispositions like regulatory focus, which are arising from an individuals' personality (Higgins 2015). Finally, extensive bodies of evidence have accumulated which show that personality affects a wide range of important outcomes at the individual and organizational level (Larsen and Buss 2010, Schneider and Smith 2004).

Despite the plethora of evidence outside the field of operations management (OM) that personality has a key impact on a wide range of outcomes, "little has traditionally been done to incorporate individual differences into OM research" (Bendoly et al. 2006, p. 741). Concretely, only a small number of studies have evaluated the benefits of using personality in optimization decisions (e.g. Juran and Schruben 2004) or considered the impact of personality traits on outcomes like inventory management (e.g. Strohhecker and Größler 2013). This lack of in-depth research on personality in the OM literature is particularly surprising given that in other fields of the management literature, a large body of research has formed focusing specifically on personality. For example, in strategic management, researchers have extensively studied how important firm-level outcomes can be explained by recourse to strategic leaders like the CEO (Hambrick 2007, Neely et al. 2020), and their personality (e.g., Gupta et al. 2019, Nadkarni and Herrmann 2010).

Building on the foundations of the personality psychology literature, this thesis aims to contribute to further our understanding of the role and impact of personality in operations management. In this endeavour, the thesis places a significant focus on an important aspect relating to firms' operations, namely product recalls (Shah et al. 2017). Very little has been done so far to explore the role of personality in explaining product recalls (Wowak and Boone 2015), as is the case in many other streams of the operations management literature.

1.2 Outline and Contribution

This thesis features three independent research projects. While these research projects are set in different contexts, consider different types of outcomes and use different methodological approaches, they share the purpose to improve

the understanding of how personality shapes important outcomes in operations management. Each of the three main chapters (Chapters 2 – 4) presents one research project.

Chapters 2 and 3 are hypothesis-based research projects located at the intersection of the operations management and strategic leadership literatures and investigate the impact of personality of CEOs' — firm's most strategic leaders — on important operations outcomes, primarily product recalls. Concretely, Chapter 2 investigates the role of CEO regulatory focus on product innovations and the number of product safety problems. Based on these findings, Chapter 3 focuses on product safety problems and investigates how CEOs' personality traits affect the handling of individual product recalls, analyzing time-to-recall. These projects clearly establish the importance of personality in operations management broadly and product recalls specifically.

Chapter 4 builds on the findings of the importance of personality and answers calls in the operations management literature to investigate "the interaction between humans and machines" (Samson and Kalchschmidt 2019, p. 1) against the backdrop of advancements in artificial intelligence (AI), which present a "major development for operations management" (Loske and Klumpp 2021, p. 1). The chapter features an exploratory research project at the intersection of operations management and information systems (IS), and is based on data from an experimental study. While this project does not directly study product recalls, its findings not only have important implications for the broader operations management literature, but also the specific literature on product recalls insofar as it highlights the importance of and effects from individual's personality traits.

We briefly outline the three research projects below. In Chapter 5, we summarize the key findings and conclude with a review of this thesis' contribution and directions for future research.

Chapter 2 examines the effect of CEOs' regulatory focus, an important aspect relating to individuals' personality, on product innovation and product safety problems. Building on regulatory focus theory (Higgins et al. 1997, Higgins 1998), which describes two distinct motivational foci, we hypothesize that CEO promotion focus and CEO prevention focus have opposing effects on product innovations and product safety problems. We posit that the exploration and achievement orientation emblematic of promotion focus is associated with more subsequent product innovations. However, since this orientation can also lead individuals to take hasty and risky decisions, we hypothesize that CEO promotion focus is also associated with more product safety problems. As prevention focus is associated with a focus on safety, protection and vigilance, but also slower and more risk-averse decision making, we hypothesize that CEO prevention focus is associated with fewer product safety problems and product innovations. We conduct an empirical study using a panel of healthcare firms listed in the S&P 1500 health care index comprising 1,197 firm-years across 136 firms. We gather data from the U.S. Food and Drug Administration (FDA) on approvals and recalls of medical devices and find strong support for three and partial support for one of our four hypotheses. Additionally, we are able to show that the effect of CEO prevention focus on product safety problems is strongest in the largely unexplored situations where CEOs also hold the title of COO. This work makes a number of contributions, including to the operations management and product recall literatures. Importantly, we complement and extend prior research on product recalls which has thus far primarily focused on proxies of leaders' cognition (e.g., Wowak et al. 2015, 2021), rather than aspects of their personality like regulatory focus.

Chapter 2 is joint work with Prof. Dr. Andreas Fügener, Prof. Dr. Lorenz Graf-Vlachy, and Prof. Daniel L. Gamache, Ph.D.. I collected, processed, and analyzed the empirical data and wrote all sections of the chapter. The co-authors

provided input on the empirical design and analysis, and provided feedback on the paper and its positioning.

Chapter 3 builds on the findings of Chapter 2 and more closely investigates the role of CEOs' personality traits in the handling of individual product recall events. Specifically, the paper focuses on recall timing, or *time-to-recall*, which is the time span between a firm becoming aware of a product defect and issuing a recall for the affected product. We build on personality theory to hypothesize that more neurotic CEOs (i.e., those CEOs with higher levels of neuroticism, which is a key personality trait) are more vigilant when they face uncertainty, as they do when making recall decisions, and strive to avoid the potential threat of social disapproval from recalling late, leading to lower time-to-recall. We contextualize our argumentation by considering two moderators which go to the heart of our theoretical argumentation. First, we take a configurational approach to personality to hypothesize that another personality trait, CEO agreeableness, increases the salience of the threat of social disapproval and thus positively moderates the main relationship. Second, we build on trait activation theory to argue that more negative recent media coverage of a CEO provides a trait-relevant situational cue which reinforces CEOs' neurotic tendencies and thus also positively moderates the main relationship. We analyze 1,573 product recalls overseen by 110 CEOs of healthcare firms and find robust support for our hypotheses. Our work makes a number of contributions to the operations management and recall literatures, but also to the strategic leadership literatures. For product recalls, we demonstrate that the personality of strategic leaders not only impacts the number of product safety problems (and resulting recalls) that firms suffer, but also the handling of specific recall events. We show that multiple traits of these individuals' personalities, as well as situational factors interact to shape time-to-recall.

Chapter 3 is joint work with Prof. Dr. Andreas Fügener and Prof. Dr. Lorenz

Graf-Vlachy and is based on Gass et al. (2022), which was presented at the *Academy of Management Annual Meeting* held in Seattle in 2022. I collected, processed, and analyzed the empirical data and wrote most of the sections of the chapter. Both co-authors provided input on the empirical design and analysis, the positioning of the paper, and line-edited the paper. The paper further benefited from participants of the *Strategic Management Society Conference (SMS)* set in Toronto in 2021 and the *Manufacturing and Service Operations Management Society Conference (MSOM)* held in Munich in 2022.

Chapter 4 explores the role of personality in Human-AI collaboration. As recent research suggests that humans and AI will continue to work side by side — in spite of recent advancements in artificial intelligence (AI) — findings ways how humans can work collaboratively with AI is critically important. A key obstacle that has been identified is that humans' ability to capture complementary knowledge while working with AI is critically constrained by our limited ability to assess what we know, i.e., metaknowledge. However, extant research has largely overlooked a critical difference between human and AI: Human behavior is shaped by personality. Building on data from an image classification experiment with 281 participants, we explore how differences arising from personality traits impact metaknowledge, delegation behavior towards AI and collaborative performance of humans working with AI. We find that higher levels of conscientiousness improve, while higher levels extraversion harm humans' ability to work collaboratively with AI. We find that differences in metaknowledge help to explain this relationship. Analyzing a Treatment condition, we show that providing humans with information on the AI's certainty has heterogeneous effects on human performance. While such an intervention is beneficial for individuals with higher levels of conscientiousness and neuroticism, it is harmful for individuals with higher levels of extraversion. Concretely, this intervention harms their ability to distinguish their confidence between images they know and ones they do not know, and significantly increases

their confidence in incorrect decisions. Our work in this project makes a number of contributions to the information systems and operations management literatures. Next to specific contributions on Human-AI collaboration and metaknowledge, this work also contributes to the wider operations management field as it highlights the desirable and undesirable effects of personality traits on individuals' cognitions. Clearly, these have a wide-ranging impact in operations management, including when it comes to the handling of product safety problems and product recalls.

Chapter 4 is single-authored. This project is an exploratory analysis based on data from Taudien et al. (2022). I provided input to the original experimental design, including on the selection and integration of the Big-Five Inventory (BFI; John et al. 1991). I ran the analyses and wrote all sections of the chapter. I presented an earlier version of Chapter 4 at the *International Conference on Information Systems (ICIS)* in Hyderabad in 2023. This version is published as Gass (2023).

Chapter 2

"To do good or to do no harm": The effect of CEO regulatory focus on product innovations and product safety problems

Regulatory focus theory proposes that individuals pursue goals via one of two motivational foci, which can easily be described by key phrases from the Hippocratic Corpus. Promotion focus (aiming to "do good") is a desire to achieve an ideal state of being, which makes individuals strive for gain. Prevention focus (making sure to "do no harm") is characterised by a sensitivity to loss, which makes individuals desire stability and security. We build on existing research showing that the regulatory focus of CEOs' — firms' most important strategic leaders — impacts firm outcomes and posit that CEO promotion focus and CEO prevention focus have counter-acting effects on product innovations and product safety problems. Concretely, we hypothesize that CEO promotion focus encourages a higher number of product innovations, but that the same characteristic is are also linked adversely (i.e., positively) with a firm's incidence of product safety problems. Reversely, we propose that CEO prevention focus is linked with a lower number of product safety problems, but that the vigilance associated with CEOs' prevention focus contributes to a lower output of product innovations. We find strong support for three of our four hypotheses using a unique data set of product innovations and product safety problems at 136 medical device firms across 1,197 firm-years. Additionally, we run a number of supplementary analyses, in which we are able to demonstrate that the effect of CEO prevention focus — in contrast to CEO promotion focus — is strongest in the largely unexplored constellation where CEOs also hold the title of COO.

2.1 Introduction

The physician must [. . .] have two special objects in view with regard to disease, namely to do good or to do no harm.

— *Hippocrates*

Philosophic principles, such as those outlined in the Hippocratic Corpus, have long guided not only physicians, but most of those working in the healthcare industry. Millions of employees in healthcare firms are seeking to find innovative and safe ways to cure diseases and benefit patients. Yet as the principles of “doing good” and “do[ing] no harm” highlight, there is an inherit tension between innovating novel approaches and delivering safe products to patients. If healthcare professionals think outside of the proverbial box, dream big, and take chances, they are more likely to develop and release innovative products which “do good” for their patients. At the same time, this focus means they are typically more prone to error. In the context of the medical products, such errors can manifest themselves in product safety problems with detrimental impact on patients. On the flip side, if professionals are concentrated on staying safe, anticipating and preventing such safety problems and “doing no harm”, this typically inhibits their ability to develop and release these innovative products in the first place.

Reconciling this tension is one of the many challenging tasks facing Chief Executive Officers (CEOs), who are arguably healthcare firms’ most important strategic leaders. In their highly demanding roles, CEOs — among a host of other responsibilities — set the overall strategic direction of their firms and decide how to best allocate spare resources (Mintzberg 1973, Herrmann and Nadkarni 2014). Because of their unique position, researchers have been able to explain a variety of

important firm outcomes by recourse to CEO personality (Finkelstein et al. 2009).

We posit that CEO’s regulatory focus — a widely used psychological theory describing a key aspect of an individual’s personality — is critical to understand the role of CEOs in product innovations and product safety problems. Regulatory focus describes individuals’ motivational focus using two distinct self-regulatory systems: promotion focus and prevention focus (Higgins 1997). Individuals with a strong promotion focus strive toward an ideal state of being and ensuring hits (“doing good”), while individuals with a pronounced prevention focus have a desire to avoid negative outcomes (“doing no harm”) (Crowe and Higgins 1997, Higgins 1998).

Surprisingly, extant literature has not jointly studied the conflicting roles of CEOs’ regulatory foci — CEO promotion and CEO prevention focus — in product innovation and product safety problems. While considerable work has been done to understand the role of CEO regulatory focus on accounting measures of inputs to the innovation process, such as R&D expenditure (e.g., Scoresby et al. 2021, Mount and Baer 2022), prior studies have not investigated effects on product innovation outcomes, such as new product approvals. In the literature on product safety problems, the role of CEOs has previously been examined only through proxies of executive cognition such as compensation measures (e.g., Wowak et al. 2015), rather than more direct measures of CEO’s personality like regulatory focus.

Concretely, we hypothesize that CEO promotion focus and CEO prevention focus have opposing influence on product innovations and product safety problems. Individuals with a promotion focus self-regulate their behavior to achieve hits, i.e., making them more alert to possible opportunities (Crowe and Higgins 1997). This motivational focus drives exploration and innovative behaviour (McMullen and Kier 2016, Spanjol et al. 2011). Clearly, this focus should drive CEOs (and those

they lead) to launch more product innovations. At the same time, the focus on "doing good" also makes individuals with a promotion focus more prone to make mistakes and underestimate risks (Brockner et al. 2004, Higgins 1997). It's not difficult to expect that this will manifest in a larger incidence of safety problems for the products which are released. Accordingly, we hypothesize that CEO promotion focus is positively associated with product innovations and positively associated with product safety problems.

By contrast, individuals with a prevention focus are keen to avoid potential risks and act with vigilance to prevent mistakes (Crowe and Higgins 1997). In particular, they are keen to minimize errors of commission, i.e., failing to detect a problem (Higgins 2000). As a result, these individuals place a greater emphasis on security and safety (Förster et al. 2003). These characteristics should clearly drive CEOs (and the firms they lead) to "do no harm" and help to prevent safety problems. As this behaviour comes at the expense of lower decision-making speed (Förster et al. 2003) and is associated with lower exploratory activity (Kammerlander et al. 2015), it should lower the pace of successful development of product innovations. Thus, we hypothesize that CEO prevention focus is negatively associated with product safety problems and negatively associated with product innovations.

We test our hypotheses on a unique data set of product approvals and product recalls from a sample of healthcare firms which develop and market medical devices. We leverage data from the U.S. Food and Drug Administration (FDA), the world's preeminent regulatory agency in the healthcare industry, to build a comprehensive data set of newly approved medical devices (which we use to capture successful product innovations) and product recalls (which is the process through which firms must address product safety problems). Our final data set consists of 1,197 firm-years across 136 firms, spanning the period 2005 through 2017. We find strong support for three (and partial support for one) of our four

hypotheses.

Our study makes three primary contributions. First, we contribute to the growing literature on CEO regulatory focus by developing and empirically validating a theoretical model linking CEO regulatory focus to two important firm-level outcomes: product innovations and product safety problems. Within a single context (the medical device industry) and using one unified sample, we are able to demonstrate that the positive effect of CEO promotion focus on product innovations and the negative effect of prevention focus on product safety problems are inherently associated with less desirable, reverse effects on the respective other outcome. Additionally, we answer calls by Scoresby et al. (2021, p. 415) to investigate the “outcome of R&D investments”. By investigating product innovation outcomes directly, we extend and complement prior research which has thus far focused on the effect of CEO regulatory focus on accounting measures of innovation inputs.

Second, our findings also allow us to contribute to the literature on product safety problems and product recalls. We find that CEO promotion focus is conducive to greater subsequent incidence of product recalls, while CEO prevention focus appears to inhibit a higher incidence thereof. With these findings, we demonstrate that CEO cognition as shaped by aspects of their personality like regulatory focus affects the incidence of product safety problems. Thus, we complement and extend prior research which has thus far primarily focused on proxies of leaders’ cognition (e.g., Wowak et al. 2015, 2021).

Third, our supplementary analyses allow us to make a number of smaller contributions to the literature on strategic leaders. We find that the under-explored role of the COO (Hambrick and Cannella 2004), and the combination of the CEO and COO roles specifically, is important to understand the impact of CEO prevention focus on product safety problems. Additionally, our finding on the impact of CEO

gender on product innovations and product safety problems adds to the growing evidence that the gender of strategic leaders impacts important firm outcomes (e.g., Wowak et al. 2021).

2.2 Theory and Hypotheses

As Hambrick and Mason argued in their seminal work on upper echelons theory, organizational outcomes can be viewed as the “reflections of the values and cognitive bases of powerful actors” (1984, p. 193). In particular, such powerful actors like a firm’s strategic leaders act based on their personal interpretations of the situations they find themselves in (Finkelstein et al. 2009). Their interpretations of these situations, in turn, are shaped by their prior experiences, values, and personalities (Hambrick 2007). These key ideas have laid the groundwork for what has become a vast research program on strategic leaders (see Neely et al. 2020 for a recent review).

Out of all strategic leaders, CEOs have received the most attention. CEOs are usually firms’ most publicly visible figures and frequently become a personification of their firms (Finkelstein et al. 2009, Love et al. 2017). Additionally, they have a disproportionate influence on the overall success of the firm they lead (Quigley and Hambrick 2015), including on a wide range of areas and processes like corporate social responsibility programs (e.g., Wernicke et al. 2022), innovation (e.g., Hirshleifer et al. 2012), and product recalls (e.g., Liu et al. 2016).

Given this wide-ranging influence of CEOs, it is not surprising that researchers have repeatedly explained organizational outcomes by recourse to a variety of CEO characteristics. Traditionally, these include positional (e.g., tenure), demographic (e.g., age), and compensation-related (e.g., stock option pay) factors, which are readily observable without requiring “intrusive access to large numbers of

executives” (Hambrick 2007, p. 337). Recently, researchers have extended their reach beyond such proxy measures, using novel and unobtrusive techniques to obtain more direct measures of CEOs cognitive processes, such as CEO cognitive complexity (e.g., Graf-Vlachy et al. 2020), CEO extraversion (e.g., Malhotra et al. 2018), and CEO regulatory focus (e.g., Gamache et al. 2015).

Regulatory focus theory is a psychological framework that explains individuals’ motivational focus using two distinct self-regulatory systems: promotion focus and prevention focus (Higgins 1997). The theory, which has been widely adopted inside and outside of the field of psychology, posits that individuals align (i.e., self-regulate) their behavior to achieve goals by pursuing aspirations and striving for positive outcomes, or by preventing losses and avoiding negative outcomes (Brockner and Higgins 2001). These foci have "a major impact on people’s feelings, thoughts, and actions” (Higgins 1998, p. 37).

Even though regulatory focus theory describes two contrasting foci, these foci are “not two ends of a continuum but orthogonal” (Kammerlander et al. 2015, p. 585). As the two foci are independent inclinations, individuals generally self-regulate their behavior using a combination of both prevention and promotion foci (Higgins 1998). This means that although many individuals may have a predominant inclination, “individuals can be high or low on promotion or prevention, or both” (Améndola et al. 2022, p. 4).

Regulatory focus is commonly viewed as a chronic dimension relating to individual’s personality (Kammerlander et al. 2015). Meta-analyses show that regulatory focus is generally stable over time, particularly within a work context (Lanaj et al. 2012). Although research does acknowledge that “momentary situations are also capable of temporarily inducing either promotion focus concerns or prevention focus concerns” (Higgins et al. 2003, p. 1141), we focus on chronic, rather than situational regulatory focus.

Extant research on CEOs has used CEO prevention and promotion focus to explain a number of important firm outcomes. For example, Gamache et al. (2015) found that CEO promotion focus and prevention focus were positively and negatively associated, respectively, with the number and size of mergers and acquisitions. Researchers have also been able to explain risk tasking (e.g., Mount and Baer 2022), stakeholder strategy (e.g., Gamache et al. 2020), and marketing controversies (e.g., Kashmiri et al. 2019) by recourse to CEO promotion and prevention focus.

2.2.1 CEO Promotion Focus

Promotion focus describes individuals’ desire to approach an ideal state of being (Higgins 1998). Those with a promotion focus act with eagerness and self-regulate their behavior to “insure hits and insure against errors of omission” (Crowe and Higgins 1997, p. 120). As a result, a promotion focus induces individuals to be particularly alert to opportunities, which they seek to capture (McMullen and Kier 2016). This means that unlike those with prevention focus, promotion-focused individuals see anything other than an improvement to the status quo — including its preservation — as a failure (Higgins 2014). In order to try to achieve these gains, promotion-focused individuals explore more distinct alternatives (Crowe and Higgins 1997).

Numerous studies have shown that the focus on capturing different opportunities drives exploratory and innovative behaviour. In a laboratory experiment with university students, Spanjol et al. (2011) demonstrated that a small group with a pre-dominant promotion focus introduced a higher number of product innovations and did so more quickly than did those groups with a prevention focus. Similarly, a meta study on regulatory focus in the workplace found that promotion focus was positively related to innovative performance (Lanaj et al. 2012). In a field study

of departments in the semiconductor industry, Tuncdogan et al. (2017) found a positive positive effect of promotion focus on exploratory innovation.

We build on this broad evidence, as well as specific research into the effects of CEO promotion focus, to suggest that CEO promotion focus will be positively related to the firm’s number of product innovations. Prior studies into CEO promotion focus have found positive effects on exploration (Kammerlander et al. 2015), R&D spending (Scoresby et al. 2021) and R&D intensity (Kashmiri et al. 2019). We propose that these tendencies not only manifest themselves in measures of innovation input, such as R&D spending, but also in the output (i.e., the number) of production innovations. We suggest that the effect of CEOs’ promotion focus on product innovations manifests itself through at least one of two possible avenues. First, CEOs will take decisions (e.g., support novel ideas), shape policies (e.g., reward successful product innovations), and allocate financial resources (e.g., for research and development) in ways that are consistent with their regulatory focus. Second, research has shown that the regulatory focus of leaders “trickles down” to influence the situational regulatory focus of followers, who would thus be more likely to act in a promotion-focused way (Johnson et al. 2017). We formalize:

Hypothesis 2.1. *Higher levels of CEO promotion focus are positively associated with the number of subsequent product innovations.*

While the eagerness of individuals with a promotion focus helps them to achieve “hits” (Crowe and Higgins 1997, p. 120), this specific tendency also carries significant downsides in other domains. Because promotion-focused individuals self-regulate their behaviour not to miss opportunities, they are more prone to make mistakes and underestimate risks (Higgins 1997, Brockner et al. 2004). This leads them to take riskier decisions (Florack and Hartmann 2007), prioritize speed over accuracy (Förster et al. 2003) and support judgements with feelings rather than facts (Higgins and Cornwell 2016). This focus also makes individuals rely

more on positive, rather than negative information (Cunningham et al. 2005, Yoon et al. 2012).

Specific downsides of these tendencies have also been shown in research on CEOs. For example, Kashmiri et al. (2019) found that promotion focused CEOs were more likely to encounter marketing controversies. In a study on workplace safety, Qian et al. (2023) found the employee injury rate to be higher in sites of firms run by a CEO with a high promotion focus.

We propose that these tendencies also manifest themselves in a higher number of product safety problems. A safety problem is a defect in a product which poses a safety (e.g., health) risk to its users (Wowak and Boone 2015). In regulated industries such as the healthcare industry, firms issue product recalls to remove these unsafe products from the market and prevent harm to users (Thirumalai and Sinha 2011). Research consistently indicates that when serious safety problems occur, firms and their leaders have little discretion whether to issue a product recall or not (e.g., Wowak et al. 2015, 2021).

We believe there are at least three ways in which CEO promotion focus could reasonably influence the number of product safety problems. First, promotion-focused CEOs will take decisions (e.g., approving product launches), shape policies (e.g., escalation processes) and allocate financial resources (e.g., quality control) in ways that are more risky, including when launching new products. This could raise the chances that potential safety problems develop in the first place. Second, because promotion focused individuals more aptly process positive rather than negative information (e.g. Yoon et al. 2012), they should also be less attuned to early warning signs of potential safety problems. Finally, CEO promotion focus may also shape situational regulatory focus of company employees (e.g., Johnson et al. 2017). Accordingly, we expect that:

Hypothesis 2.2. *Higher levels of CEO promotion focus are positively associated*

with the number of subsequent product safety problems.

2.2.2 CEO Prevention Focus

Prevention focus describes individual's desire to avoid negative outcomes (Higgins 1998). Individuals with a prevention focus act with vigilance to achieve goals (Crowe and Higgins 1997), and are generally concerned with "protection, safety, and responsibility" (Higgins 1998, p. 27). Prevention focused individuals are motivated to ensure that there is no deterioration compared to a status quo and would thus also see the preservation of the status quo as success in itself (Higgins 2014). This caution manifests itself in lower risk seeking behaviour (Florack and Hartmann 2007), higher prioritization of accuracy over speed (Förster et al. 2003), and increased sensitivity towards negative, rather than positive information (Yoon et al. 2012).

Studies have also shown the impact of this emphasis on duty and safety in the workplace. In a meta study, Lanaj et al. (2012) demonstrate that prevention focus is positively associated with a number of safety-related outcomes. Qian et al. (2023) specifically study the impact of CEO prevention focus on workplace safety. They find that the rate of employee injuries is lower in locations led by CEOs with high prevention focus, demonstrating that CEO prevention focus positively impacts firm-level safety outcomes.

Accordingly, we propose that CEO prevention focus will manifest itself in a lower incidence of product safety problems. Again, we suggest that CEOs' prevention focus could impact product safety problems through at least three ways. Higher levels of prevention focus are associated with increased sensitivity towards negative information, making them more sensitive to early warning signs of potential safety problems. Clearly, prevention-focused CEOs will take decisions, shape policies

and allocate resources in ways which are consistent their “vigilance” strategy (Higgins 2000, p. 1200). Additionally, CEO prevention focus may strengthen the situational prevention focus of followers (e.g. Johnson et al. 2017). We formalize:

Hypothesis 2.3. *Higher levels of CEO prevention focus are negatively associated with the number of subsequent product safety problems.*

While the deliberate planning and vigilance of prevention-focused individuals should be beneficial when it comes to preventing or minimizing product safety problems, we expect these exact qualities to be detrimental to their ability to lead their firms to be successful in product innovation. Prevention focus leads individuals to prioritze accuracy, even if this comes at the expense of lower speed (Förster et al. 2003). In a lab experiment, this translated into a lower speed and lower output of product innovations (Spanjol et al. 2011). Similar detrimental effects of prevention focus on exploration and innovation have also been shown for management teams (e.g. Tuncdogan et al. 2017), and CEOs specifically (e.g. Kammerlander et al. 2015). This leads us to hypothesize a direct effect between CEO prevention focus and output of product innovations, such that:

Hypothesis 2.4. *Higher levels of CEO prevention focus are negatively associated with the number of subsequent product innovations.*

2.3 Data and Methodology

2.3.1 Sample

Our analyses are based on a sample of publicly traded healthcare firms listed in the S&P 1500. We chose the S&P 1500 because it ensures the availability of comparable data on critical control variables due to the firms’ homogeneous

reporting requirements while simultaneously enhancing the generalizability of our findings by allowing for substantial variance in firm size. Specifically, we collected data on all firms listed in the S&P 1500 health care sector index (GICS 35). This ensures that we are focusing on firms which are primarily healthcare firms and not large conglomerates which have a small base in the healthcare industry, which would unduly bias our empirical estimations (Chatterji and Fabrizio 2014). Additionally, we excluded firms in those sub-industries which focus on non-physical healthcare services, such as the operation of healthcare facilities like hospitals (see Appendix 2.A). We gathered data on healthcare firms listed in the index for at least three consecutive fiscal years between 2005 and 2018 to ensure the availability of data for our (lagged) control variables.

Focusing on U.S. healthcare firms provided us with a single regulatory context as these firms are tightly regulated by the *U.S. Food and Drug Administration (FDA)*. The FDA regulates a wide variety of product categories, including healthcare (e.g., medical devices) and non-healthcare products (e.g., food). Dedicated branches of the agency, known as “centers”, oversee each product category. These centers provide legally binding regulations as well as guidelines for key processes, including how to make applications to register new products and how to recall unsafe products.

In order to ensure that differences between the center do not unduly bias our results, we focused our study on product innovations and product safety problems in medical devices and gather data from the *Center for Devices and Radiological Health (CDRH)*. Healthcare products are overseen by three centers. The *Center for Drug Evaluation and Research (CDER)* oversees pharmaceutical products (i.e., drugs), the *Center for Biologics Evaluation and Research (CBER)* oversees biologics (including blood products and vaccines), and the *CDRH* oversees medical devices. Not only is the nature of the product categories themselves very different

(e.g., product cycles are much longer in the pharmaceutical industry than in the medical device industry), but regulatory processes and powers also differ considerably by center. For example, unlike the CDRH, which can mandate firms to recall unsafe medical devices, the CDER cannot easily mandate firms to recall unsafe drugs (Medical Device Recall Authority 1996, FDA 2020).

We compiled data-sets on product innovations and product safety problems using both publicly available data from the CDRH section of the FDA website and several Freedom of Information Act requests and interactions with the agency. Our final data set with all required variables includes 1,197 firm-years across 136 firms, which are overseen by 229 unique CEOs.

2.3.2 Variable Definitions

Dependent Variables

Number of product approvals. We construct a firm-year count of the number of newly approved medical devices, which “provides a reliable record that is a reasonable approximation of innovation outcomes” (Chatterji and Fabrizio 2014, p. 1432). We thus go beyond measuring innovation by analyzing process inputs (e.g., Scoresby et al. 2021 analyzed changes to R&D spending) or intermediate measures (e.g., Balkin et al. 2000 analyzed patent applications) of innovation. In this way, we answer calls to investigate innovation outcomes more directly (e.g., Scoresby et al. 2021).

New medical devices are submitted to the FDA for approval through one of three regulatory processes. Depending on the novelty of the product and the criticality of its planned usage, complexity and rigour of the approval process varies considerably. First, if firms believe that the product is “substantially equivalent” to a product that is already being marketed by any firm, firms can submit an

application using the simplified *510(k)* approval process (FDA 2022). Second, if this is not the case and the product is not considered to be high risk[1], firms can submit a *De Novo* request (FDA 2022). To ensure that we capture all product innovations, we deviate from prior literature (e.g., Chatterji and Fabrizio 2014) and also consider these approvals. Third, life-sustaining or otherwise critical novel devices without substantially equivalent predecessors or competitor products need to undergo the more rigorous *premarket approval (PMA)* process (FDA 2022, Singh 2007). We only consider the initial submissions as a product innovation, and disregard all supplemental PMA applications as these are filed when firms make small changes to the way existing products are designed (e.g., new packaging) or produced (e.g., new manufacturing location) (FDA 2019).

We gathered all approved applications made through these three processes, and proceeded to match them to our sample firms. For this purpose, we compiled two lists of company names. First, we collected a list of historical company names by company and fiscal year using data from COMPUSTAT and 10-K filings with the U.S. Securities and Exchange Commission (SEC). This ensured that we could match the innovation even if the company had since changed name. Second, we collected a list of all consolidated subsidiaries for each firm-year using their 10-K filings with the SEC[2]. Being able to match even those innovations submitted by a subsidiary of our sample firm is critical to ensure that different corporate structures or naming conventions for subsidiaries do not bias our results. We used the application date — rather than the decision date — for matching applications to firm-years to ensure that differences in the duration of the approval process do

[1]Specifically, the FDA classifies products into one of three classes (Class I, II, and III). Devices are classified into the high-risk Class III if “the device is life-supporting or life-sustaining, or for a use which is of substantial importance in preventing impairment of human health, or if the device presents a potential unreasonable risk of illness or injury” (21 CFR 860.3; Medical Device Classification Procedures 2023). Note that these classes are di□erent to the classification of recall severity, which also uses a system of three classes.

[2]Specifically, firms are required by law to list all subsidiaries which are in their majority ownership and thus consolidated in their financial reporting in Exhibit 21.1 of their 10-K annual report (17 CFR § 229.601; Exhibits 2023).

not adversely affect the validity of our measure.

Finally, we aggregated the number of approved innovations by firm for each year *t* into the *Number of product approvals* variable, which is a count variable ranging from 0 to 96.

Number of product recalls. We used product recalls as a way to measure product safety problems. Not only is this an intuitive operationalization — recalls are a clear sign that a product has a safety problem — but also well established in the recall literature (e.g., Shah et al. 2017). In line with prior research (e.g., Wowak et al. 2015), we focus on those safety problems which may cause permanent harm or even death (Class 1 recalls) or otherwise jeopardize the health of their users (Class 2 recalls) (FDA 2020). It is important to note that because the industry is so closely regulated, firms are legally prohibited from covertly addressing product safety problems in the field without formally declaring a recall (FDA 2020, Hensley 2010). Accordingly, prior research suggests that firms have little discretion for launching recalls for such safety problems (Wowak et al. 2021). This means that the number of product recalls should be a rather comprehensive measure of safety problems these firms experience. We did not include Class 3 recalls as these reflect instances where a product is recalled for problems which are unlikely to have any safety impact (Ball et al. 2018c)[3]. Given that firms have more discretion whether they undertake a recall to rectify such problems, recall behavior for Class 3 recalls may not be consistent across firms (Ball et al. 2018c).

We matched product recalls of medical devices as reported by the CDRH to our sample firms using the same process described above for innovation (i.e., matching by name of the firm or any of its registered subsidiaries). We used the recall initiation date to match recalls to the respective fiscal year. Finally, we aggregated the number of recall events by firm for each year *t* into the *Number of product*

[3] In section 2.4.2, we nonetheless also report results using an alternative operationalization which includes Class 3 recalls.

recalls variable, which is a count variable ranging from 0 to 59.

Independent Variables

We obtained unobtrusive measures of *CEO promotion focus* and *CEO prevention focus* using content analysis of CEO communication. This method is well established and widely used for measuring CEO cognition broadly (e.g., DesJardine and Shi 2021), and CEO regulatory focus specifically (e.g., Gamache et al. 2015).

We followed Chen et al. (2018) and used CEO speeches in earning conference calls to capture the strength of their regulatory focus. We gathered all available transcripts of quarterly earnings conference calls for our sample firms from Thomson StreetEvents and Fair Disclosure wires on LexisNexis. Then, we processed each transcript of each call by extracting the opening speech made by the CEO. We collated all available speeches by the CEO at the focal firm in the prior two fiscal years. We used the Linguistic Inquiry and Word Count (LIWC) software with dictionaries developed and comprehensively validated by Gamache et al. (2015) to capture the strength of CEOs' prevention and promotion foci. These dictionaries have also been extensively used in other recent work on CEO regulatory focus (e.g., Mount and Baer 2022, Gamache et al. 2020). Our final variables *CEO promotion focus* and *CEO prevention focus* are continuous variables indicating the percentage of words matched to the respective dictionaries (Gamache et al. 2015). Accordingly, they are unaffected by the length of CEO speeches.

Control Variables

We controlled for a variety of firm- and CEO-specific factors in year $t - 1$ that may have an impact on our dependent variables. This time lag is in line with prior research in both the recall (e.g., Wowak et al. 2015, 2021) and innovation

literatures (e.g., Hirshleifer et al. 2012). All models also included year dummy variables to capture possible effects arising from time-based changes in the overall business or regulatory environments.

We controlled for a host of possible confounding firm-level factors using data from COMPUSTAT. We controlled for firm size using the firm's *total assets* (natural log of total assets in USD million). In line with previous research on recalls (e.g., Thirumalai and Sinha 2011), we controlled for *R&D intensity* (R&D expenditure divided by sales), and *firm profitability* (return on assets). We also included *total debt/assets* to account for firm risk-taking (Cronqvist et al. 2012, Malhotra et al. 2018) and *Altman's Z-score* to capture financial distress (Altman 1968). Finally, we controlled for the complexity of the firm's corporate structure using the *number of subsidiaries* reported in the firm's 10-K filing.

In light of the recent findings by Wowak et al. (2021), we also controlled for the *female board ratio* (share of female board members) and included a control variable for *board independence* (percentage of non-executive directors on the board) using corresponding data from BoardEx.

We controlled for a variety of CEO-level factors, including ones relating to demographics, position and compensation using data from Execucomp. To capture potential effects arising from CEO demographics, we included a continuous variable for *CEO age* and a dummy variable for *CEO gender* indicating whether the CEO was female. We measured *CEO tenure* as the number of years a CEO has held their position at the focal firm (e.g., Wowak et al. 2015) and additionally included a dummy variable for *CEO change*. We used proxies to measure CEOs' power to influence their organizations (Malhotra et al. 2018). Specifically, we controlled for *CEO duality* using a binary variable indicating whether the CEO is also the chairperson of the board, and *CEO stock ownership*, a continuous measure of the percentage of voting shares owned by the CEO. We also included

a dummy variable *CEO also COO* indicating if the CEO also held the title of COO (or an equivalent title like President; Hambrick and Cannella 2004) during the year. Finally, to account for possible effects arising from the way in which CEOs are compensated (e.g., Wowak et al. 2015, 2017), we control for *CEO option pay* (percentage of total pay in stock options using their ex-ante value), *CEO non-option pay* (natural log-transformed) and *CEO exercisable options* using the natural log of the value of in-the-money stock options.

2.3.3 Model Specification

We analyzed our data using generalized estimating equations (GEEs), a technique pioneered by Liang and Zeger (1986). GEEs are widely used in research on CEOs broadly (e.g. Certo et al. 2017), and in research into the role of CEOs in firm innovation (e.g., Kashmiri et al. 2019) and product safety problems (e.g., Shah et al. 2017, Wowak et al. 2015) specifically. They are uniquely suited for our analyses for three reasons. First, GEEs are well suited for longitudinal studies with panel data in which repeated (and thus inter-temporally correlated) measurements are obtained for the same sample firms. We specified an auto-regressive (AR1) correlation structure (e.g., Quigley et al. 2020, Hambrick and Quigley 2014). Second, GEEs do not require the outcome variable to be normally distributed. Because our dependent variables are count variables (i.e., non-negative integer values which frequently take on the value zero), we specified a negative binomial distribution with a log link function (Ballinger 2004). Third, standard errors can be calculated in a way that they remain valid “even if the correlations within the group deviate from the specified correlation structure” (Rhee et al. 2006, p. 508). Specifically, we used Huber-White robust standard errors, clustered at the firm level, to avoid any potential issues arising from heteroskedasticity (Wooldridge 2002). We estimated our models using the xtgee function in Stata 17.0.

Table 2.1: Summary statistics and correlations

Variables	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
(1) DV: Number of product approvals	4.24	9.04	1.00																					
(2) DV: Number of recalls	3.03	7.61	0.71*	1.00																				
(3) Total assets	8258.01	18333.58	0.41*	0.41*	1.00																			
(4) RD intensity	0.24	0.94	-0.08*	-0.07*	-0.13*	1.00																		
(5) Return on assets	0.05	0.14	0.08*	0.05	0.14*	-0.35*	1.00																	
(6) Total debts / assets	0.18	0.19	0.01	0.05	0.29*	0.04	-0.16*	1.00																
(7) Altman's Z-score	7.17	9.21	-0.08*	-0.10*	-0.23*	-0.09*	0.31*	-0.42*	1.00															
(7) Number of subsidiaries	53.52	105.39	0.40*	0.35*	0.55*	-0.08*	0.05	0.11*	-0.16*	1.00														
(9) Female board ratio	0.14	0.09	0.23*	0.21*	0.49*	-0.04	-0.04	0.25*	-0.20*	0.21*	1.00													
(10) Board independence	0.86	0.06	0.03	0.10*	0.23*	0.03	-0.07*	0.20*	-0.12*	0.11*	0.22*	1.00												
(11) CEO age	54.74	7.04	-0.11*	-0.11*	0.06*	-0.06*	0.05	-0.07*	0.16*	-0.10*	-0.04	-0.05	1.00											
(12) CEO gender	0.03	0.16	-0.07*	-0.06*	-0.08*	0.00	-0.03	0.03	-0.05	-0.06*	0.07*	-0.08*	0.04	1.00										
(13) CEO tenure	7.29	7.28	-0.10*	-0.10*	-0.09*	-0.01	0.05	-0.14*	0.22*	-0.06	-0.11*	-0.20*	0.38*	-0.06	1.00									
(14) CEO change	0.20	0.40	0.01	-0.01	0.05	-0.01	0.00	0.06*	-0.09*	-0.02	0.02	0.02	-0.15*	0.07*	-0.49*	1.00								
(15) CEO duality	0.44	0.50	0.14*	0.06*	0.23*	-0.03	0.09*	0.00	0.03	0.09*	0.11*	-0.04	0.23*	0.00	0.29*	-0.15*	1.00							
(16) CEO stock ownership	1.39	4.48	-0.07*	-0.10*	-0.22*	-0.02	0.04	-0.14*	0.11*	-0.11*	-0.11*	-0.19*	0.00	-0.01	0.24*	-0.12*	0.15*	1.00						
(17) CEO also COO	0.64	0.48	-0.08*	0.01	-0.09*	-0.07*	0.00	-0.01	0.02	-0.01	0.03	0.14*	-0.09*	0.01	-0.17*	0.04	-0.30*	-0.10*	1.00					
(18) CEO option pay	0.29	0.25	0.04	0.03	-0.01	0.02	0.00	-0.03	0.02	-0.02	-0.10*	0.04	-0.10*	-0.05	-0.02	-0.03	-0.02	-0.05	-0.01	1.00				
(19) CEO non-option pay	4659.44	5891.50	0.24*	0.23*	0.72*	-0.05	0.06*	0.28*	-0.24*	0.36*	0.41*	0.17*	0.02	-0.02	-0.06	0.02	0.24*	-0.17*	-0.12*	-0.25*	1.00			
(20) CEO exercisable options	18892.71	59249.16	0.11*	0.10*	0.21*	-0.01	0.07*	0.06	0.02	0.11*	0.01	0.03	-0.08*	0.02	0.04	-0.22*	0.05	-0.08*	-0.02	0.24*	0.19*	1.00		
(21) CEO promotion focus	1.77	0.65	0.34*	0.31*	0.35*	-0.19*	0.14*	0.18*	-0.18*	0.29*	0.27*	0.10*	-0.24*	-0.13*	-0.23*	0.04	0.01	-0.08*	0.09*	-0.01	0.30*	0.10*	1.00	
(22) CEO prevention focus	0.18	0.16	-0.12*	-0.14*	-0.13*	0.11*	-0.11*	0.01	0.00	-0.08*	-0.09*	0.01	0.04	0.09*	-0.05	-0.03	-0.05	0.03	0.01	0.04	-0.09*	-0.04	-0.18*	1.00

Un-transformed summary statistics (mean, standard deviation) shown for interpretability.
Year dummy variables not included.
Correlations highlighted with a * are significant at the $p < 0.05$ level.

2.4 Results

Table 2.1 presents descriptive statistics and correlations. Although observed correlations are generally low, we nonetheless tested for multicollinearity by calculating variance inflation factors (VIFs) for our models 2 and 4. The mean VIFs were 2.10 and 1.99, respectively, suggesting that multicollinearity does not materially affect our inferences (Kutner et al. 2005).

2.4.1 Test of Hypotheses

Table 2.2 displays the results of our hypothesis tests. Models 1 and 2 predict the number of product approvals (i.e., our innovation DV) while models 3 and 4 predict the number of product recalls (i.e., our product safety problem DV). In Hypothesis 2.1, we posited that higher levels of CEO promotion focus are positively associated with subsequent product innovation. The coefficient for CEO promotion focus in Model 2 is positive ($\beta = 0.265$) and highly significant ($p < 0.001$), providing support for Hypothesis 2.1. In Hypothesis 2.2, we posited that higher levels of CEO promotion focus are also positively associated with subsequent product safety problems. The coefficient for CEO promotion focus in Model 4 is positive ($\beta = 0.399$) and highly significant ($p < 0.001$), thus also providing clear support for Hypothesis 2.2.

We also tested our hypotheses on CEO prevention focus. The coefficient for the effect of CEO prevention focus on product recalls in Model 4 is negative ($\beta =$ -2.400) and highly significant ($p < 0.001$). This result indicates that firms who are led by CEOs with higher levels of prevention focus experience fewer subsequent product recalls, providing support for Hypothesis 2.3. The coefficient for the effect of CEO prevention focus on product approvals in Model 2 is negative ($\beta =$ -0.315),

Table 2.2: GEE models predicting the number of product recalls and product approvals

	Product approvals		Product recalls	
	(1)	(2)	(3)	(4)
Constant	-5.275	-5.973	-11.312	-11.515
	(1.630)	(1.625)	(1.787)	(1.871)
Lagged DV	0.007	0.002	0.014	0.012
	(0.011)	(0.010)	(0.011)	(0.011)
Total assets	0.435	0.433	0.639	0.593
	(0.076)	(0.074)	(0.091)	(0.089)
R&D intensity	-2.355	-1.957	-2.796	-2.142
	(0.379)	(0.327)	(0.771)	(0.659)
Return on assets	-0.733	-0.748	-1.273	-1.517
	(0.329)	(0.316)	(0.498)	(0.494)
Total debt/assets	-0.123	-0.119	-1.047^{+}	-1.115^{+}
	(0.460)	(0.445)	(0.623)	(0.594)
Altman's Z-Score	0.002	0.004	-0.009	-0.006
	(0.005)	(0.006)	(0.010)	(0.009)
Number of subsidiaries	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Female board ratio	-0.176	-0.414	-0.356	-0.778
	(0.590)	(0.597)	(0.887)	(0.939)
Board independence	-1.999	-2.056	1.351	1.495
	(0.657)	(0.661)	(1.035)	(1.008)
CEO age	-0.011	-0.004	-0.018	-0.005
	(0.009)	(0.008)	(0.012)	(0.012)
CEO gender	-1.794	-1.279	-1.391	-1.136
	(0.751)	(0.344)	(0.504)	(0.412)
CEO tenure	-0.018^{+}	-0.020	-0.025	-0.026^{+}
	(0.010)	(0.009)	(0.016)	(0.016)
CEO change	-0.113	-0.101	0.029	0.050
	(0.112)	(0.106)	(0.127)	(0.129)
CEO duality	0.265	0.274	0.026	0.043
	(0.098)	(0.098)	(0.128)	(0.127)
CEO stock ownership	0.007	0.007	-0.016	-0.011
	(0.009)	(0.008)	(0.014)	(0.013)
CEO also COO	0.016	0.008	0.019	-0.016
	(0.074)	(0.072)	(0.116)	(0.127)
CEO option pay	0.099	0.134	-0.102	-0.020
	(0.157)	(0.160)	(0.180)	(0.180)
CEO non-option pay	-0.028	-0.025	-0.218	-0.232
	(0.065)	(0.063)	(0.082)	(0.084)
CEO exercisable options	-0.003	-0.005	0.033	0.033
	(0.009)	(0.008)	(0.013)	(0.015)
CEO promotion focus		0.265		0.399
		(0.080)		(0.114)
CEO prevention focus		-0.315		-2.400
		(0.307)		(0.573)
Number of observations (firm-years)		1,197		
Number of clusters (firms)		136		

Robust standard errors in parentheses. Year dummy variables included but not shown.
$^{+}$ $p < 0.10$, $^{\square}$ $p < 0.05$, $^{\square\square}$ $p < 0.01$, $^{\square\square\square}$ $p < 0.001$

i.e., in line with our hypothesis. However, this effect failed to met the threshold for statistical significance. Therefore, we did not find support for Hypothesis 2.4, in which we posited that CEO prevention focus would be associated with lower subsequent innovation output.

2.4.2 Robustness Checks

We conducted several analyses to demonstrate the robustness of the tests of our hypotheses.

Sensitivity to bias and omitted variables

In order to test the sensitivity of our inferences to bias, we used the robustness of inference to replacement (RIR), a method developed by Frank et al. (2013) that is finding increasing adoption in the management literature (e.g., Harrison et al. 2018, Quigley et al. 2020). To invalidate our inferences on CEO promotion focus at the $p < 0.05$ level (i.e., to make them only marginally significant), 41% and 44% of the estimate for product approvals and product recalls, respectively, would have to be due to bias. Put differently, this analysis shows that 488 and 526 observations, respectively, would need to be replaced with cases for which no effect exists in order for our findings to be only marginally significant. For the effect of CEO prevention focus on product recalls, 54% or 652 of the observations would need to be replaced. This analysis thus bolsters our confidence in our findings.

Alternative operationalizations of both dependent variables

We also validated the robustness of our results by testing alternative operationalizations of both dependent variables. We report the results of these analyses in Table 2.3. In all of these models, the lagged dependent variable follows the adjusted definition of the dependent variable.

We tested two alternative operationalizations of innovation. In model 5, we use the number of novel product approvals to operationalize radical innovation (e.g., Chatterji and Fabrizio 2014). Concretely, we use only the count of *De Novo* and *PMA* approvals, i.e., those products without equivalent predecessor or competitor products. While both effects are directionally consistent with those reported in model 2 (CEO promotion focus $\beta = 0.347$, CEO prevention focus $\beta =$ -0.587), only the effect of CEO promotion focus is marginally significant ($p < 0.1$). For model 6, we used the FDA's strict product classification hierarchy to consider whether the type of product that firms sought to receive approval for was novel to the firm or not. Concretely, we counted only approvals in those product lines in which firms had not received any approval in the prior fiscal year (i.e., new product lines). The effect for CEO promotion is directionally consistent and highly significant ($\beta = 0.398$, $p < 0.001$). Interestingly, the effect of CEO prevention focus on innovation in new product lines is in line with our theoretical predictions (i.e., negative; $\beta =$ -1.194) and significant ($p < 0.05$).

Put together, these robustness checks for innovation provide strong support for the associations between CEO promotion focus and subsequent firm innovation. We also found partial support for Hypothesis 2.4, in which we hypothesized an association between CEO prevention focus and product innovation.

For the number of product recalls, we tested four alternative operationalizations. For model 7, we operationalized the number of recalls as the count of all product recalls by firm-year, irrespective of the severity of the recall (in our main DV, we considered only recalls for safety problems, i.e., those problems which could impact patient health). Results of this operationalization reported in model 7 are consistent with those reported in model 2 for both CEO promotion focus ($\beta =$ 0.391, $p < 0.001$) and CEO prevention focus ($\beta =$ -2.361, $p < 0.001$).

Next, we created three operationalizations based on the root cause of the safety

Table 2.3: GEE models predicting alternative operationalization of product innovation (5–6) and product safety problems (7–10)

	Product approvals		Product recalls			
	Radical innovation (6)	Innovation in new product lines (7)	All recalls (8)	Design-related recalls (9)	Manuf.-related recalls (10)	Other recalls (11)
Constant	-7.268^{+}	-1.371	-11.079	-12.005	-14.609	-8.653
	(4.097)	(1.874)	(1.856)	(2.233)	(2.221)	(2.399)
Lagged DV	0.571	0.127	0.014	0.121	-0.013	0.093
	(0.130)	(0.020)	(0.011)	(0.026)	(0.031)	(0.023)
Total assets	0.179	0.156^{+}	0.577	0.478	0.685	0.457
	(0.244)	(0.086)	(0.090)	(0.113)	(0.121)	(0.122)
R&D intensity	-1.608^{+}	-3.559	-2.128	-3.939	-1.145	-3.151
	(0.951)	(1.126)	(0.738)	(1.308)	(0.698)	(1.018)
Return on assets	-1.110	-1.698	-1.349	-2.521	-0.959	-1.970
	(2.034)	(0.498)	(0.458)	(0.799)	(0.449)	(0.744)
Total debt/assets	-0.476	-0.143	-1.063^{+}	-0.559	-1.097	-0.447
	(1.002)	(0.516)	(0.642)	(0.833)	(0.828)	(0.720)
Altman's Z-Score	-0.018	-0.004	-0.005	-0.001	0.012	0.008
	(0.043)	(0.009)	(0.009)	(0.012)	(0.011)	(0.013)
Number of subsidiaries	0.002	0.002	0.001	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female board ratio	2.729	0.586	-0.476	0.836	-0.715	-0.380
	(1.671)	(0.825)	(0.901)	(0.939)	(1.041)	(1.179)
Board independence	-0.022	-1.667^{+}	1.399	1.329	1.433	0.687
	(1.748)	(0.997)	(0.948)	(1.301)	(1.340)	(1.335)
CEO age	0.024	-0.006	-0.003	-0.005	0.001	-0.014
	(0.021)	(0.012)	(0.012)	(0.017)	(0.016)	(0.019)
CEO gender	-9.163	-1.745	-0.960	-5.237	-0.256	-1.253^{+}
	(0.541)	(0.271)	(0.398)	(0.699)	(0.296)	(0.665)
CEO tenure	-0.033	-0.014	-0.027^{+}	-0.015	-0.045	-0.012
	(0.025)	(0.013)	(0.016)	(0.019)	(0.023)	(0.019)
CEO change	-0.128	-0.064	0.041	0.084	-0.139	0.008
	(0.401)	(0.133)	(0.126)	(0.213)	(0.137)	(0.182)
CEO duality	0.391	0.217	0.024	-0.004	-0.057	0.092
	(0.363)	(0.174)	(0.129)	(0.155)	(0.178)	(0.195)
CEO stock ownership	-0.315	0.003	-0.007	-0.151	0.019	-0.046
	(0.278)	(0.015)	(0.012)	(0.058)	(0.013)	(0.033)
CEO also COO	-0.499	-0.116	-0.033	0.073	-0.076	0.048
	(0.314)	(0.132)	(0.123)	(0.169)	(0.126)	(0.187)
CEO option pay	0.322	0.414	0.014	0.185	0.190	-0.221
	(0.829)	(0.268)	(0.180)	(0.329)	(0.191)	(0.315)
CEO non-option pay	-0.126	-0.110	-0.242	-0.213^{+}	-0.292	-0.227^{+}
	(0.356)	(0.097)	(0.088)	(0.125)	(0.139)	(0.120)
CEO exercisable options	0.047	0.026	0.034	0.033	0.024	0.045
	(0.046)	(0.017)	(0.013)	(0.017)	(0.015)	(0.017)
CEO promotion focus	0.347^{+}	0.398	0.391	0.505	0.241^{+}	0.616
	(0.184)	(0.101)	(0.113)	(0.136)	(0.124)	(0.157)
CEO prevention focus	-0.587	-1.194	-2.361	-1.746	-1.946	-2.192
	(1.172)	(0.484)	(0.593)	(0.657)	(0.717)	(0.883)
Number of firm-years			1,197			
Number of firms			136			

Robust standard errors in parentheses. Year dummy variables included but not shown.
$^{+}$ $p < 0.10$, $^{\square}$ $p < 0.05$, $^{\square\square}$ $p < 0.01$, $^{\square\square\square}$ $p < 0.001$

problem which necessitated the product recall (for details, please refer to Appendix 2.B). We chose to create these operationalizations to investigate whether there are any differential effects on recalls which are design-related (and thus more closely related to the product innovation and development process) as opposed to such recalls which become necessary because of problems in manufacturing. In model 8, we operationalize the DV as the count of design-related recalls. In model 9, we count only those recalls caused by manufacturing issues. In model 10, we count all other recalls.

The results of models 8 – 10 not only provide further evidence of the general robustness of our results, but also interesting additional insights. The effect of CEO prevention focus was consistent and significant for both design-related recalls (model 8, β = -1.746, $p < 0.01$), manufacturing-related recalls (model 9, β = -1.946, $p < 0.01$) and other recalls (model 10, β = -2.192, $p < 0.05$). The effect of CEO promotion focus is also consistent for both design- (model 8, β = 0.505, $p < 0.001$), manufacturing-related recalls (model 9, $\beta = 0.241$, $p < 0.1$) and other recalls (model 10, $\beta = 0.616$, $p < 0.001$). Interestingly, in contrast to design-related recalls, the effect on manufacturing-related recalls appears to be weaker and is only marginally statistically significant. This suggests that CEO promotion focus may have a more nuanced influence on product safety problems than CEO prevention focus, which appears to be associated with a lower incidence of such problems more broadly.

2.4.3 Supplemental Analysis

We performed a supplemental analysis to explore potential differences in the effects of CEO regulatory focus on innovation and safety problems between CEOs who are also Chief Operating Officer (COO) and those CEOs who are not. In contrast to CEOs, COOs have received comparatively little attention in research

Table 2.4: GEE models predicting interaction effect of CEO regulatory focus and binary variable *CEO also COO* on product innovation (models 2 and 11) and product safety problems (models 4 and 12)

	Product approvals		Product recalls	
	(2)	(11)	(4)	(12)
CEO also COO	0.008	0.077	-0.016	0.314
	(0.072)	(0.185)	(0.127)	(0.263)
CEO promotion focus	0.265	0.288	0.399	0.377
	(0.080)	(0.085)	(0.114)	(0.103)
CEO prevention focus	-0.315	-0.337	-2.400	-0.607
	(0.307)	(0.349)	(0.573)	(0.467)
CEO promotion focus		-0.037		0.012
× CEO also COO		(0.092)		(0.111)
CEO prevention focus		0.041		-2.608
× CEO also COO		(0.418)		(0.870)
Number of observations (firm-years)		1,197		
Number of clusters (firms)		136		

All control variables included in line with prior models but not shown in table.
Robust standard errors in parentheses.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

of strategic leaders (Hambrick and Cannella 2004). We build on work by Zhang (2006) to consider whether a combination of these roles impacts firm outcomes.

Concretely, we tested interaction effects between our independent variables, *CEO prevention focus* and *CEO promotion focus*, and the dummy variable *CEO also COO* indicating if the CEO also held the title of COO (or an equivalent title like President; Hambrick and Cannella 2004) during the financial year. This allowed us to separate the effects between CEOs who also are COOs, and those who are not. We report the results of the original models 2 and 4, as well as the interaction models 11 and 12, in Table 2.4. Table 2.5 shows the marginal effects.

Our results show nuanced effects. The main effect of CEO promotion focus on product approvals (model 11) and product recalls (model 12) is unaffected by the interaction (i.e., remains significant). Put differently, also holding the title of COO does not partially affect how CEO promotion focus affects these organizational outcomes. This suggests that the influence of CEO promotion focus affects the

Table 2.5: Marginal e□ect (dydx) of CEO prevention focus on product innovation and product recalls based on whether the CEO also holds the title of COO.

	Average marginal effect on DV			
	CEO is not also COO		CEO is also COO	
	dy/dx	p-value	dy/dx	p-value
Product innovation (model 11)				
CEO promotion focus	0.989	0.004	0.866	0.012
CEO prevention focus	-1.158	0.335	-1.024	0.423
Product recalls (model 12)				
CEO promotion focus	1.024	0.002	1.055	0.010
CEO prevention focus	-1.650	0.202	-8.731	0.001

organization primarily through the role of the CEO and does not require the operational involvement which is part of a COO’s role. Model 12 shows a negative (β = -2.608) and significant ($p < 0.01$) interaction between CEO prevention focus and our dummy variable on the number of recalls. By contrast, the interaction effect for CEO promotion focus is not significant. However, even in the presence of the interaction term, the main effect of CEO promotion focus remains consistent ($\beta = 0.377$) and highly significant ($p < 0.001$). An analysis of the marginal effects (shown in Table 2.5) indicates that the effect of CEO prevention focus on product recalls appears to be primarily driven by CEOs who are also COOs. Put differently, the higher operational involvement which is emblematic of the COO role appears to be required for CEO prevention focus to reduce the incidence of product safety problems[4].

[4]We also ran several alternative interaction models using some our control variables for CEO power, specifically *CEO duality* and *CEO stock ownership*. We do not find any significant interaction e□ects in these models, suggesting — at a minimum — that CEO power is not the primary driver of the e□ect

2.5 Discussion

2.5.1 Implications

In this study, we sought to examine the role of CEO promotion focus and CEO prevention focus in product innovations and product safety problems. Building on regulatory focus theory (Higgins et al. 1997, Higgins 1998), we hypothesized that CEO promotion focus and CEO prevention focus have opposing effects on product innovations and product safety problems. We posited that the exploration and achievement orientation emblematic of promotion focus is associated with more subsequent product innovations. However, since this orientation can also lead individuals to take hasty and risky decisions, we hypothesized that CEO promotion focus is associated with more product safety problems. As prevention focus is associated with a focus on safety, protection and vigilance, but also slower and more risk-averse decision making, we hypothesized that CEO prevention focus is associated with fewer product safety problems and product innovations. Using a unique data-set with product approvals and product recalls for medical devices from the FDA comprising 1,197 firm-years across 136 firms, we find strong support for three and partial support for one of our four hypotheses. In addition to testing our hypotheses, we also completed a number of supplementary analyses, providing additional interesting insights.

We make important contributions to several literature streams with our findings. We contribute to strategic leadership literature by theoretically developing and empirically validating a theoretical model linking CEO promotion focus and CEO prevention focus to two important firm-level outcomes: product innovations and product safety problems. Using the closely regulated medical devices industry as our context, we demonstrate that the desirable, positive effect of CEO promotion

focus on product innovations and the desirable (and negative) effect of prevention focus on product safety problems are inherently associated with less desirable, reverse effects on the respective other outcome. This highlights a critically important tension in the role of CEOs’ regulatory focus on firm performance. By investigating “outcome of R&D investments” rather than accounting-based measures of innovation inputs such as R&D (Scoresby et al. 2021, p. 415), we also complement and extend prior research on the role of strategic leaders like CEOs on firm innovation.

We also contribute to the Operations and Supply Chain (OSCM) literature on product recalls (Wowak and Boone 2015) by showing that CEO cognition — as shaped by aspects of their personality such as regulatory focus — affects the incidence of product safety problems. Specifically, we demonstrate that CEO promotion focus is conducive to a greater subsequent incidence of product safety problems, while CEO prevention focus is associated with a lower incidence of these safety problems. We hope that this contribution sparks further research into the role of strategic leaders, and strategic leader personality, which can complement the valuable prior research on measures which act as proxies of leaders’ cognition (e.g., Wowak et al. 2015, 2021).

We also make a number of smaller contributions to the literature on strategic leadership which opens interesting avenues for future research. We find that the combination of the CEO and COO roles is instrumental to understanding the impact of CEO prevention focus on product safety problems. By contrast, the effect of CEO promotion focus is stable no matter whether the CEO simultaneously acts as the firm’s COO or not. This indicates that the role of the COO in general, and the duality of CEO-COO roles in particular, may be interesting to understand the impact of CEO prevention focus. Additionally, these results indicate that CEO prevention and promotion focus may influence organizational outcomes

through different ways. Finally, our findings on CEO gender that female CEOs are associated with fewer innovations and safety problems add to the growing evidence (e.g., Wowak et al. 2021) that the gender of strategic leaders impacts important firm-level outcomes.

Our research also has important implications for practitioners across organizations. Not only do we demonstrate the effect of CEO regulatory focus on two important firm-level outcomes, but we also highlight the contrasting and counter-acting nature of these effects. Our findings suggest a need to ensure a balance of regulatory foci is required to achieve desired firm outcomes. For example, firms may want to ensure that the strength of regulatory foci of other members of the strategic leadership team, such as Chief Supply Chain Officers (e.g., Roh et al. 2016) or Chief Innovation Officers (e.g., Aktan et al. 2021), is complementary to those of the CEO. As our theoretical predictions build on a widely used theory from the field of psychology, it is reasonable to assume that regulatory focus of other employees — from middle managers to floor supervisors — will similarly affect outcomes within their organizational perimeters.

2.5.2 Limitations and Avenues for Future Research

The findings of our study should be viewed in the context of its limitations, which present numerous opportunities for future research. One limitation is that we are considering our two dependent variables, product innovation and product safety problems, in separate empirical models. These models build on shared theory, and we provide some initial insights that suggest that these outcomes may share common mechanisms. However, researches may want to explore more closely whether the negative impact of CEO prevention focus on product recalls can in fact be (at least partially) explained by the lower number of product innovations launched under such CEOs. To do so, researchers could consider

running a mediation analysis. Alternatively, researchers may want to attempt to map each recall back to the original product innovation, exploring whether the odds of having a recall, the number of recalls, or the time to the first recall (e.g., Hora et al. 2011) differ with respect to the regulatory foci of the CEO under which the innovation was developed and launched.

Our overall findings also open several further avenues of future research related to the product recall and strategic leadership literatures. Our supplementary analyses highlight that strategic leadership roles beyond the CEO — in particular the COO role — may be critical to understanding the impact of CEOs in general, and CEO regulatory focus in particular. Researchers in the strategic leadership literatures may in particular want to explore impacts of CEO-COO duality (i.e., CEOs also holding the title COO or an equivalent title) on other important firm outcomes. Furthermore, it may also be interesting to deepen our understanding of the interaction between CEOs and COOs when these roles are held by different individuals (e.g., Hambrick and Cannella 2004).

Finally, there are countless opportunities to deepen our understanding of the impact of CEO regulatory focus on important firm outcomes beyond product innovations and product safety problems. For example, it may be worthwhile to study potential impacts of CEO regulatory focus on other indications of safety issues, such as FDA warning letters, or to study impacts on the handling of the product recalls themselves. In the strategic leadership literature, researchers could attempt to validate whether CEOs’ regulatory focus also impacts other safety-based corporate events, such as data breaches, employee harassment, or corporate misconduct.

2.5.3 Conclusion

This study provides insights into the counter-acting effects of CEO promotion focus and CEO prevention focus on product innovations and product safety problems. Our analyses show that CEO promotion focus positively contributes towards a higher number of product innovations, but also adversely affects (i.e., increases) the number of product safety problems their firms experience. Reversely, while CEO prevention focus is linked with a lower number of product safety problems, it in turn adversely affects (i.e., decreases) the output of product innovations. We show that this effect — in contrast to CEO promotion focus — is strongest in the largely unexplored situations where CEOs also hold the title of COO. We hope that our findings can contribute to spur additional research in both the product recall and strategic leadership literatures.

Appendix of Chapter 2

2.A Industry Overview

In our sample, we included all relevant sub-industries of the healthcare sector, as defined in the Global Industry Classification Standard (GICS).

Table 2.6: Summary of the sub-industries of the healthcare sector index (GICS 35) with focus on recall-able o□erings, which are in scope for our sample. As indicated, the description column contains the original description as defined in the index and is thus not an original work of the authors.

Sub-Industry Name (Code)	**Sub-Industry Description** (Source: S&P Global Market Intelligence 2018)
Healthcare Equipment (35101010)	Manufacturers of healthcare equipment and devices. Includes medical instruments, drug delivery systems, cardiovascular orthopedic devices, and diagnostic equipment
Healthcare Supplies (35101020)	Manufacturers of healthcare supplies and medical products not classified elsewhere. Includes eye care products, hospital supplies, and safety needle syringe devices.
Healthcare Distributors (35102010)	Distributors and wholesalers of healthcare products not classified elsewhere.
Biotechnology (35201010)	Companies primarily engaged in the research, development, manufacturing and/or marketing of products based on genetic analysis and genetic engineering. Includes companies specializing in protein-based therapeutics to treat human diseases. Excludes companies manufacturing products using biotechnology but without a healthcare application.
Pharmaceuticals (35202010)	Companies engaged in the research, development, or production of pharmaceuticals. Includes veterinary drugs.
Life Sciences Tools & Services (35203010)	Companies enabling the drug discovery, development, and production continuum by providing analytical tools, instruments, consumables supplies, clinical trial services and contract research services. Includes firms primarily servicing the biotechnology and pharmaceutical industries.

The sub-industries of the medical industry listed below are not included in our

sample since they provide services, operate facilities, manage healthcare service provision, or operate data processing software, all of which cannot be recalled.

Table 2.7: Summary of the sub-industries of the healthcare sector index (GICS 35) with focus on non-recall-able o□erings, such as the operation of healthcare facilities. As indicated, the description column contains the original description as defined in the index and is thus not an original work of the authors.

Sub-Industry Name (Code)	**Sub-Industry Description** (Source: S&P Global Market Intelligence 2018)
Healthcare Services (35102015)	Providers of patient healthcare services not classified elsewhere. Includes dialysis centers, lab testing services, and pharmacy management services. Also includes companies providing business support services to healthcare providers such as clerical support services, collection agency services, staffing services and outsourced sales & marketing services.
Healthcare Facilities (35102020)	Owners and operators of healthcare facilities, including hospitals, nursing homes, rehabilitation and retirement centers and animal hospitals.
Managed Health Care (35102030)	Owners and operators of Health Maintenance Organizations (HMOs) and other managed plans.
Health Care Technology (35103010)	Companies providing information technology services primarily to healthcare providers. Includes companies providing application systems and/or data processing software, internet-based tools, and I.T. consulting services to doctors, hospitals or businesses operating primarily in the healthcare sector.

2.B Details on Methodology

We created variants of our dependent variable based on the cause of the safety problem for which the recall was issued. For our *Number of design-related recalls* variable, we counted the product recalls which were design-related. For the *Number of manufacturing-related recalls* variable, we counted all recalls which were caused by manufacturing issues. Finally, any remaining product recalls that could be clearly attributed to design- or manufacturing causes were counted in the *Number*

of other recalls variable.

We determined the recall cause using a two-step process: In the first step, we used the FDA-provided Root cause description, a categorical variable with 101 unique values which describe the root cause of the safety problem that caused the recall. Examples of these descriptions include “Labeling mix-ups”, “Device design” and “Use error”. Root causes are not provided for all recalls. Two coders independently mapped each root cause into “Design-related”, “Manufacturing-related” or “Other” (which included those recalls without a root cause description and any root causes that could not readily be attributed to Design or Manufacturing). Any recall for which no root cause was provided by the FDA or which were linked to root cause descriptions for which the coders were not in agreement were mapped to “Other”.

In the second step, we undertook additional efforts to complement the root cause information provided by the FDA. One coder reviewed the full reason provided for a randomly drawn sample of 300 recalls (100 recalls for the Devices, Drugs, and Biologics categories, respectively) and identified keywords which could be used to clearly identify a recall as being design- or manufacturing-related. In total, the coder identified 30 keywords/ phrases, such as “Design change” (Design) or “Incorrectly assembled” (Manufacturing). We then proceeded to run a keyword search on all recalls which remained in the “Other” category after the first step. While we report results based on the full two-stage process in section 2.4.2, the results can also be replicated using the first step of the described process only.

Chapter 3

"Better safe than sorry": How CEO neuroticism influences product recall timing

Extant research on product recalls provides initial insights on operational factors that affect how quickly firms decide to recall defective products from the market (i.e., time-to-recall). Additionally, research highlights the importance of strategic leaders in such decisions on recall timing. However, little is known about the psychological underpinnings of strategic leaders' decision-making in this context. We focus on CEOs—firms' most important strategic leaders—to study how neuroticism influences recall timing in the healthcare industry. Building on theory in personality psychology, we hypothesize that more neurotic CEOs are more vigilant when they face uncertainty, as they do when making recall decisions, and strive to avoid the potential threat of social disapproval from recalling late, leading to lower time-to-recall. We contextualize this argument by taking a configurational approach to personality to hypothesize that another personality trait, CEO agreeableness, increases the salience of the threat of social disapproval and thus positively moderates the main relationship. Further, we invoke trait activation theory to argue that more negative recent CEO media tenor provides a trait-relevant situational cue which reinforces CEOs' neurotic tendencies and thus also positively moderates the main relationship. We find broad support for our theory in analyses on a sample of 110 CEOs of U.S. healthcare firms and 1,573 product recalls they oversee.

3.1 Introduction

When firms are faced with an impending product recall due to a product defect, timing is crucial but hard to get right. Unfortunately for CEOs and their management teams, there is no ex ante obviously correct moment to trigger a recall. On the one hand, firm leaders may wish to take time to gain clarity on important aspects of the recall, such as the source of the underlying defect, its scope, and the appropriate remedy to be offered to users. The more time they take, the clearer the picture gets, allowing for more targeted recalls in which only products that are actually affected by the defect are recalled and in which the most cost-effective remedy is offered to users. This way, leaders may be able to limit costs for their firm. On the other hand, there are substantial risks related to taking more time. Not only could users of the defective products suffer potentially fatal harm while the product defect goes unaddressed (Fung and Cucho 2018), but details of the defect might become publicly known, potentially making a targeted recall infeasible. Thus, not recalling quickly enough can cause a host of adverse reputational, financial, and legal consequences for firms and social disapproval for their leaders (Van Heerde et al. 2007).

Against the backdrop of this challenge of finding the ideal moment to trigger a product recall, operations management scholars have recently begun to empirically investigate the timing of recalls. Specifically, Ni and Huang (2018) and Hora et al. (2011) identified a number of recall characteristics, such as the defect type, as antecedents of recall timing. Further, Mukherjee et al. (2022) demonstrated herding effects in recall announcements, suggesting that managers have considerable discretion as to when they issue a recall.

Particularly recent research has focused on the role of strategic leaders, i.e., firms’ top managers and directors. In this vein, Wowak et al. (2021) demonstrated that

the share of female board members influences how quickly firms issue product recalls once high-severity defects become known. Mayo et al. (2022) focused on the role of the Chief Executive Officer (CEO)—arguably the most important top manager—and showed that CEOs’ tenure influences their propensity to issue recalls.

However, while these studies highlight the importance of strategic leaders in decisions on recall timing, they focus on observable characteristics of strategic leaders rather than the psychological underpinnings of their decision-making. Specifically, extant studies on recall timing have focused on demographic proxies (i.e., gender and tenure), which provide us with important but severely limited insights into the cognitive processes and motivations of strategic leaders (Neely et al. 2020, Priem et al. 2016). Even within the broader recall literature, where research into the antecedents of recall incidence (Wowak and Boone 2015) or types of remedy offered during a recall (Liu et al. 2016) has additionally considered strategic leaders’ compensation, research has yet to use more “substantive” measures (Priem et al. 2016, p. 944) of the “underlying psychological characteristics” that drive strategic leaders’ decision-making (Colbert et al. 2014, p. 354), perhaps most critically their personality.

This absence of research on strategic leaders’ personality and recalls is surprising because research in other fields has long established that personality in general, and strategic leaders’ personality in particular, is highly consequential. In the psychology literature, a large body of research demonstrates that an individual’s personality is a dominant predictor of their cognitions and behaviors (Schneider and Smith 2004). Correspondingly, in the strategic leadership literature, numerous studies focusing specifically on the personality of strategic leaders (e.g., Herrmann and Nadkarni 2014, Benischke et al. 2019, Harrison et al. 2020) have found wide-ranging implications of personality for important firm-level outcomes.

We thus attempt to address this issue by studying how CEO personality influences recall timing. Specifically, we explain time-to-recall, which is the length of the period between the discovery of a defect and the announcement of a corresponding recall, as a function of CEOs’ personality.

Importantly, when top managers need to decide on recall timing, they find themselves in challenging circumstances in that they face substantial uncertainty regarding when a recall best be issued, and they risk considerable social disapproval for their handling of a recall. As already indicated above, on the one hand, if CEOs quickly initiate a recall before complete information about the defect and the affected products is available, they may overreact, for example by offering remedies that may be more extensive than would ultimately be required. This incurs unnecessary costs. If, on the other hand, CEOs wait too long and allow defective products to remain on the market, this may result in potentially fatal harm to customers or patients. This, in turn, may expose them and their firms to significant legal, financial, and reputational risks.

The personality trait of neuroticism is likely particularly critical for CEOs’ decisions on recall timing because neuroticism specifically predicts how individuals resolve situations of uncertainty and potential social disapproval. Research in psychology has linked neuroticism to a lower tolerance for and a heightened response to uncertainty (Berenbaum et al. 2008, Hirsh and Inzlicht 2008). Individuals with higher levels of neuroticism try to resolve this uncertainty in a way that makes social disapproval less likely by using “a better safe than sorry strategy” (Lommen et al. 2010, p. 1005). Given the uncertainty inherent in a recall decision for the firm and the CEO personally, as well as the potential for broad social disapproval, we hypothesize that CEOs with higher levels of neuroticism will act in a “better safe than sorry” way and trigger recalls more quickly, i.e., their firms will exhibit lower time-to-recall.

We contextualize this main effect through two moderators that go to the heart of the theoretical mechanism centering on uncertainty and social disapproval, namely how attuned the CEO is to the perceptions of others, and how favorably others currently regard the CEO. On the first point, we build on configurational approaches to personality (Shoss and Witt 2014) to theorize that CEO agreeableness (Costa and McCrae 1985), another personality trait, amplifies the salience of any threat of social disapproval. Accordingly, we hypothesize that the negative influence of CEO neuroticism on time-to-recall will be stronger for CEOs who are more agreeable. On the second point, we build on theory indicating that trait-relevant situational cues can activate dispositions arising from an individual’s personality traits (Tett and Guterman 2000). Specifically, we argue that more unfavorable recent media coverage of a CEO constitutes a trait-relevant situational cue and thus facilitates the activation of any dispositional neurotic tendencies that CEO may have. We therefore propose that the negative influence of CEO neuroticism on time-to-recall will be stronger for CEOs with more unfavorable recent media coverage.

We test our hypotheses on a sample of 110 CEOs of U.S. healthcare firms, covering 1,573 recalls over a fourteen-year period. We find robust support for our hypotheses.

Our study makes several important contributions. First, we deepen the linkages between the operations and supply chain management (OSCM) and the strategic leadership literatures, and additionally integrate personality literature (Anand and Gray 2017), by highlighting that the personality of CEOs has a substantial impact on the handling of product recalls, i.e., an important area of operational decision-making. Perhaps most importantly, in doing so, we contribute to the OSCM literature more broadly by showing that the personality of strategic leaders can have a significant impact on operational outcomes.

Second, we develop and empirically validate a theoretical model that links CEO neuroticism, a critical yet generally under-explored personality trait, to time-to-recall. We thus contribute to the emerging literature on the role of strategic leaders in product recall timing by showing how a deep-level factor, namely a key personality trait of the CEO, affects time-to-recall. With this important finding we complement and go beyond prior research on recalls that focused exclusively on leaders’ demographic proxies (e.g., Wowak et al. 2021), financial incentives (e.g., Wowak et al. 2015), or strategic considerations (e.g., Mayo et al. 2022).

Finally, we contribute by adding nuance and context to the relationship between CEO neuroticism and time-to-recall through two critical moderators. Not only does this contribute to the literature on recalls by making the boundary conditions of our theory explicit and—since both moderators are directly linked to the hypothesized main mechanism—further corroborating our primary theoretical argument, but it also allows us to make two additional contributions. For one, we advance the literature by demonstrating that multiple personality traits jointly affect managers’ actions. We demonstrate that such interactive effects, which have thus far been neglected in OSCM and management research more broadly, are critical to a holistic perspective on the consequences of managers’ personality (Shoss and Witt 2014). For another, we show that important situational cues moderate how managers’ personality traits affect firm-level outcomes. In doing so, we answer explicit calls for more research on the interplay of managers’ personality and the situations they find themselves in (Harrison et al. 2019).

3.2 Theory and Hypotheses

3.2.1 Product Recalls

A product recall is the process through which firms remove a product from market circulation because the product suffers from one or more defects that cast doubt on whether this product is safe to be used for its intended purpose (Haunschild and Rhee 2004). Recalls are initiated for reasons ranging from rather minor issues with limited impact on the end users to major problems that can cause grave bodily harm and even death (Thirumalai and Sinha 2011).

The recall process starts once a firm becomes aware of a defect, either through internal sources, such as quality monitoring, or external sources, such as adverse event reports filed by users (Mukherjee and Sinha 2018). Thereafter, firms need to perform several complex activities under uncertainty, including the identification of the source and extent of the defect, and the coordination of the removal of products from the market. Regarding the former, firms typically seek to identify the root cause of the defect and the range of products that are impacted by the defect so they can avoid overreacting and, for instance, unnecessarily recalling too many products (e.g., Smith et al. 1996). Regarding the latter, firms publicly announce the recall, and proceed to notify distributors and end-users about the remedy they are offering for affected products (Ni and Huang 2018). Recalling firms nearly always incur significant cost from remedying defects, be it partially (e.g., by repairing products) or fully (e.g., by replacing products or asking for them to be returned) (Muralidharan et al. 2019).

Because defective products may adversely impact the safety of users interacting with them, the recall process is often tightly regulated. In the U.S., for example, agencies like the Food and Drug Administration (FDA) define in which specific

circumstances firms need to recall defective products and set out detailed procedures for executing recalls. For instance, firms that are overseen by the FDA must notify the agency when they initiate a recall for a defective product, and provide detailed data on the reason for the recall, its scope, and the planned strategy to execute it (FDA 2020). Firms then provide continuous updates to and liaise with the FDA on the execution of the recall (FDA 2010). Firms cannot covertly remove a defective product from the market without declaring such action to the FDA (Hensley 2010).

Beyond the regulatory agencies like the FDA, firms and their executives are also managing and working with a range of other stakeholders. End customers and intermediaries like distributors or retailers are generally looking for clarity on the scope of products affected, information on remedies and compensation offered for those products, and, importantly, assurances that other (and which other) products are not affected (Ball et al. 2018b). Shareholders and investors are keen to see the impact on the firm’s financial performance minimized, which requires both the costs of the recall itself and public upset caused by it to be as low as possible (Liu et al. 2017).

Due to the importance and complexity of recalls, a corresponding research stream has emerged in the OSCM domain. For example, researchers have identified a number of factors associated with a higher incidence of product recalls, such as firms’ global sourcing (Steven et al. 2014), manufacturing plant utilization (Shah et al. 2017), product competition (Ball et al. 2018c), or research and development focus (Thirumalai and Sinha 2011).

Going beyond the study of such antecedents of recall incidence, and recognizing that it may not be feasible to avoid recalls altogether (Berman 1999), scholars have recently begun to more closely investigate how firms handle recall situations. For example, Ball et al. (2018a) have identified and tested a set of operational factors

that determine how managers handle recalls, and Raithel and Hock (2021) have found that firms may experience unintended adverse reputational consequences by exceeding customer expectations while handling a recall.

A key aspects of how firms handle product recalls is the timing of recalls. Researchers have identified several characteristics of the defect underlying a recall that are associated with recall timing. First, Hora et al. (2011) investigated the total time between the initial market launch and the recall of a product in the context of the U.S. toy industry. They found that firms initiate recalls more quickly following injuries or deaths to consumers, and more slowly if the defect is caused by design flaws or involves an upstream supply chain entity.

Second, Ni and Huang (2018) investigated a series of operational factors influencing time-to-recall. Time-to-recall, perhaps the most important aspect of recall timing, is the time span between a firm becoming aware of a product defect and issuing a recall for the affected product. Using data on automotive recalls, the authors proposed that a greater need for time-consuming information processing (e.g., coordinating with a supplier), operational complexity (e.g., resolving a design flaw), and lack of organizational learning (i.e., little experience managing recalls) are lengthening the time period required to complete a thorough investigation of the defect and devise a comprehensive plan to address the defect, increasing time-to-recall.

Third, Mukherjee et al. (2022) found herding effects in recall timing in the automotive industry. In their study, the authors showed that leading recalls are followed by a considerable number of other firms’ recalls in close temporal proximity. Importantly, these findings suggest that managers have considerable discretion as to when they issue a product recall.

3.2.2 Strategic Leaders and Recall Timing

The OSCM literature has a rich tradition of considering the role of managers in operations management. Research has considered a variety of managers across hierarchies, ranging from project managers (e.g., Salvador et al. 2021), purchasing managers (e.g., Goebel et al. 2018), and plant managers (e.g., Smith et al. 2009) to strategic leaders such as Chief Risk Officers (e.g., Li et al. 2022) or Chief Supply Chain Officers (e.g., Roh et al. 2016).

In line with this tradition, recent research on recall timing has focused on the particular role of strategic leaders. Wowak et al. (2021) investigated the effect of the board of directors’ gender composition, an important demographic characteristic, on recall timing. Using data from the healthcare industry, they demonstrated that a higher share of female board members reduces time-to-recall for high-severity defects. The authors argued that female board members set the tone for the firm’s managers to be more attuned to the needs of customers who could be harmed by product defects. Mayo et al. (2022) were the first to focus on the role of the CEO, arguably the most important strategic leader with a disproportionate influence on firm outcomes. Their research on consumer product recalls highlights that CEO appointments impact recall timing, as CEOs are more likely to issue product recalls early in their tenures, and less likely to do so later in their tenures as CEO.

While this research clearly establishes the importance of strategic leaders, such as CEOs, in decisions on recall timing, it focuses on observable proxies rather than the psychological underpinnings that actually drive strategic leaders’ behavior. This limits our understanding of the motivations and cognitions of strategic leaders (Neely et al. 2020, Priem et al. 2016) regarding recalls, and leads to what some researchers have come to call a “black box problem” (Hambrick 2007, p. 335).

Outside the field of OSCM, specifically in the strategic leadership literature, re-

searchers have sought to overcome this problem by considering deep-level variables like personality traits to explain a host of firm-level outcomes. In this regard, Chatterjee and Hambrick (2007), for example, demonstrated the impact of CEO narcissism on strategic dynamism and firm performance. Other researchers have shown how strategic leaders’ personality traits influence strategic change (Hermann and Nadkarni 2014), strategic risk-taking (Benischke et al. 2019), or M&A activity (Malhotra et al. 2018).

3.2.3 CEO Neuroticism and Time-to-Recall

In this paper, we explain time-to-recall, the most frequently studied aspect of recall timing, by recourse to CEO personality. We focus on CEOs as their firms’ most senior strategic leaders because it is well established that they generally have a significant impact on firm outcomes (Finkelstein et al. 2009, Quigley and Hambrick 2015). Correspondingly, research has shown that they also play an important role in the handling of product recalls (e.g., Mayo et al. 2022). In order to gain a deep understanding of the impact of CEOs on recall timing, we focus our theorizing on one uniquely relevant aspect of their personality, namely the personality trait of neuroticism.

Neuroticism, which is the tendency to exhibit emotional instability, is considered as “one of the most salient higher-order personality traits” (Widiger 2009, p. 1, Tackett and Lahey 2016). Neuroticism forms part of several important personality models, including the comprehensive Five-Factor Model (FFM) pioneered by McCrae and Costa (1987) which is dominant in modern psychology (Bainbridge et al. 2022). Neuroticism is thus a crucial dimension of an individual’s personality. The fact that prior scholars of strategic leader personality conducted studies that center exclusively on this one single trait (e.g., Ormiston et al. 2022) further underscores the importance of neuroticism.

Neuroticism is uniquely suited to explain the speed with which CEOs and their firms initiate product recalls[1]. We choose to focus on neuroticism because the theory underpinning this personality trait provides a clear directional indication of how individuals respond to and resolve the uncertainty around recall timing. Specifically, psychologists repeatedly examined how neuroticism affects individuals' responses to uncertainty, demonstrating that neurotic[2] individuals are highly averse to uncertainty and therefore react strongly to corresponding stimuli. Hirsh and Inzlicht (2008), for instance, investigated the electrophysical responses of individuals to different types of stimuli. They found that, unlike their non-neurotic counterparts, neurotic individuals were highly responsive to uncertain stimuli. In fact, the reactivity of neurotic individuals to uncertainty was even greater than their sensitivity to negative stimuli, a sensitivity neurotic individuals are well known for (Larsen and Ketelaar 1989, Rusting 1998). Together with the findings of other studies (e.g., de Bruin et al. 2007), which show that neuroticism is associated with intolerance of uncertainty, this evidence indicates that individuals with higher levels of neuroticism are more averse to uncertainty and react more strongly to it than those with lower levels of neuroticism.

This strong aversive response to uncertainty, which has been labeled "(hyper)-vigilance" (Rauthmann et al. 2015, p. 117), stems from neurotic individuals' desire to minimize potential threats (Nettle 2006). In any situation characterized by uncertainty, the exact outcomes and associated consequences are not readily predictable (Monat et al. 1972). In such a situation, neurotic individuals act strongly and impulsively to remedy the anxiety they experience (Judge et al. 2013) and to minimize potential future threats that may materialize ("the devil you don't know"), even if this means accepting some more concrete negative

[1] In Appendix 3.B, we provide an overview of all other FFM personality traits and explain why they are theoretically less suitable than neuroticism to explain time-to-recall.

[2] We follow the convention established by Bendersky and Shah (2013) and others and use the adverb "neurotic" to describe individuals with relatively higher instead of necessarily clinical levels of neuroticism.

outcomes (“the devil you know”) (Hirsh and Inzlicht 2008, p. 966). In other words, neurotic individuals respond to uncertainty with a “better safe than sorry strategy” (Lommen et al. 2010, p. 1005).

This tendency is particularly pronounced in uncertain situations that could give rise to threats of a social nature. Not only do more neurotic individuals have a heightened reactivity to social stressors compared to less neurotic individuals (Denissen and Penke 2008a), but neurotic individuals also exhibit greater reactivity to social stimuli compared to other stimuli (Denissen and Penke 2008b). Thus, when facing uncertainty with potential social threats, such as “appearing incompetent” by not taking action, they act particularly vigilantly to “avoid social disapproval” (Bendersky and Shah 2013, p. 389, Elliot and Thrash 2002).

CEOs facing an impending product recall find themselves in a situation characterized precisely by both significant uncertainty and substantial social threat. First, CEOs deciding about when to initiate the product recall must do so under uncertainty. If they act swiftly to initiate a product recall, they may not have complete information about the underlying product defect, which products are affected by it, or what the least expensive remedy for it may be. As a result, their firms will potentially incur greater costs, e.g., by recalling more products than strictly necessary or offering more extensive remedies than required (Ball et al. 2018b). For example, firms may offer to replace all products, although it might later turn out that a less costly remedial action, such as a partial repair, would have been sufficient. The option of initiating a recall quickly, although costly, immediately reduces or even fully resolves the uncertainty. Conversely, if CEOs wait longer, they continue to face uncertainty. On the one hand, they will gain additional information on the scale and scope of the problem, allowing them to deploy more targeted and thus potentially more cost-effective remedies. On the other hand, if they are unlucky, the defective product may cause potentially fatal

harm to consumers in the meantime, and expose the CEOs and their firms to substantial legal, financial, and social risk (Van Heerde et al. 2007).

Second, CEOs making a recall decision face a specific threat of broad social disapproval. This threat is particularly likely to materialize when they are slow to initiate a recall because audiences are likely to attribute adverse outcomes to the CEO (Meindl et al. 1985, Wiesenfeld et al. 2008). Coverage of two separate recalls by the same company, Baxter International, by the same journalist in a large U.S. newspaper illustrates this. In the first instance, the journalist praised then-CEO Harry Kraemer Jr. for quickly initiating a recall, even “without seeing definite proof” of an association between the product, the suspected defect, and the reported deaths (Greising 2001, p. 31). In the second case, during the CEO tenure of Robert Parkinson, the same journalist raised questions over a supposed “10-day delay” between the defects being known and the product being recalled (Greising and Japsen 2008, p. 5). This comparison illustrates the risk of social disapproval for CEOs and their firms from not recalling defective products swiftly.

Put together, we expect that CEO neuroticism affects how CEOs handle product recalls. Specifically, we propose that CEOs with higher levels of neuroticism resolve the uncertainty around when to trigger a product recall in a way that makes social disapproval from perceived inaction unlikely—by being vigilant and recalling the defective products quickly in a “better safe than sorry” approach. We formalize:

Hypothesis 3.1. *Higher levels of CEO neuroticism are negatively associated with time-to-recall.*

3.2.4 Moderating Effects of CEO Agreeableness and CEO Media Tenor

Contextualizing our main hypothesis, we propose two contingencies that capture two distinct but related influences on our main effect: how attuned the CEO is to the perceptions of others and how favorably others currently regard the CEO. These contingencies provide complementary perspectives since they operate at different levels. The first one—how attuned a CEO is to the perceptions of others—is an aspect of their personality, while the second one—how favorably others currently regard the CEO—is a characteristic of the respective situation.

In our theoretical model, these contingencies take the form of two moderators. For the first moderator, we build upon configurational approaches to personality (Shoss and Witt 2014) by considering a trait-interaction effect of CEO neuroticism with *CEO agreeableness*, another FFM personality trait, which particularly captures how attuned CEOs are to the perceptions of others. For the second moderator, we build on the logic of trait activation theory (Tett and Guterman 2000) and select the valence of recent media coverage of a CEO, i.e., *CEO media tenor*, as the moderating variable to assess how favorably or unfavorably relevant others currently view that CEO.

CEO agreeableness. Agreeableness is a personality trait that describes interpersonal behavior with a focus on the quality of interactions with others “along a continuum from compassion to antagonism” (Costa and McCrae 1985, p. 2). This focus on social quality sets it apart from all other personality traits in the five factor model (Costa et al. 1991).

Agreeableness has been associated with a range of prosocial motives and behaviors. For example, more agreeable individuals have a higher “desire to maintain positive relations with others” (Gleason et al. 2004, p. 45), are more likely to help others

(Graziano and Eisenberg 1997), and place greater emphasis on their likeability and their social reputation (Graziano and Tobin 2016, Hogan et al. 1996). Conversely, less agreeable individuals are more antagonistic and less concerned about their interpersonal relationships (Graziano and Eisenberg 1997).

Research shows that more agreeable individuals are more attuned to perceptions of others. Individuals with higher levels of agreeableness are more equipped to take the perspective of others (known as perspective taking), and to feel emphatic concern for the needs of others (Graziano and Tobin 2009, 2016). In fact, this “propensity or motivation to attend to the mental states of others may be central to the personality dimension of Agreeableness” (Nettle and Liddle 2008, p. 323).

Accordingly, we expect more agreeable CEOs to be more aware of the needs of the many audiences and stakeholders of a product recall, such as customers, regulators, distributors, and regulators. We expect that this ability to be more attuned to the perception of others will increase the salience of the threat of social disapproval that comes with recall decisions and that neurotic CEOs are keen to minimize. Conversely, more disagreeable CEOs, who are less inclined to consider the perspective of others in their handling of a product recall, will be less likely to perceive a significant social threat from not taking swift action. Put together, this leads us to hypothesize an interactive effect of CEO neuroticism and CEO agreeableness:

Hypothesis 3.2. *The negative influence of CEO neuroticism on time-to-recall will be stronger for CEOs with higher levels of agreeableness than for CEOs with lower levels of agreeableness.*

CEO media tenor. CEOs are carefully attuned to how others, specifically the media, perceive them. For one, CEOs often attract considerable evaluative media coverage. Notably, they are not merely frequently mentioned in the media, but they are specifically subject to constant evaluations (Hayward et al. 2004). These

evaluations of CEOs range from them being hailed as “saviors” (Lovelace et al. 2018, p. 420) to being cast as villains who are “singled out and penalized far out of proportion to their culpability for their companies’ problems” (Wiesenfeld et al. 2008, p. 231).

For another, continuously negative media evaluations can have decidedly negative effects on CEOs. Such evaluations can, for example, harm CEOs’ social standing, causing them to lose “prestigious directorships” and rendering them unable to find suitable executive positions (Wiesenfeld et al. 2008, p. 231). They might also lead to CEOs’ outright dismissal, as boards tend to attach great weight to the assessments of external evaluators when making judgments about CEOs (Khurana 2002) because they deem external sources particularly credible (Menon and Pfeffer 2003). In fact, negative media coverage of CEOs has been linked to changes to CEOs’ future compensation (Vergne et al. 2018) and with a higher chance of dismissal (Bednar 2012). Consequently, CEOs are likely to be keenly aware of their portrayal in the media, making their media coverage a salient situational factor that potentially impacts their decision-making.

A host of research in personality psychology shows that the dispositions arising from an individual’s personality are particularly consequential in situations where they match the situational conditions (Bowers 1973, Higgins 2000). In fact, prior research has explicitly shown that such a “match between trait and state motivational cues” (Tamir 2005, p. 455) is relevant for the expression of neuroticism. We follow this line of thought and build on trait activation theory, which posits that trait-relevant situational cues trigger the behavioral expression of personality dispositions (Tett and Guterman 2000). This is to say that if the situation holds cues that are consistent with a particular trait, trait activation theory predicts that this would facilitate the expression of the dispositions associated with that trait.

We thus expect that a match between situational factors and CEO personality affects how CEOs make decisions regarding product recalls. In our main hypothesis 3.1, we argued that more neurotic CEOs act vigilantly by recalling defective products more quickly, at least partially to avoid social threats. We propose that unfavorable media coverage is a situational cue that is trait-relevant for a CEO high in neuroticism, i.e., an individual prone to “anxious monitoring of social exclusion” (Rauthmann et al. 2015, p. 117). When a neurotic CEO experiences negative media coverage, this should strengthen the expression of their neurotic dispositions. Consequently, we expect the relationship between CEO neuroticism and time-to-recall to be more pronounced for CEOs who recently experienced more unfavorable media coverage than for CEOs with more favorable recent media coverage. Formally put:

Hypothesis 3.3. *The influence of CEO neuroticism on time-to-recall will be stronger for CEOs who recently experienced relatively unfavorable media coverage than for CEOs who recently experienced relatively favorable media coverage.*

3.3 Data and Methodology

3.3.1 Sample

Our study is based on a sample of recalls made by publicly traded healthcare firms listed in the S&P 1500. We focused on the U.S. healthcare industry as it provides us with a uniform regulatory context. Specifically, the FDA provides clear guidelines and legally binding regulations to firms on many key processes, including product recalls. We collected data on all firms listed in the S&P 1500 Health Care Sector index (GICS 35) for at least three fiscal years between 2005 and 2018, a fourteen-year period. We excluded sub-industries with offerings that

cannot be recalled, such as the operation of healthcare facilities like hospitals (see Appendix 3.A for details). We retrieved the names of the CEOs leading those firms from Execucomp.

Next, we compiled a list of all product recall events for the sample period, using the FDA’s website. We dropped recalls in four product categories (cosmetics, food, tobacco, and veterinary products) as these do not relate to the healthcare sector. For the remaining recalls, we manually complemented the data with additional attributes, including those used for our dependent variable, through several Freedom of Information Act requests and interactions with the FDA.

We then proceeded to match recalls to CEOs based on the name of the “most responsible individual” designated for each recall, as reported by the FDA, and retained only recalls that listed the respective CEO in this capacity. When a CEO name matched but the firm name of the recalling firm did not match the name of the CEO’s firm in our sample, we validated that the recalling firm was a subsidiary of the sample firm using its annual report filings with the Securities and Exchange Commission. We could validate all matched recalls in this way. Our final sample comprises 1,573 recall events in 79 firms between 2005 and 2018. These recalls were overseen by 110 CEOs.

3.3.2 Variable Definitions

Dependent Variable. We followed Wowak et al. (2021) and defined time-to-recall as the number of days that elapsed between a firm becoming aware of a quality defect and the firm initiating a recall for it. We retrieved both dates, the “firm awareness date” and the “recall initiation date,” from the respective reports provided by the FDA. A smaller time-to-recall value indicates a faster decision to issue a recall. Since time-to-recall is a highly skewed variable, we used its natural

logarithm as our dependent variable.

Independent Variables. We assessed CEO personality using the extensively validated open-language tool developed by Harrison et al. (2019, 2020), which is experiencing increasing popularity in the management literature. The tool uses a machine learning approach to infer CEOs’ personality from the transcripts of spontaneous statements made by the respective CEOs during quarterly earnings conference calls with capital market analysts. We describe the data-gathering process as well as the tool and its validity in detail in Appendix 3.C and outline key steps below.

We first retrieved all available transcripts of quarterly earnings conference calls for the sample firms during the sample time frame from Fair Disclosure wires on LexisNexis and Thomson StreetEvents. In total, we gathered more than 8,000 transcripts. We then processed each transcript by attributing every statement made during these calls to a specific individual.

In line with prior research (e.g., Malhotra et al. 2018), we proceeded to filter all statements that were made by a firm’s CEO in direct response to an analyst’s question. We collated all statements made by each CEO across our entire sampling frame and scored each of the five personality traits of the FFM for each CEO using Harrison et al.’s open-language tool (2019, 2020). Basing the personality assessment on all calls in the sampling frame, and thus assuming time-invariance of the FFM traits, is in line with prior studies in the management literature (e.g., Harrison et al. 2020) and the broader personality literature, which indicates that these traits are stable in adult life (Roberts and DelVecchio 2000).

Our primary independent variable, CEO neuroticism, was scored on a 7-point scale. A higher score indicates a higher level of neuroticism. Our first moderator, CEO agreeableness, was measured using the exact same process as used for CEO neuroticism. It was thus also scored on a 7-point scale where a higher value

indicates a higher level of agreeableness.

Our second moderator, CEO media tenor, gauges how positive (or negative) the media coverage of a CEO was before the firm first became aware of the focal defect (i.e., the firm awareness date). Assessing the media coverage immediately before the defect became known to the firm ensures that the focal media coverage is unaffected by the defect or the recall itself. To compute CEO media tenor, we followed prior studies (e.g., Love et al. 2017). We began by retrieving articles covering our sample firms from three eminent U.S.-based newspapers—*The Wall Street Journal*, *The New York Times*, and *The Washington Post*. We choose these three newspapers specifically as they are considered “prestige media” (Deephouse and Suchman 2008, p. 56). Educated and sophisticated audiences like the corporate elite are “closely attuned” to the coverage in these kinds of outlets (Li et al. 2021, p. 1998), making it likely that CEOs are aware of their own coverage in them. In total, we retrieved 1,268 articles published during the 30-day period prior to the firm awareness date of each recall. Following Love et al. (2017), we focused on articles that cover the CEOs, not just the firms they lead, by retaining only those articles that mention the CEO by name at least three times.

We precisely followed Pfarrer et al. (2010) and Love et al. (2017) in aggregating the coverage into a single continuous measure per period. First, we assessed the degree of positive and negative affective language of each article using the Linguistic Inquiry and Word Count tool (Pennebaker et al. 2001). Using this tool and its validated dictionaries ensures that we obtain reliable measurements for the valence of these texts. Second, we created ratios of the positive and negative affective content, respectively, to the total affective content in the article. Third, we coded each article based on these ratios. Again, we followed Pfarrer et al. (2010) and designated an article as positive if the positive ratio was at least 0.6, and as negative if the negative ratio was at least 0.6. All remaining articles were

coded as neutral. Fourth, we aggregated the coverage across all articles during the period using the Janis-Fadner coefficient of imbalance (Janis and Fadner 1943). The resulting CEO media tenor score can range from -1 (all negative coverage) to 1 (all positive coverage), with zero indicating neutral or no coverage.

Control Variables. We controlled for a variety of recall-, firm- and CEO-specific factors that may have an impact on the recall process. Our models include dummy variables for the year in which a focal recall took place to control for possible changes in the overall business and regulatory environments. We also included dummies for the type of product being recalled to capture differences in product complexities and specificities in regulatory processes. This classification was provided by the FDA. Our sample included 1,263 events of medical device recalls, 288 recall events concerning pharmaceuticals, and 22 biologics recall events. We used the most common class, medical devices, as the base class. Furthermore, we included a control variable to capture the severity of the product defect. We use the official recall classification of the FDA, which assigns each recall to one of three classes, as our measure for defect severity. We used the most common class (class II) as the base class. Finally, we also include a dummy variable indicating whether the products being recalled were manufactured in-house at the recalling firm, or by another firm (e.g., Steven et al. 2014). We obtained this information directly from the FDA.

Additionally, we controlled for a host of possibly confounding firm-level factors, mostly using data from Compustat. In line with previous research on recalls (e.g., Thirumalai and Sinha 2011), we controlled for firm size (natural log of total assets in USD millions), R&D intensity (R&D expenditure divided by sales), and firm profitability (return on assets). We also included total debt/assets to account for firm risk-taking (Cronqvist et al. 2012, Malhotra et al. 2018) and immediate slack (working capital divided by sales) to control for the availability

of spare resources (Finkelstein and Hambrick 1990). Like prior recall research, we controlled for recall experience (number of recalls in the prior fiscal year). In light of the recent findings by Wowak et al. (2021), we also controlled for the female board ratio (share of female board members) and included a control variable for board independence (percentage of non-executive directors) using corresponding data from BoardEx.

Finally, we also included controls for several CEO attributes using Execucomp data. We measured CEO tenure as the number of years a CEO has held their position at the focal firm (e.g., Wowak et al. 2015). We used proxies for CEOs’ power to measure their influence over their organizations (Malhotra et al. 2018). Specifically, we controlled for CEO duality, a binary variable indicating if the CEO is also the chairperson of the board, and CEO stock ownership, a continuous measure of the percentage of voting shares owned by the CEO. To capture possible confounding influences from other personality traits, we also included the remaining FFM personality traits CEO openness, CEO conscientiousness, and CEO extraversion, which were measured exactly like CEO neuroticism and CEO agreeableness.

3.3.3 Model Specification

We followed prior research on product recalls (e.g., Thirumalai and Sinha 2011, Wowak et al. 2021) and estimated firm fixed effects models to test our hypotheses. These fixed effects models help us to address unobserved heterogeneity arising from potentially omitted time-invariant firm characteristics (Wooldridge 2013). We used Huber-White robust standard errors, clustered at the firm level, to avoid potential issues arising from heteroskedasticity (Wooldridge 2002). Finally, to enhance interpretability, we centered and standardized all continuous independent and control variables, except for CEO media tenor, as this variable already has a readily interpretable scale.

3.4 Results

3.4.1 Test of Hypotheses

Table 3.1 presents descriptive statistics and correlations. The untransformed mean of time-to-recall of 80.33 days is in line with research using a similar context, for example the 95 days reported as mean time-to-recall by Wowak et al. (2021). As expected, we observe several significant correlations between various FFM personality dimensions. The signs of these correlations are in line with those found in meta-reviews of the FFM dimensions in general (e.g., van der Linden et al. 2010) and prior CEO personality research in particular (e.g., Harrison et al. 2020).

Despite substantial correlations between some of our independent variables, multicollinearity is unlikely to negatively affect the statistical validity of our findings. Firstly, we apply a widely adopted methodological test for multicollinearity by calculating variance inflation factors (VIF) for our main Model 2. The mean VIF was 2.9 and all individual VIFs were below 10. This common indicator supports that multicollinearity does not materially affect our inferences (Kutner et al. 2005). Even for a simplified regression model using only the personality traits as independent variables, all individual VIFs are clearly below 10. In this constellation, the mean VIF is 3.8, suggesting that these correlations do not materially affect our results.

Table 3.1: Summary statistics and correlations

Variables	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(1) Time-to-recall	80.33	142.19	1.00																	
(2) Total assets	18.16	24.32	0.03	1.00																
(3) RD intensity	0.08	0.06	-0.06	0.14	1.00															
(4) Return on assets	0.05	0.07	0.03	0.09	-0.18	1.00														
(5) Total debt / assets	0.26	0.12	-0.07	0.36	-0.01	-0.40	1.00													
(6) Immediate slack	0.36	0.21	-0.06	-0.10	0.03	0.11	-0.15	1.00												
(7) Recall experience	21.39	17.07	0.11	0.55	-0.11	0.01	0.27	-0.11	1.00											
(8) Female board ratio	0.18	0.08	-0.02	0.60	0.16	-0.23	0.38	-0.05	0.23	1.00										
(9) Board independence	0.90	0.05	0.11	0.42	0.05	-0.06	0.28	-0.03	0.32	0.40	1.00									
(10) CEO tenure	5.01	5.13	-0.10	-0.16	0.02	0.05	-0.03	-0.08	-0.16	-0.13	-0.28	1.00								
(11) CEO duality	0.50	0.50	-0.04	0.35	-0.02	0.28	0.03	-0.09	0.06	0.17	0.04	0.23	1.00							
(12) CEO stock ownership	0.36	1.07	-0.07	-0.44	-0.00	-0.00	-0.11	0.11	-0.29	-0.30	-0.49	0.50	0.05	1.00						
(13) CEO media tenor	0.02	0.15	0.02	0.11	0.13	-0.01	0.03	0.04	0.03	0.08	0.02	-0.02	0.01	-0.04	1.00					
(14) CEO openness	4.88	0.32	0.01	0.28	0.15	0.10	0.01	0.16	0.04	0.17	0.23	-0.19	0.28	-0.19	-0.00	1.00				
(15) CEO conscientiousness	5.33	0.25	0.03	0.51	0.07	-0.02	0.25	0.15	0.28	0.29	0.27	-0.31	0.14	-0.28	0.05	0.76	1.00			
(16) CEO extraversion	5.24	0.32	0.09	0.46	-0.11	-0.09	0.25	-0.04	0.34	0.24	0.28	-0.19	0.16	-0.29	0.04	0.52	0.64	1.00		
(17) CEO agreeableness	4.36	0.45	-0.02	0.39	0.22	0.20	-0.00	0.05	-0.00	0.26	0.27	-0.21	0.35	-0.23	0.01	0.88	0.68	0.40	1.00	
(18) CEO neuroticism	3.13	0.33	0.01	-0.45	-0.32	-0.27	0.09	-0.13	-0.07	-0.23	-0.24	0.18	-0.26	0.18	-0.05	-0.77	-0.62	-0.27	-0.86	1.00

Unstandardized and untransformed summary statistics (mean, standard deviation) shown for interpretability. Total assets shown as USD billions.
Dummy variables for categorical or binary variables (product type, recall class, manufactured in-house) and year not shown.
Correlations greater than |0.05| are significant at $p < 0.05$.

Secondly, Lindner et al. (2020, p. 296) recently demonstrated that multicollinearity generally “does not introduce bias” to regression models. Put differently, this means that potential multicollinearity (which does not seem to be a significant concern based on the analysis of the VIF above) is unlikely to bias the results of our regression analysis. The authors additionally explain that “variance inflation factors are indicators of standard errors that are too large, not too small” (Lindner et al. 2020, p. 296). In other words, the significance levels reported in our models may be conservative, but not overly aggressive.

In Table 3.2, we present the results of our fixed effects regression models. Model 1 includes all controls and Model 2 adds our focal independent variable, CEO neuroticism. In Hypothesis 3.1, we posited that higher levels of CEO neuroticism are negatively associated with time-to-recall. The coefficient for CEO neuroticism in Model 2 is negative (β = -0.656) and significant ($p < 0.05$), providing support for Hypothesis 3.1. A CEO with a level of neuroticism of one standard deviation above the mean corresponds to an acceleration of recall initiation by nearly 50%, or 39 days.

Models 3 and 4 include the interaction terms for CEO neuroticism with CEO agreeableness and CEO media tenor, respectively. The coefficients for both interaction terms are in line with our theoretical predictions. The coefficient for the interaction with CEO agreeableness in Model 3 is negative (β = -0.261) and significant ($p < 0.01$), indicating that higher levels of CEO agreeableness strengthen the negative effect of CEO neuroticism on time-to-recall, and supporting Hypothesis 3.2. The coefficient for the interaction with CEO media tenor in Model 4 is positive ($\beta = 0.288$), which indicates that the negative effect of CEO neuroticism on time-to-recall is stronger when CEO media tenor is more negative, and weaker when CEO media tenor is more positive. This finding is marginally significant ($p < 0.1$) and thus provides some support for Hypothesis 3.3.

Table 3.2: Fixed e□ects regression on time-to-recall (natural log-transformed)

	(1)	(2)	(3)	(4)	(5)
Constant	3.620	3.631	3.442	3.645	3.452
	(0.320)	(0.319)	(0.319)	(0.320)	(0.323)
Product type: Drugs	-0.899	-0.904	-0.912	-0.886	-0.893
	(0.244)	(0.244)	(0.246)	(0.234)	(0.235)
Product type: Biologics	-0.124	-0.117	-0.118	-0.120	-0.120
	(0.374)	(0.372)	(0.383)	(0.373)	(0.384)
Recall classification: Class 1	-0.023	-0.021	-0.020	-0.020	-0.020
	(0.165)	(0.167)	(0.164)	(0.165)	(0.162)
Recall classification: Class 3	-0.149	-0.176	-0.172	-0.173	-0.169
	(0.180)	(0.185)	(0.186)	(0.185)	(0.186)
Manufactured in-house	0.016	0.020	0.011	0.023	0.014
	(0.133)	(0.131)	(0.132)	(0.129)	(0.130)
Total assets	-0.211	-0.147	-0.078	-0.157	-0.087
	(0.214)	(0.180)	(0.155)	(0.175)	(0.153)
R&D intensity	0.050	0.055	0.056	0.057	0.059
	(0.069)	(0.066)	(0.071)	(0.067)	(0.073)
Return on assets	-0.070	-0.099	-0.087	-0.100	-0.087
	(0.064)	(0.066)	(0.065)	(0.066)	(0.066)
Total debts/assets	0.003	-0.055	-0.060	-0.062	-0.067
	(0.100)	(0.102)	(0.101)	(0.102)	(0.101)
Immediate slack	-0.186	-0.155	-0.133^{+}	-0.154	-0.131^{+}
	(0.073)	(0.069)	(0.068)	(0.068)	(0.068)
Recall experience	0.052	0.057	0.012	0.055	0.009
	(0.071)	(0.069)	(0.069)	(0.069)	(0.068)
Female board ratio	0.013	-0.006	-0.006	-0.018	-0.019
	(0.079)	(0.080)	(0.081)	(0.081)	(0.082)
Board independence	0.126	0.114	0.139	0.117	0.142
	(0.093)	(0.089)	(0.087)	(0.089)	(0.087)
CEO tenure	0.035	0.071	0.035	0.086	0.050
	(0.071)	(0.073)	(0.085)	(0.077)	(0.090)
CEO duality	-0.202	-0.132	-0.096	-0.138	-0.102
	(0.186)	(0.175)	(0.148)	(0.175)	(0.148)
CEO stock ownership	0.324	0.266	0.184	0.261	0.177
	(0.130)	(0.123)	(0.118)	(0.125)	(0.120)
CEO media tenor	0.151	0.102	0.093	0.156	0.151
	(0.258)	(0.258)	(0.252)	(0.193)	(0.186)
CEO openness	-0.099	-0.096	0.226	-0.112	0.216
	(0.219)	(0.238)	(0.274)	(0.244)	(0.279)
CEO conscientiousness	0.282	0.394^{+}	0.102	0.415^{+}	0.117
	(0.213)	(0.226)	(0.277)	(0.238)	(0.282)
CEO extraversion	-0.057	0.024	-0.060	0.078	-0.004
	(0.159)	(0.140)	(0.143)	(0.150)	(0.152)
CEO agreeableness	0.018	-0.523	-0.373^{+}	-0.606	-0.459^{+}
	(0.176)	(0.239)	(0.202)	(0.277)	(0.232)
CEO neuroticism		-0.656	-0.359	-0.750	-0.454^{+}
		(0.263)	(0.231)	(0.286)	(0.254)
CEO neuroticism × CEO agreeableness			-0.261		-0.267
			(0.094)		(0.095)
CEO neuroticism × CEO media tenor				0.288^{+}	0.310
				(0.155)	(0.138)
Firm fixed-effects			Included		
Year dummy variables			Included		
Number of observations (recalls)			1,573		
Number of firms			79		

Robust standard errors in parentheses. Year dummy variables included but not shown.
$^{+}$ $p < 0.10$, $^{□}$ $p < 0.05$, $^{□□}$ $p < 0.01$, $^{□□□}$ $p < 0.001$

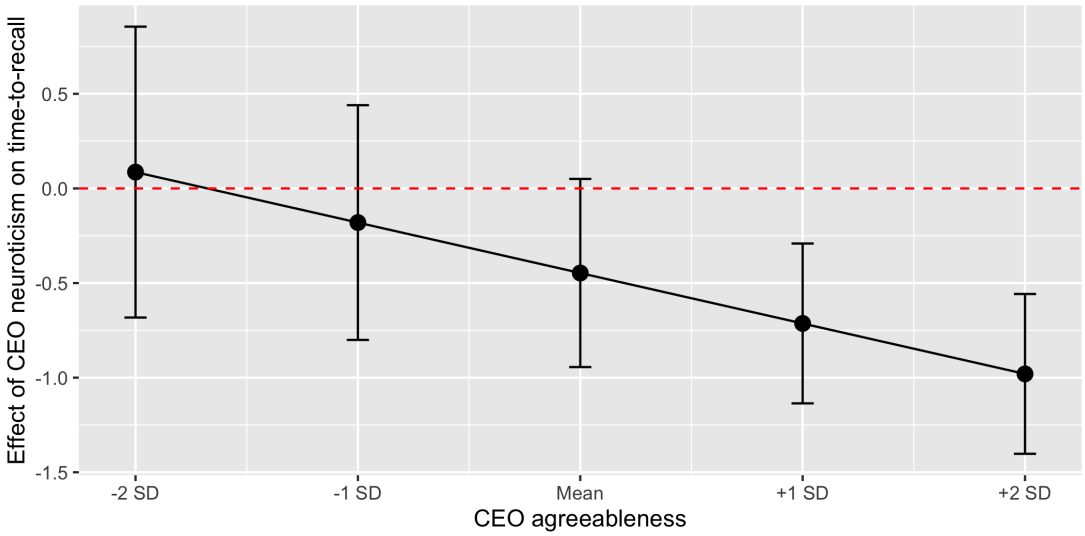

Figure 3.1: Average marginal e□ect of CEO neuroticism on time-to-recall as CEO agreeableness varies in model 5, including 95% confidence interval.

Model 5 integrates both interactions within a single saturated model. In this model, the signs of both interaction terms are as hypothesized (β = -0.267 for CEO agreeableness, $\beta = 0.310$ for CEO media tenor) and both coefficients are significant ($p < 0.01$ and $p < 0.05$, respectively). This analysis thus provides further support for hypotheses 3.2 and 3.3.

In Figures 3.1 and 3.2, we show the marginal effect plots for both interactions from the saturated model (Model 5), along with the 95% confidence intervals. We provide the marginal effects of these interactions from the individual interaction models (Model 3 and 4), which are very similar to those in Figure 1, in figures 3.4 and 3.5 in Appendix 3.D. The plots indicate that there is no significant effect of CEO neuroticism on time-to-recall at or below mean levels of CEO agreeableness, and for positive or neutral media coverage. However, the plots show that above-average levels of agreeableness and negative media coverage indeed substantially strengthen the hypothesized main effect. Accordingly, we see that both variables moderate our main effect, even when taking into account the respective other interaction effect. Put together, these findings provide strong support for hypotheses 3.2 and 3.3.

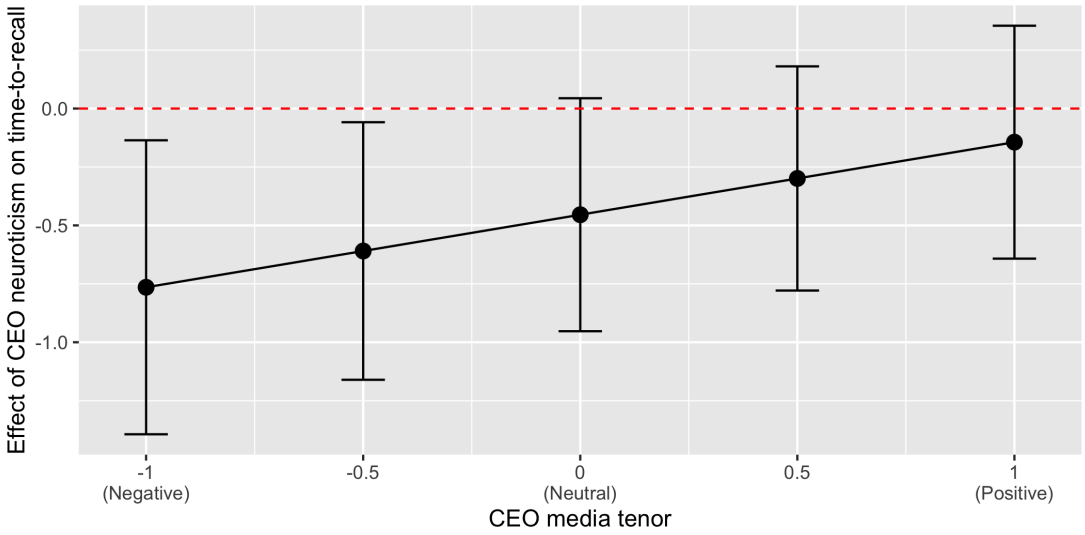

Figure 3.2: Average marginal e□ect of CEO neuroticism on time-to-recall as CEO media tenor varies in model 5, including 95% confidence interval.

3.4.2 Robustness Checks

Sensitivity of results to bias and omitted variables. In order to test the sensitivity of our inferences to bias and omitted variables, we used two methods that have recently been increasingly adopted in the management literature (e.g., Harrison et al. 2018, Quigley et al. 2020). We focused these analyses on the main effect (Hypothesis 3.1), which we analyzed in Model 2, as neither method has yet been formalized for interactive effects (Busenbark et al. 2022).

The first method is an analysis of the robustness of inference to replacement, which tests the robustness of our inference to various sources of bias (Busenbark et al. 2021). To fully invalidate the significant inference of CEO neuroticism on time-to-recall (i.e., $p > 0.1$), 34% of the estimate would have to be due to bias. Accordingly, more than 500 observations would have to be replaced with cases for which there is an effect greater than zero (i.e., a positive as opposed to negative association between CEO neuroticism and time-to-recall). This is a considerable number of cases for which the relationship would not only need to be weaker than prescribed, but even be opposite to what we have theoretically

derived and empirically found. While there is no established benchmark for a high level of robustness to replacement (Frank et al. 2023), 34% is comparable to the robustness of findings in some other recent studies of CEOs (e.g., 33% in Andrei et al. (2022), 23%–56% in Pollock et al. (2023) and 36%–44% in Harrison et al. (2022)). This analysis thus bolsters our confidence in our findings.

Although our models include a considerable number of relevant control variables capturing various aspects of the recall, the firm, and the CEO, we also assessed the potential impact of an omitted variable by determining the impact threshold of a confounding variable (ITCV). The ITCV indicates the minimum impact an omitted variable would need to have to render an inference insignificant (Frank 2000). In line with prior research (Busenbark et al. 2022, Gamache and McNamara 2019), we thus evaluate the impact required to invalidate our inference (i.e., $p >$ 0.1). For Model 2, the corresponding ITCV is -0.021. Put differently, an omitted variable would have to be correlated at 0.146 with time-to-recall and at -0.146 with CEO neuroticism, or the other way around, to invalidate the inference.

Since a potentially omitted variable is, by definition, not known, the ITCV score is best understood in relation to existing control variables. Accordingly, we compare the impact measure of our existing control variables with the ITCV of a potential confounding variable (Busenbark et al. 2022). Since the association between CEO neuroticism (our independent variable) and time-to-recall (our dependent variable) is negative, only variables which are negatively correlated with either our dependent or independent variable, but not both, could weaken or invalidate the significance of our findings (Frank 2000, Larcker and Rusticus 2010). In all other cases, the omitted variable would strengthen rather than weaken the significance of our findings. Of the 21 variables included in our main model, only two variables – CEO extraversion and board independence – have this property and additionally have a marginally stronger negative observed impact than the ITCV. While this

result of a blanket comparison to all control variables is encouraging, a likely more informative comparison is one between the ITCV and a specific canonical control variable in the recall literature, like recall experience. This comparison shows that the impact of an omitted confounding variable would need to be almost three times bigger than the impact of this widely-used control to invalidate our inference. Overall, this analysis thus suggests that our results are robust.

Model Specification. To test the robustness of our findings to the specification of the empirical model, we tested our hypotheses using generalized estimating equations (GEE) with a Gaussian distribution and an identity link function, which is commonly used with Gaussian distributions. We used an exchangeable correlation matrix and again calculated robust standard errors.

Overall, the results of this analysis (which can be found in Table 3.3) again show that our findings are robust. The signs of the coefficients for our main effect and interaction effects are identical to those in our fixed effects models. The coefficient for CEO neuroticism remains statistically significant at the $p < 0.05$ level. The coefficient for the interaction term with CEO agreeableness also remains consistently significant. While the coefficient for the interaction term with CEO media tenor occasionally loses significance in some models, the corresponding margin plots are essentially identical to those we reported for the fixed effects models. Consequently, the GEE models provide further substantial support for our findings.

Table 3.3: GEE models on time-to-recall (natural log-transformed)

	(1)	(2)	(3)	(4)	(5)
Constant	3.447	3.405	3.323	3.414	3.332
	(0.300)	(0.303)	(0.275)	(0.302)	(0.272)
Product type: Drugs	-0.809	-0.872	-0.897	-0.859	-0.883
	(0.193)	(0.194)	(0.202)	(0.186)	(0.195)
Product type: Biologics	-0.036	-0.041	-0.042	-0.046	-0.048
	(0.357)	(0.351)	(0.365)	(0.349)	(0.364)
Recall classification: Class 1	-0.018	-0.013	-0.026	-0.012	-0.025
	(0.158)	(0.158)	(0.155)	(0.157)	(0.154)
Recall classification: Class 3	-0.065	-0.086	-0.091	-0.080	-0.084
	(0.177)	(0.177)	(0.176)	(0.177)	(0.176)
Manufactured in-house	-0.021	-0.009	-0.004	-0.006	-0.000
	(0.127)	(0.129)	(0.127)	(0.127)	(0.126)
Total assets	-0.026	-0.109	-0.065	-0.112	-0.067
	(0.111)	(0.115)	(0.107)	(0.115)	(0.108)
R&D intensity	0.038	0.015	0.001	0.014	-0.001
	(0.065)	(0.053)	(0.048)	(0.054)	(0.048)
Return on assets	-0.061	-0.084	-0.081	-0.082	-0.079
	(0.053)	(0.054)	(0.057)	(0.054)	(0.058)
Total debts/assets	-0.046	-0.031	-0.011	-0.034	-0.014
	(0.072)	(0.073)	(0.074)	(0.073)	(0.073)
Immediate slack	-0.168	-0.152	-0.145	-0.155	-0.147
	(0.060)	(0.056)	(0.055)	(0.055)	(0.055)
Recall experience	0.041	0.048	0.012	0.048	0.011
	(0.068)	(0.070)	(0.064)	(0.070)	(0.064)
Female board ratio	-0.010	-0.008	-0.016	-0.013	-0.021
	(0.066)	(0.068)	(0.066)	(0.069)	(0.066)
Board independence	0.184	0.165	0.198	0.166	0.201
	(0.066)	(0.066)	(0.064)	(0.066)	(0.064)
CEO tenure	-0.007	0.030	0.002	0.034	0.006
	(0.047)	(0.051)	(0.057)	(0.052)	(0.058)
CEO duality	-0.195	-0.164	-0.161	-0.166	-0.164
	(0.148)	(0.146)	(0.135)	(0.147)	(0.135)
CEO stock ownership	0.228	0.202	0.215	0.200	0.212
	(0.061)	(0.059)	(0.058)	(0.058)	(0.057)
CEO media tenor	0.169	0.128	0.121	0.189	0.191
	(0.248)	(0.242)	(0.237)	(0.195)	(0.182)
CEO openness	0.078	-0.051	0.084	-0.064	0.073
	(0.171)	(0.182)	(0.166)	(0.185)	(0.166)
CEO conscientiousness	0.103	0.171	0.055	0.173	0.055
	(0.172)	(0.177)	(0.153)	(0.181)	(0.155)
CEO extraversion	-0.023	0.058	0.078	0.078	0.102
	(0.100)	(0.093)	(0.082)	(0.095)	(0.086)
CEO agreeableness	-0.019	-0.293	-0.291^{+}	-0.307^{+}	-0.309
	(0.185)	(0.178)	(0.149)	(0.186)	(0.155)
CEO neuroticism		-0.406	-0.266^{+}	-0.432	-0.293
		(0.163)	(0.141)	(0.169)	(0.143)
CEO neuroticism × CEO agreeableness			-0.148		-0.152
			(0.045)		(0.045)
CEO neuroticism × CEO media tenor				0.246	0.285^{+}
				(0.183)	(0.153)
Year dummy variables			Included		
Number of observations (recalls)			1,573		
Number of firms			79		

Robust standard errors in parentheses. Year dummy variables included but not shown.
$^{+}$ $p < 0.10$, $^{\square}$ $p < 0.05$, $^{\square\square}$ $p < 0.01$, $^{\square\square\square}$ $p < 0.001$

3.5 Discussion

3.5.1 Implications

In this study, we sought to examine the role of CEO personality, specifically CEO neuroticism, in decisions on recall timing. Based on research in personality psychology, we argued that CEOs with greater levels of neuroticism would take action more quickly in a recall situation to resolve the inherent uncertainty and fend off any threat of social disapproval. In so doing, they would follow a “better safe than sorry strategy” (Lommen et al. 2010, p. 1005) and initiate recalls earlier. Our empirical findings—based on analyses of a large sample of 1,573 recalls in the U.S. healthcare industry—provide compelling evidence in support of this hypothesis.

With this finding, we make important contributions to two literature streams and strengthen the linkages between them. By highlighting that the personalities of CEOs substantially impact how product recalls are handled, we deepen the linkages between the OSCM and strategic leadership literatures (Anand and Gray 2017), and additionally integrate key concepts from the personality literature. We specifically contribute to the OSCM literature by demonstrating that operational outcomes can be more comprehensively understood by taking into account the personality of decision makers, in particular of strategic leaders like the CEO. Our findings are thus also a foundation for future researchers in the field to consider how the personality of operational decision makers such as plant or project managers can help us better understand other operational outcomes.

We further advance the product recall literature more specifically (Wowak and Boone 2015) by showing that CEOs’ personality has a substantial influence on recall timing. In particular, we develop and empirically validate a theoretical

model that links CEO neuroticism, a critical yet under-explored personality trait, to time-to-recall, a critical aspect of recall timing. We thus contribute to the emerging literature on the role of strategic leaders in recall timing by showing how a deep-level factor, specifically a key personality trait of the CEO, affects time-to-recall. With this important finding, we move beyond prior findings showing effects of demographic proxies (e.g., Wowak et al. 2021), financial incentives (e.g., Wowak et al. 2015), and strategic considerations (e.g., Mayo et al. 2022) of strategic leaders on recall timing.

Based on our findings on CEO neuroticism, we are also able to make a contribution to the wider strategic leadership and management literatures. Extant literature has linked CEO neuroticism primarily to undesirable outcomes (e.g., Peterson et al. 2003). We challenge this prevailing view of neuroticism as a predominantly negative trait in CEOs. While we do not wish to go as far to argue that faster recalls are always better, or that neuroticism would even be universally desirable in CEOs, we conjecture that the particular vigilance of CEOs with higher levels of neuroticism might potentially be beneficial in a variety of high-stakes situations they face.

We further contextualized and refined the relationship between CEO neuroticism and time-to-recall through two moderators that capture how attuned a CEO is to the perception of stakeholders, and how others currently regard the focal CEO. Our theorizing around these moderators ties in directly with the hypothesized main effect, and our empirical findings support our arguments. For CEOs high in agreeableness, who are more attuned to the perceptions of others, we proposed that the threat of social disapproval from not taking swift action – which is central to the main relationship we hypothesize – would be more salient. This is borne out in our models, which show that CEOs’ agreeableness significantly amplifies CEOs’ neurotic dispositions at higher levels of agreeableness. Similarly, negative

external assessments by the media strengthen CEOs’ neurotic tendencies, while relatively positive evaluations appear not to provide the trait-relevant situational cues that activates neurotic tendencies.

We make two additional contributions through these findings. First, it is of great importance to note that, by theoretically arguing and empirically demonstrating an interaction between CEO neuroticism and CEO agreeableness, we highlight that multiple personality traits can jointly affect a manager’s actions. While this may not be surprising to scholars of personality psychology, who have long considered more integrated approaches (e.g., Hofstee et al. 1992), this is fundamentally new to the OSCM and the broader management literature. Prior studies on CEO personality have studied several FFM traits simultaneously, but only in separate direct-effect hypotheses (e.g., Herrmann and Nadkarni 2014). Our findings show that interactive effects between several personality traits, which have thus far been completely neglected in the management literature, are critical to obtaining a holistic perspective on the consequences of decision makers’ personality (Shoss and Witt 2014). Second, we respond to the recent call to “consider how relevant situational factors may differentially activate or constrain the various dimensions of each of CEOs’ personality traits” (Harrison et al. 2019, p. 1327). We do so by considering an interaction between CEO neuroticism, a personality trait, and CEO media tenor, which describes a situational circumstance surrounding a CEO’s actions. Our findings suggest that such interactive effects are indeed consequential and thus worthwhile to consider for future inquiry into the consequences of managers’ personalities.

Our research also has important implications for practitioners at all hierarchical levels, especially against the backdrop of the potentially substantial consequences of product recalls (Van Heerde et al. 2007). At the board level, directors may want to monitor that CEO personality does not unduly influence recall decisions.

For instance, taking measures to ensure that CEOs are properly accountable for their decision processes may help improve the efficacy of such decisions (Lerner and Tetlock 1999). At the same time, boards selecting a CEO may also wish to recognize that even seemingly undesirable personality traits like neuroticism may have potentially beneficial consequences.

At the level of top managers, not only should CEOs be aware of their own personal tendencies—as well as the additional influence of outside factors like media coverage—to be able to make the best possible decisions for their firms and stakeholders. Instead, other executives tasked with OSCM-related tasks, such as Chief Risk Officers (e.g., Li et al. 2022) or Chief Supply Chain Officers (e.g., Roh et al. 2016) share the same responsibility. In addition, they might wish to consider actively acting as a counterweight to any undesirable tendencies the CEO may exhibit. Finally, our research hints at the importance of individuals’ personality in OSCM contexts more generally. It hence holds implications for middle managers and even front-line employees. For one, they might similarly need to be aware of their own decision tendencies. For another, and arguably more importantly, they might wish to deliberately choose upward influence or communication strategies (Cable and Judge 2003) that account for their superiors’ personality dispositions to enable optimal operations decisions.

3.5.2 Limitations and Avenues for Future Research

The findings of our study should be viewed in the context of its limitations, which present opportunities for future research. One limitation is that we are considering only a single dependent variable, time-to-recall, which captures but one specific aspect of a larger recall episode. It is certainly a crucial variable, as it provides insights into whether CEOs and their firms take decisive action to quickly remove unsafe products from the market. Nevertheless, future studies may want

to investigate additional aspects of the management of recalls. Treating recalls as “product harm crises” (Van Heerde et al. 2007, p. 230) and building bridges to the broader crisis management literature (Schaedler et al. 2022), it might, for instance, be insightful to investigate the role taken by CEOs in communication efforts towards affected customers and the broader public when a recall is announced.

Our overall findings also open several further avenues of future research related to the recall, OSCM, and strategic leadership literatures. In the recall literature, researchers may want to investigate other, non-personality-related explanatory factors of our dependent variable, time-to-recall. Additionally, our empirical results suggest that CEO agreeableness may also be directly associated with shorter time-to-recall. While extant literature did not provide a clear ex ante indication for this association (see Appendix 3.B), researchers may want to examine this finding more closely to understand if and how this personality trait directly shapes recall timing. Also, it may prove valuable to understand the downstream consequences of shorter and longer time-to-recall. For example, researchers may want to investigate whether stock-market reactions to recalls (as studied by Ni et al. 2014 and others) or reactions by external stakeholders (such as consumers or the media) differ based on the speed with which firms recall defective products.

In the wider OSCM literature, there are near boundless opportunities to deepen our understanding of the impact of decision makers’ personality on important operational processes and outcomes. For example, future studies may want to explore how the personality of CEOs or supply chain managers affects their behavior in large-scale supply chain disruption events (Cohen et al. 2022).

In the strategic leadership literature more broadly, it may be worthwhile to study the determinants of the swiftness of executives’ reactions to other uncertainty-fraught corporate events with significant importance to firms and the potential for social threat to CEOs. While there are likely many such types of events,

investigating how a firm responds to data breaches, lawsuits, or consumer advocacy group campaigns may be particularly topical.

3.5.3 Conclusion

This study provides insights into how CEO personality shapes decision-making on product recalls. Our results show that CEOs who are more neurotic respond to product defects by initiating product recalls more quickly than less neurotic CEOs. We show that this effect is particularly strong for CEOs who are also more agreeable and who recently experienced negative media coverage. These findings not only answer theoretically important questions but also raise new ones, specifically in the literatures on recall timing and strategic leader personality. We hope that our findings will spur additional research in both fields and at their interface.

Appendix of Chapter 3

3.A Industry Overview

In our sample, we included all relevant sub-industries of the healthcare sector, as defined in the Global Industry Classification Standard (GICS).

Table 3.4: Summary of the sub-industries of the healthcare sector index (GICS 35) with focus on recall-able o□erings, which are in scope for our sample. As indicated, the description column contains the original description as defined in the index and is thus not an original work of the authors.

Sub-Industry Name (Code)	**Sub-Industry Description** (Source: S&P Global Market Intelligence 2018)
Healthcare Equipment (35101010)	Manufacturers of healthcare equipment and devices. Includes medical instruments, drug delivery systems, cardiovascular orthopedic devices, and diagnostic equipment
Healthcare Supplies (35101020)	Manufacturers of healthcare supplies and medical products not classified elsewhere. Includes eye care products, hospital supplies, and safety needle syringe devices.
Healthcare Distributors (35102010)	Distributors and wholesalers of healthcare products not classified elsewhere.
Biotechnology (35201010)	Companies primarily engaged in the research, development, manufacturing and/or marketing of products based on genetic analysis and genetic engineering. Includes companies specializing in protein-based therapeutics to treat human diseases. Excludes companies manufacturing products using biotechnology but without a healthcare application.
Pharmaceuticals (35202010)	Companies engaged in the research, development, or production of pharmaceuticals. Includes veterinary drugs.
Life Sciences Tools & Services (35203010)	Companies enabling the drug discovery, development, and production continuum by providing analytical tools, instruments, consumables supplies, clinical trial services and contract research services. Includes firms primarily servicing the biotechnology and pharmaceutical industries.

The sub-industries of the medical industry listed below are not included in our

sample since they provide services, operate facilities, manage healthcare service provision, or operate data processing software, all of which cannot be recalled.

Table 3.5: Summary of the sub-industries of the healthcare sector index (GICS 35) with focus on non-recall-able o□erings, such as the operation of healthcare facilities. As indicated, the description column contains the original description as defined in the index and is thus not an original work of the authors.

Sub-Industry Name (Code)	**Sub-Industry Description** (Source: S&P Global Market Intelligence 2018)
Healthcare Services (35102015)	Providers of patient healthcare services not classified elsewhere. Includes dialysis centers, lab testing services, and pharmacy management services. Also includes companies providing business support services to healthcare providers such as clerical support services, collection agency services, staffing services and outsourced sales & marketing services.
Healthcare Facilities (35102020)	Owners and operators of healthcare facilities, including hospitals, nursing homes, rehabilitation and retirement centers and animal hospitals.
Managed Health Care (35102030)	Owners and operators of Health Maintenance Organizations (HMOs) and other managed plans.
Health Care Technology (35103010)	Companies providing information technology services primarily to healthcare providers. Includes companies providing application systems and/or data processing software, internet-based tools, and I.T. consulting services to doctors, hospitals or businesses operating primarily in the healthcare sector.

3.B Five-Factor Model (FFM) Traits

Table 3.6: Overview of the personality traits described in the Five-Factor Model (FFM) and brief commentary on potential links to time-to-recall

Trait & Brief Description	**Comment on potential links to time-to-recall**
Neuroticism is the tendency to exhibit emotional instability (Tackett and Lahey 2016). Occasionally, the opposite of neuroticism, "emotional stability", is also used described this personality trait (Hills and Argyle 2001). While individuals with high levels of neuroticism are said to be more insecure, irrational, and act more impulsively, those with low levels of neuroticism are characterized as unemotional or calm (Costa et al. 1991, DeYoung et al. 2007). Neuroticism is "strongly associated" with social anxiety and a "desire to avoid social disapproval" (Bendersky and Shah 2013, p. 389, Elliot and Thrash 2002).	Psychological theory on neuroticism clearly indicates that neurotic individuals respond strongly and impulsively to uncertainty in an e□ort to minimize potential downsides that may emerge ("the devil you don't know"), even if this means accepting some concrete negative outcomes ("the devil you know"), i.e., by following a "better safe than sorry" approach (Hirsh and Inzlicht 2008, p. 966, Lommen et al. 2010) Consequently, based on this literature, we deem neuroticism the only trait that provides a clear directional indication on how individuals who score highly on this dimension resolve the uncertainty that prevails when a CEO must decide when to recall a defective product. We develop Hypothesis 3.1 on the relationship between neuroticism and time-to-recall in greater detail in the main body our paper.

Continued on next page

Table 3.6: Overview of FFM traits and potential links to time-to-recall (Continued)

Agreeableness describes interpersonal behavior with a focus on the quality of interactions with others "along a continuum from compassion to antagonism" (Costa and McCrae 1985, p. 2). Agreeableness has been associated with a range of prosocial motives and behaviors (e.g., Gleason et al. 2004, Graziano and Eisenberg 1997).	Agreeable individuals are more attuned to perceptions of others and this "propensity [...] to attend to the mental states of others may be central to [...] agreeableness" (Nettle 2006, p. 323) Accordingly, we expect more agreeable CEOs to be more keenly aware of the needs of the multiple stakeholders of a product recall, such as customers, regulators, distributors, and regulators. As these stakeholders, however, may have very di□erent preferences regarding the timing of a recall, we have no indication on how being attuned to these di□erent perceptions would impact CEOs' decision making on recall timing in a clear and directionally predictable way. Accordingly, we cannot hypothesize a direct e□ect between agreeableness and time-to-recall. At the same time, psychological theory suggests that an agreeable individual's greater propensity to be attuned to the perceptions and needs of others is likely to generally make stakeholders' needs more salient. Such an elevated level of salience may now exert an e□ect in connection with neurotics' "desire to avoid social disapproval" (Bendersky and Shah 2013, p. 389). If CEOs are high in agreeableness, and thus highly aware of the perception of others, this is likely to amplify their perception of—and consequently their motivation to act to minimize—a neurotic threat of social disapproval. Conversely, more disagreeable CEOs, who are less inclined to consider the perspective of others in their handling of a product recall, will be less likely to perceive a significant social threat from not taking swift action. We develop Hypothesis 3.2 on an interactive e□ect between CEO agreeableness and CEO neuroticism on time-to-recall in greater detail in the main body our paper.

Continued on next page

Table 3.6: Overview of FFM traits and potential links to time-to-recall (Continued)

Openness, sometimes also referred to as openness to experience "is seen in the breadth, depth, and permeability of consciousness, and in the recurrent need to enlarge and examine experience" (McCrae and Costa 1997, p. 826). Individuals high in Openness are said to be more liberal, intellectually engaged and more appreciative of art and beauty (Costa and McCrae 1985, Sutin 2015).	Openness to experience describes the level of intellectual curiosity of an individual and the range of their personal interests. Accordingly, it is reasonable to expect that a CEO's level of openness would have a considerable impact on some firm processes and outcomes that are closely associated with curiosity, imagination, creativity, and open-mindedness. In fact, Nadkarni and Herrmann (2010) demonstrate that CEO openness is positively associated with their firm's strategic flexibility. As we do not have any theoretical indications that a CEO's level intellectual curiosity or breadth of interests would impact decision making on recall timing, we have no grounds to hypothesize a relationship between openness and time-to-recall.

Continued on next page

Table 3.6: Overview of FFM traits and potential links to time-to-recall (Continued)

Conscientiousness describes the propensity to be "self-controlled, responsible to others, hard-working, orderly, and rule abiding" (Jackson and Roberts 2015, p. 1). Individuals high in conscientiousness are said to be achievement oriented, cautious, deliberate, and dutiful, while individuals with lower levels of this trait are said to be more unreliable or even careless (Costa et al. 1991, Costa and McCrae 2008, DeYoung et al. 2007). This trait has been associated with a positive impact across a wide array of important life outcomes (Jackson and Roberts 2015).	Psychological theory on conscientiousness suggests two potentially countervailing effects. On the one hand, the dutiful and rule-abiding nature may lead conscientious CEOs to work towards triggering recalls more quickly, such that the risk of potential further harm to patients could be minimized. This would suggest a shorter time-to-recall. On the other hand, conscientious CEOs' deliberate and process-oriented nature should lead them to coordinate a diligent process to clarify the exact scope of a defect and develop a plan to remedy it, rather than act potentially prematurely. If they can do so before the defects become publicly known or (further) harm occurs to patients, this may minimize the financial impact on the firm, thus also allowing CEOs to fulfill their duties to other stakeholders, such as the company's shareholders. Given the lack of clear prior theoretical guidance to hypothesize either a positive or a negative (or a null) effect, we do not hypothesize a relationship between conscientiousness and time-to-recall. Note that the theoretical reasoning outlined above is different from that on neuroticism, on which we base our main hypothesis. Theory on neuroticism clearly indicates that in situations of uncertainty, neurotic individuals act strongly and impulsively in an effort to minimize potential downsides that may emerge, even if this means accepting some concrete negative outcomes (Hirsh and Inzlicht 2008, Lommen et al. 2010). Put differently, they take swift action, even if this means accepting sub-optimal outcomes. This logic is not applicable to conscientiousness.

Continued on next page

Table 3.6: Overview of FFM traits and potential links to time-to-recall (Continued)

Extraversion describes individuals who are "characterized by energy, dominance, spontaneity, and sociability," while those with lower levels of extraversion (i.e., introverts) are "lethargic, inhibited, reflective, and quiet" (Wilt and Revelle 2016, p. 1). Extraversion thus describes an individual's preferred amount of interpersonal stimulation (Costa et al. 1991).	Among the FFM traits in CEOs, extraversion has been studied extensively. For example, studies have theorized and empirically validated associations between CEO extraversion and mergers and acquisitions (Malhotra et al. 2018), transformational leadership (Bono and Judge 2004), leadership emergence (Judge et al. 2002), and initiation of strategic change (Herrmann and Nadkarni 2014). Unlike in the cases described above, we do not see how the assertiveness and desire for interpersonal interaction of extraverted CEOs would impact their decision making on recall timing in a clear and directionally predictable way. We consequently do not hypothesize a relationship between extraversion and time-to-recall.

3.C Details on Methodology

We summarize the methodology used to determine the strength of each CEO’s personality traits in Table 3.7 below. In Figure 3.3 below, we show extracts from a sample transcript of an earnings conference call.

Table 3.7: Overview of the methodology used to assess CEO’s Five-Factor Model (FFM) personality traits, including CEO neuroticism and CEO agreeableness

Step	**Description**
1. Gather transcripts of earning conference calls	We first retrieved all available transcripts of quarterly earnings conference calls for sample firms during the sample time frame from Fair Disclosure wires on LexisNexis and Thomson StreetEvents. Each transcript contains information about the conference call, a list of internal and external participants, and a transcript of the call itself. We share extracts of a sample conference call transcript in Figure 3.3. Earning conference calls consist of two segments: • An opening segment, also known as the “presentation”, with the speech of one or more company executives. All individuals other than the operator and the company’s executives are muted and cannot speak during this segment. • A question-and-answer (Q&A) segment in which executives respond to questions from analysts. These exchanges may include several rounds of follow-up between analyst and several company executives.

Continued on next page

Table 3.7: Methodology used to assess CEOs’ FFM traits (Continued)

2. Process transcripts to filter attribute text to speakers	We attribute all statements made by participants in each call to the respective individual. As the text is structured in paragraphs, any paragraphs without a speaker can be attributed to the last speaker specifically mentioned. In line with prior research (e.g., Malhotra et al. 2018, Harrison et al. 2019, 2020), we retained only those statements made by CEOs in the Q&A segment. In the example in Figure 3.3, these are highlighted in the red boxes with dashed outline. We focused on these impromptu responses to previously unknown questions because the “spontaneous nature of questions and answers” a□ords us greater insights into CEO personality than the prepared remarks of CEOs at the beginning of conference calls would (Matsumoto et al. 2011, p. 1387).
3. Collate statements made by CEOs	We proceeded to collate all responses to questions made by each CEO across our entire sampling frame. As a result, our measures of CEO personality are time invariant. This is in line with prior work in the management literature, including that by Harrison et al. (2019, 2020), as well as the broader personality literature, which shows that these personality traits are stable in adult life (Roberts and DelVecchio 2000).

Continued on next page

Table 3.7: Methodology used to assess CEOs’ FFM traits (Continued)

4. Use Harrison et al.’s OLCPT to score CEO personality traits	We assessed CEO personality using the extensively validated open-language tool developed by Harrison et al. (2019, 2020), which is experiencing increasing popularity in the management literature. This tool uses a Gradient Boosting Machine learning model. Specifically, the model was trained using a data set of more than 200 CEOs, for which the developers gathered: • Responses to questions in earning conference calls (gathered in line with steps 1–3 above) • Personality scores gathered by Hill et al. (2019) using a psychometrically validated instrument • Personality scores as rated by three doctoral students using available video footage and the 50-item international personality item pool (Goldberg et al. 2006) For more details, including the extensive validation e□orts, please refer to the original publications by Harrison et al. (2019, 2020). We were given access to the tool, which is programmed in R, by the developers. We followed their specific instructions for installing the tool, loading data, and generating scores for our data-set using the pre-trained machine learning model.

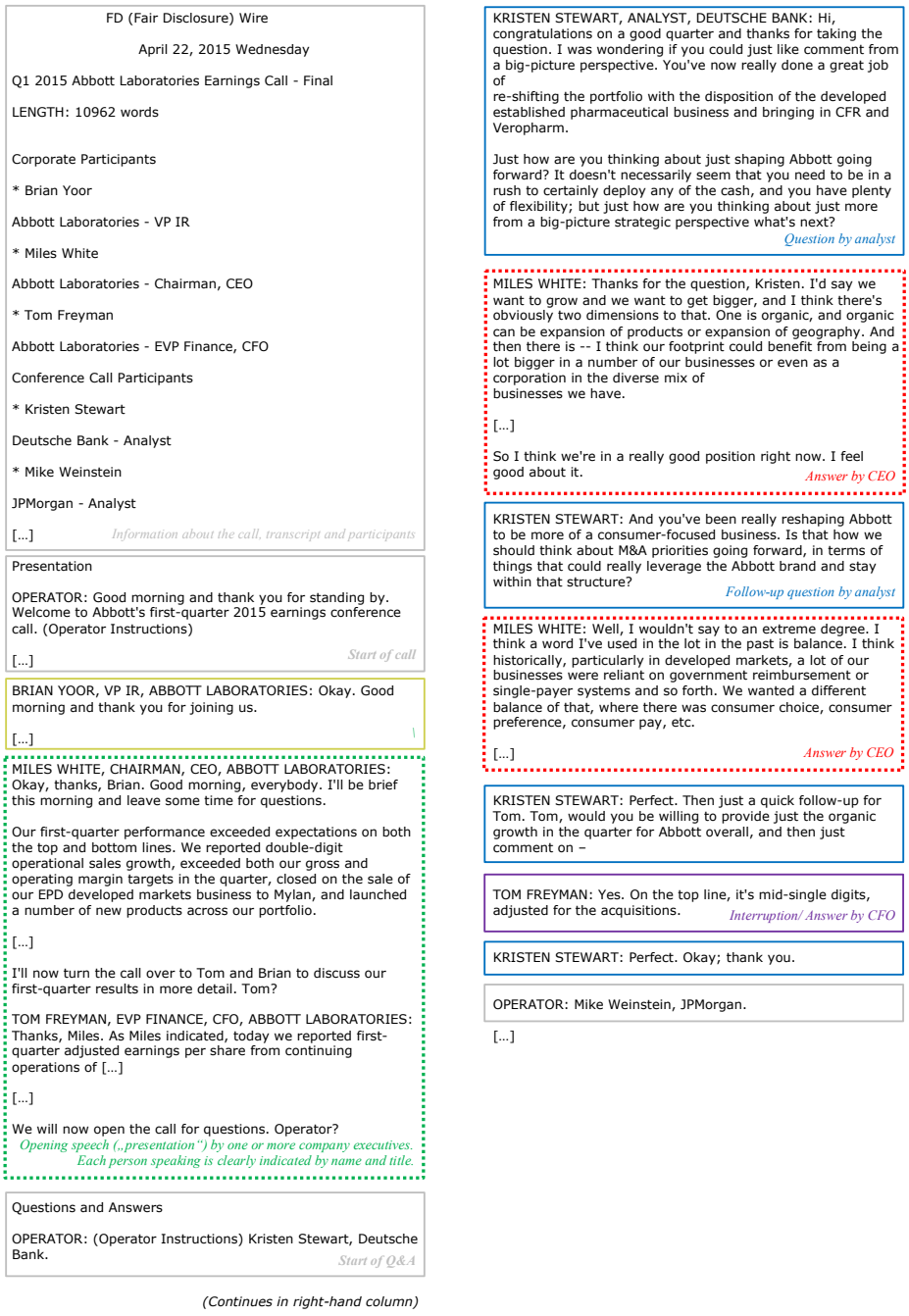

Figure 3.3: Sample transcript of an earnings conference call with highlighted sections.

3.D Additional Interaction Plots

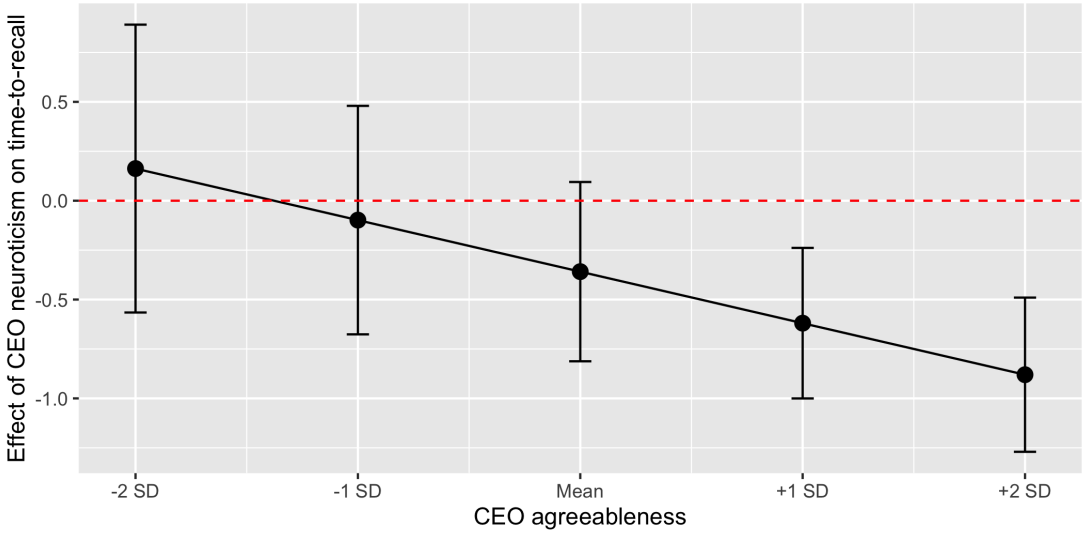

Figure 3.4: Average marginal e□ect of CEO neuroticism on time-to-recall as CEO agreeableness varies in model 3, including 95% confidence interval.

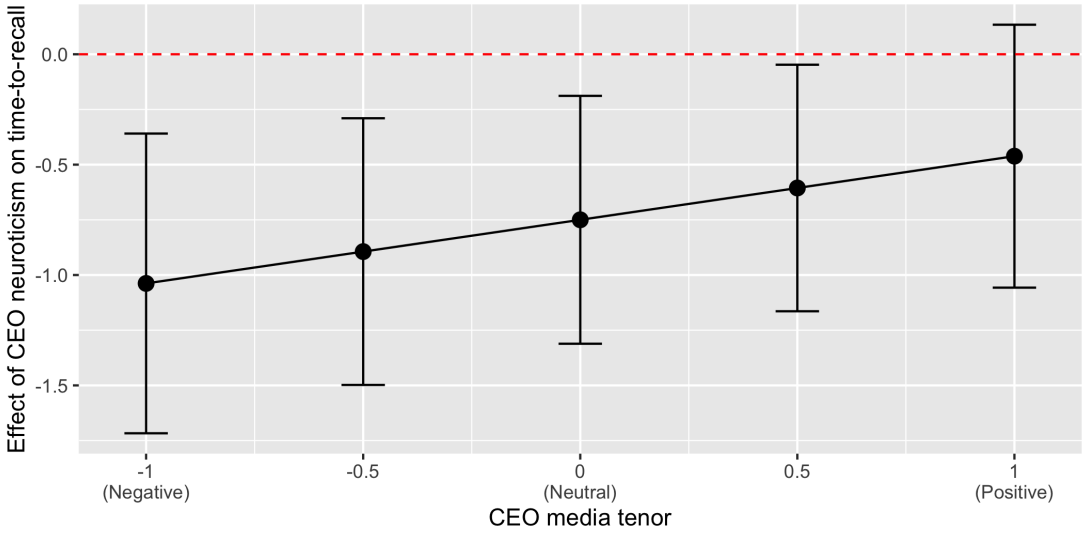

Figure 3.5: Average marginal e□ect of CEO neuroticism on time-to-recall as CEO media tenor varies in model 4, including 95% confidence interval.

Chapter 4

"We're humans, not machines": Exploring Personality-based Heterogeneity in Metaknowledge and Human-AI Collaboration

While artificial intelligence (AI) has great potential to transform how we work, it's clear that in many contexts, humans and AI will work side by side. According to recent research, humans' ability to capture complementarities while working collaboratively with AI is critically constrained by our limited ability to assess what we know, i.e., metaknowledge. However, extant research has largely overlooked a critical difference between human and AI: Human behavior is shaped by personality. We explore how differences arising from personality traits impact metaknowledge and delegation behavior towards AI. Using an image classification experiment, we find that higher levels of conscientiousness improve, while higher levels extraversion harm humans' ability to work collaboratively with AI. We find that differences in metaknowledge help to explain this relationship. Our findings motivate further research on personality in human-AI collaboration and the design of collaborative systems which foster metaknowledge and improved performance.

4.1 Introduction

Although AI systems have made considerable advances across a variety of domains, many settings remain in which humans and AI possess complementary knowledge (e.g., Patel et al. 2019, Steyvers et al. 2022). These complementarities present opportunities to design and use collaborative settings with both humans and AI that can achieve superior performance than either party could realize alone. In order to leverage these complementarities, a given task should be performed by (i.e., delegated to) the actor most capable in solving that task (Autor et al. 2003). Not only does this require an actor to be able to assess the ability of the other actor, but critically, this requires actors to estimate their own ability to perform a particular task.

This ability to assess one’s own capabilities to complete a particular task, called metaknowledge, is crucially important for successful collaboration and delegation between humans and AI systems (Evans and Foster 2011, Fügener et al. 2022). At the same time, emerging Information Systems (IS) research suggests a lack of metaknowledge is a key obstacle which prevents humans from working with AI and delegating tasks to an AI in a way that leverages their complementary skills (Zhang et al. 2020, Fügener et al. 2022).

However, extant research on human-AI collaboration broadly, and delegation between humans and AI systems specifically, has generally overlooked a critical difference between human and most AI systems: Human behavior and cognition is shaped by humans’ personality traits (Schneider and Smith 2004). This means that humans may collaborate with an AI system in different ways, depending on their personality traits. We aim to address this gap by exploring the effects of personality traits on metaknowledge and human delegation behavior and answering the question “*How do personality traits influence human metaknowledge*

and collaborative performance with an AI?"

To answer this question, we use data from an image classification experiment with 281 participants. In the experiment, participants classified a set of 100 images, indicated the confidence in their response, and were then given the opportunity to delegate the final decision to an AI. We varied whether participants saw the AI's certainty prior to taking this delegation decision. Additionally, we used a validated, widely used questionnaire to assess the five personality traits of the Five-Factor Model (FFM), the predominant model of personality structure (John and Srivastava 1991, McCrae and Costa 1987). Our preliminary findings show that two of these personality traits, conscientiousness, and extraversion, improve and harm humans' ability to work collaboratively with AI, respectively. Using a causal mediation analysis, we show that differences in metaknowledge help to explain this relationship. Using image-level data, we can demonstrate a particular effect of extraversion which means that individuals with higher levels of extraversion are more confident in incorrect decisions.

4.2 Theory

4.2.1 Human-AI Collaboration and Metaknowledge

In settings where humans are unable to adequately depict their decision rules, humans and AI have complementary knowledge (Autor et al. 2003). A key way of leveraging this complementary knowledge is through delegation from a human to an AI system, or vice versa (Baird and Maruping 2021). This transfers the responsibility of completing of a task or decision from the principal to an agent (Eisenhardt 1989). Where humans and AI have complementary knowledge, effective delegation enables the joint performance of humans and AI system

to exceed the individual performance of the better-preforming actor (Steyvers et al. 2022, Zhang et al. 2020). Fügener et al. (2022) describe two conditions for effective delegation: Firstly, this requires the delegating actor, i.e., the principal, to recognize the complementarities with the agent. This requires the principal to possess adequate metaknowledge, i.e., an ability to assess how capable they are to execute a task by themselves (Evans and Foster 2011), and a perception of the agents’ own ability to complete the same task (Zhang et al. 2020). Secondly, the principal needs to delegate tasks with known complementarities to the more capable agent. While it is straight-forward to implement such a logic in situations where an AI delegates tasks to humans, humans generally appear to struggle to consistently follow decision rules to delegate to a more capable algorithmic counterpart (Fügener et al. 2022).

Although research has demonstrated the presence of complementary knowledge, it also indicates that humans to have difficulty to capture this complementary knowledge in a way that improves joint performance. In an experiment by Lai and Tan (2019) in which participants were asked to collaborate with an algorithm to flag deceptive online hotel reviews, participants were unable to surpass the algorithm’s own performance. Thus, they were not able to capitalize on the complentary knowledge with the AI. Studies by Wang and Yin (2021), Zhang et al. (2020), and others have found similar results.

One idea that has frequently been proposed is to provide humans with more information about the algorithm, including about the algorithms’ certainty in its prediction. In theory, having information about the algorithms certainty should encourage humans to rely on the algorithms input when it is certain, and encourage them to complete tasks on their own when the algorithm is rather unsure (Tomsett et al. 2020, Wang and Yin 2021). Indeed there is some evidence that such information improves human’s mental model about the algorithms’ error

boundaries (e.g., Bansal et al. 2019) and increase unique human knowledge (e.g., Fügener et al. 2021).

4.2.2 Personality Traits

An individuals’ personality is made up of “enduring characteristics and behavior” (American Psychological Association 2022), which can help to understand how individuals feel and behave (Schneider and Smith 2004). A plethora of models has emerged, aiming to structure personality into personality traits, which are “distinguishable and stable patterns of behavior” (Shiraev 2016, p. 202). Among these, the Five-Factor Model (FFM) pioneered by McCrae and Costa (1987) has emerged as the dominant taxonomy (Bainbridge et al. 2022).

As its name suggests, the FFM describes an individual’s personality along five dimensions, which are also known as the Big 5 personality traits. These traits are conscientiousness, extraversion, agreeableness, openness to experience, and neuroticism (McCrae and Costa 1987). We describe these traits briefly; however, due to the exploratory nature of our work, we do not report any hypotheses.

Conscientiousness describes an individual’s propensity to be self-controlled, orderly, and rule-abiding (DeYoung et al. 2007). While this trait has traditionally been associated with positive outcomes (e.g., higher performance) in the workplace (Jackson and Roberts 2015), recent research suggests that individual with higher levels of conscientiousness may benefit less from working with an AI system than others (e.g., Tang et al. 2021).

Extraversion, as opposed to introversion, describes a desire for social stimulation and is characterized by a tendency to exhibit assertive behavior (Judge et al. 2002). Individuals who score highly in extraversion are more likely to take charge in interactions with others and confidently express their opinions (Wilt and Revelle

2016). However, research indicates that their tendency to be assertive may also manifest into overconfidence (Schaefer et al. 2004) and may make it less likely for them to transfer control to an AI (Goldbach et al. 2019).

Finally, Agreeableness describes individual differences in approaches to interpersonal relations (Graziano and Eisenberg 1997). Individuals who score highly on this dimension are commonly described as warm or sensitive and more concerned with cooperation (Thompson 2008). Openness to experience refers to the “experiential component” of personality and is associated with broad personal interests creativity (e.g., Sutin et al. 2011). Neuroticism is the tendency exhibit emotional instability. Individuals with higher levels of neuroticism have a greater tendency to experience anxiety, stress, and self- consciousness and are also known to act more impulsively than others (Judge et al. 2013).

To the best of our knowledge, only one prior study by Goldbach et al. (2019) has considered the effect of these personality traits on delegation in human-AI collaboration. However, this study was a specific application to a multi-actor traffic coordination problem with no ground truth. Accordingly, many of the key considerations applied in this research project, including, but not limited to metaknowledge, were not considered (and could not readily be considered) in that work.

4.3 Data and Methodology

4.3.1 Experimental Design

We use data from an image-classification experiment, building on the design used by Fügener et al. (2021, 2022). We use image classification for two reasons: First, it is a task that is well suited for both humans and AI systems. For humans, it is

a generic task that does not require specific knowledge or training. Modern AI can be trained to achieve a high level of performance in these tasks (Russakovsky et al. 2015). Second, humans and modern AI systems possess complementary knowledge in this task, allowing them to collaboratively outperform humans or the AI on their own (Fügener et al. 2022).

For a set of 100 images, participants are asked to assign an image to one class (from a list of 10 classes) which best represents that image. On the first screen, which is shown in Figure 4.3 in Appendix 4.A, participants can see the class name and thirteen other images belonging to that class for each of the 10 classes. Classes are shown in random order. An image is classified correctly if the correct class was chosen for a given image. Additionally, participants are asked to indicate how certain they are that their choice is correct using a scale from 0 (not certain) to 100 (certain).

After classifying an image and indicating their certainty, participants proceed to a second screen where they are given the option to delegate the final classification decision for that image to an AI. This second screen can also be seen in Appendix 4.A. We manipulated the information shown on this screen. In the Control group, which is our control group, participants were provided no additional information (see Figure 4.4 in Appendix 4.A). By contrast, subjects in the Treatment group were shown the AI’s certainty and their own previously stated certainty before they decided whether to delegate the image or not (Figure 4.5 in Appendix 4.A). We use this Treatment to evaluate whether providing a reference value for certainty enables participants to delegate more effectively. As part of the general instructions for both Treatments, participants were informed that if they decide to delegate the decision to the AI, the AI’s answer for that image will be used instead of their own answer.

Like Fügener et al. (2021, 2022), the experiment uses a set of 100 images with

known class labels from the ImageNet database, a large and widely used set of human-annotated images (Russakovsky et al. 2015). This set of 100 images was not included in Fügener et al.’s samples, which will also allow our study to contribute towards replicating and validating some of their original findings. All participants classified the same 100 images in a random order to make results comparable while simultaneously eliminating the possibility that an order bias could affect the results. The AI used was Inception-v3, a convolutional neural network (CNN) trained on the ImageNet database which assigns a certainty score to 1,000 possible classes per image (Szegedy et al. 2016). Specifically, the AI’s classification choice was the class with the highest certainty score as the AI’s answer for a given image. The AI was able to correctly classify 78 out of these 100 images. After classifying the images, participants responded to a set of survey questions relating to the task and themselves, including on demographics and personality. Here, participants completed the Big Five Inventory (BFI; John et al. 1991. This widely accepted and extensively used method for assessing the FFM traits contains 44 statements which participants rate on a five-point Likert scale from 1 (disagree a lot) to 5 (agree a lot) (Allik and Realo 2015).

4.3.2 Sample

Data collection for this research project was part of a larger data collection effort, on which work by Taudien et al. (2022) is based. We report all relevant methodological choices impacting our research.

The experiment ran on two consecutive days in April 2022 using a pool of Amazon Mechanical Turk (MTurk) workers, which were recruited on CloudResearch. Participants received a fixed reward of $2.50 for their participation and a performance-based reward of $0.05 for each image they classified correctly (or chose to delegate to the AI, which classified it correctly). Additionally, participants received a

further $0.50 for completing the personality-related questionnaires at the end of the survey.

Of the 780 participants originally recruited, 180 participants were prevented from proceeding into the experiment because they attempted to take the survey a second time (n=102), had an incompatible screen resolution (n=48), or failed the attendance check (n=30). 568 respondents completed the experiment. We removed eight responses ex-post as our analysis demonstrated that these were from four participants who completed the survey two times in parallel. We used respondents’ MTurk worker IDs and IP address to identify such misusage. In total, 560 valid responses were obtained in this way.

While the full data collection effort contained a 2x2 between-subject design, we focus our analysis on two conditions. We do because of the host of possible effects to analyze with 5 variables of interest (i.e., personality traits of the FFM model) and the large number of possible interaction effects with the Treatment. Concretely, we focus our analysis on a single manipulation, using a control (Treatment 1) and a single treatment condition (Treatment 2; displaying the AI’s certainty before the delegation decision is taken). However, we validated that our main findings relating to the direct effects of conscientiousness and extraversion on performance and metaknowledge (models A2, B2, C2, D2, and E2) can also be reproduced with the larger sample.

Our sample consists of 281 respondents. Participants were 18 to 74 years old with the largest share of participants reporting their age to be between 35 and 44 (30%). Most of the respondents were female (58%). 41% of respondents indicated that they had a bachelor’s degree, 16% held at least a master’s or professional degree. In total, all but 29 participants (90%) had at least a college degree. Participants received an average payout of $6.95 for their participation. Of the 281 respondents, 137 were randomly assigned into the Control group and 144 were

randomly assigned into the Treatment group.

4.4 Results

4.4.1 Individual and Collaborative Performance

We report the results of a hierarchical linear regression on the number of images classified correctly by humans, and the number of images classified correctly in collaboration between the humans and the AI, in Table 4.1.

On average, respondents classified some 72 (control) to 74 (treatment) images correctly. No significant difference between treatments can be observed (Model A1). Jointly (i.e., after any potential delegation), humans and the AI were able to classify 78 to 80 images correctly, marginally exceeding the AI’s stand-alone performance of 78 images. Put differently, humans overall were not able to work with the AI in a way that adequately captures complementary knowledge. Again, performance was not significantly influenced by the treatment (Model B1).

A simple ex-post analysis reveals that considerable complementarities could have been captured if humans had followed a simple decision rule: If their own certainty exceeds the certainty of the AI, they complete the task on their own (i.e., do not delegate to the AI). In all other cases, the user would delegate the task to the AI. Applying this decision rule for the Treatment group (where participants could have applied this exact decision rule) in an ex-post analysis indicates that an average of 82 images could have been classified correctly, highlighting that substantial complementarities were not captured by human respondents. This finding, which is substantially similar for the Control group, is in line with other papers which also found that humans are not able to fully leverage complementarities (Fügener et al. 2022, Zhang et al. 2020).

Table 4.1: Results of Ordinary Least Squares (OLS) linear regression on image classification performance

	Performance, i.e., number of images correctly classified by ...					
	Humans alone			Humans + AI		
	(A1)	(A2)	(A3)	(B1)	(B2)	(B3)
Constant	72.101	52.649	67.705	78.104	60.686	65.971
	(1.963)	(8.876)	(11.914)	(1.611)	(7.178)	(9.666)
Treatment	-1.084	-1.292	-33.029	1.034	0.792	-13.563
	(1.396)	(1.357)	(16.597)	(1.145)	(1.098)	(13.465)
Openness		2.144	0.884		1.010	1.199
		(0.975)	(1.551)		(0.789)	(1.258)
Conscientiousness		3.220	2.404		3.413	3.128
		(1.161)	(1.506)		(0.939)	(1.222)
Extraversion		-2.316	-2.463		-2.127	-2.172
		(0.802)	(1.131)		(0.648)	(0.917)
Agreeableness		0.587	-0.504		0.876	0.155
		(1.158)	(1.642)		(0.936)	(1.332)
Neuroticism		0.943	-0.001		0.882	0.087
		(0.947)	(1.229)		(0.766)	(0.997)
Openness × Treatment			1.762			-0.582
			(1.988)			(1.613)
Conscientiousness × Treatment			2.532			1.218
			(2.285)			(1.854)
Extraversion × Treatment			0.366			0.053
			(1.612)			(1.308)
Agreeableness × Treatment			2.103			1.538
			(2.311)			(1.875)
Neuroticism × Treatment			2.245			2.092
			(1.848)			(1.499)
Observations	281	281	281	281	281	281
Adjusted R^2	0.026	0.085	0.082	0.019	0.104	0.096

Standard errors in parentheses. One observation is one subject.
Control variables for demographics, i.e., age, gender, and education level are included, but not shown in table
$^{+}$ $p < 0.10$, $^{\square}$ $p < 0.05$, $^{\square\square}$ $p < 0.01$, $^{\square\square\square}$ $p < 0.001$

We proceed to analyze the impact of respondent personality on performance. In models A2 and B2, we included all five of the FFM traits of the respondent as additional regressors. We also ran models which included interaction effects between the FFM traits and the treatment to capture potential effects of the personality traits that were specifically elicited by the Treatment. We report the result for these models in column A3 and B3, respectively.

Extraversion is negatively associated with performance of humans (β = -2.316, $p < 0.01$, Model A2) and humans collaborating with AI (β = -2.127, $p < 0.01$, Model B2). This means that individuals with higher levels of extraversion, who are more assertive but also prone to overconfidence (DeYoung et al. 2007, Schaefer et al. 2004), are less able to classify images correctly alone and when working with an AI system than are individuals with lower levels of extraversion.

In line with theoretical predictions which stress the hard-working and rule-abiding nature of conscientious individuals (Jackson and Roberts 2015), conscientiousness is positively associated with the number of images identified correctly working alone ($\beta = 3.220$, $p < 0.01$, Model A2) and with the AI ($\beta = 3.413$, $p < 0.001$, Model B2). The effect of human performance is directionally consistent, but not significant ($\beta = 2.404$, $p = 0.112$) in the control treatment.

4.4.2 Delegation

We also analyzed delegation behavior using the number of delegations, i.e., the number of times humans opted to delegate the final decision to the AI. Across both groups, respondents delegated 22.3 decisions on average. However, there was considerable heterogeneity in delegation behavior ($\sigma = 16.9$). We show the results of a hierarchical linear regression on the number of delegations in Table 4.2. Using a dummy variable for the treatment and control variables for demographics, the analysis indicates that the treatment is associated with a significantly higher number of delegations (model C1, $\beta = 5.110$, $p < 0.05$). In the main model C3 reported in Table 4.2, we do not observe significant effects of the FFM traits on the overall number of delegations to the AI. Interestingly, this also means we are unable to replicate the effect of extraversion on the number of delegations shown by Goldbach et al. (2019), albeit in a very different context and experimental setup.

To better understand which images participants are delegating, we compute an additional measure, which we call rule-compliant delegation decisions. This measure describes the number of times respondents chose to delegate to the AI or not to delegate in compliance with the simple decision rule outlined previously. This decision rule would prescribe humans to delegate any tasks to the AI for which their certainty is lower than the AI’s and complete all tasks for which their certainty is higher than the AI’s by themselves. This would allow humans to capitalize on complementarities with the AI. The number of such delegations was generally higher for the Treatment group, also when controlling for demographics and personality traits ($\beta = 11.865$, $p < 0.001$, model D2).

The regressions including personality traits provide a valuable perspective into how humans’ personality traits are affecting delegation. Higher levels of conscientiousness are positively associated with more rule-compliant delegation decisions ($\beta = 8.129$, $p < 0.001$, model D2). Our detailed analyses suggests that this effect is primarily driven by decisions of these individuals not to delegate in situations in which they are more certain in their answer than the AI is in its answer. Higher levels of extraversion are negatively associated with rule-compliant delegation ($\beta = -2.930$, $p < 0.01$, model D2). Table 4.3 presents an analysis of the marginal effects considering the interaction effects between personality traits and the treatment from model D3. This analysis reveals that the effect of extraversion is only significant for the Treatment (dydx = -3.451, $p < 0.05$) and not for the Control group (dydx = -2.282, $p = 0.153$). Finally, neuroticism is associated with more rule-compliant delegation ($\beta = 4.142$, $p < 0.01$, model D2). This effect is consistent and at least marginally significant in both the Control and Treatment groups.

Table 4.2: Results of Ordinary Least Squares (OLS) linear regression on number of final decisions delegated

	Number of final decisions delegated					
	Total			Rule-compliant		
	(C1)	(C2)	(C3)	(D1)	(D2)	(D3)
Constant	26.466 (2.882)	30.713 (13.518)	10.240 (18.023)	65.524 (2.852)	15.510 (12.495)	29.615^{+} (16.772)
Treatment	5.110 (2.049)	4.986 (2.067)	40.580 (25.106)	12.090 (2.028)	11.865 (1.911)	-18.207 (23.364)
Openness		-1.488 (1.485)	0.646 (2.346)		1.754 (1.373)	0.484 (2.184)
Conscientiousness		-0.262 (1.768)	-0.687 (2.279)		8.129 (1.635)	8.567 (2.121)
Extraversion		0.182 (1.221)	1.983 (1.710)		-2.930 (1.129)	-2.282 (1.592)
Agreeableness		0.996 (1.763)	3.084 (2.483)		1.777 (1.630)	-1.148 (2.311)
Neuroticism		-0.647 (1.442)	-0.389 (1.858)		4.142 (1.333)	3.352^{+} (1.730)
Openness × Treatment			-3.472 (3.008)			1.598 (2.799)
Conscientiousness × Treatment			0.792 (3.456)			0.043 (3.217)
Extraversion × Treatment			-3.953 (2.438)			-1.170 (2.269)
Agreeableness × Treatment			-3.872 (3.495)			5.735^{+} (3.253)
Neuroticism × Treatment			0.119 (2.795)			2.081 (2.601)
Observations	281	281	281	281	281	281
Adjusted R^2	0.035	0.024	0.033	0.155	0.254	0.252

Standard errors in parentheses. One observation is one subject.
Control variables for demographics, i.e., age, gender, and education level are included, but not shown in table
$^{+}$ $p < 0.10$, $^{\square}$ $p < 0.05$, $^{\square\square}$ $p < 0.01$, $^{\square\square\square}$ $p < 0.001$

4.4.3 Metaknowledge

To better understand the role that metaknowledge plays in shaping performance, we proceed to analyze how the personality traits influence metaknowledge at the participant level. For this purpose, we define metaknowledge as the difference in respondent-indicated certainty on images classified correctly and the certainty on

Table 4.3: Marginal e□ects of the respective personality trait on the number of rule-compliant delegations, model D3

	Control group		Treatment group	
	dy/dx	p-value	dy/dx	p-value
Openness	0.484	0.825	2.082	0.243
Conscientiousness	8.567	0.000	8.610	0.001
Extraversion	-2.282	0.153	-3.451	0.033
Agreeableness	-1.148	0.620	4.587	0.047
Neuroticism	3.352	0.054	5.433	0.008

images classified incorrectly. A larger positive value indicates that respondents are more able to distinguish between what they know and what they do not know, i.e., have higher levels of meta-knowledge. Reversely, a negative value indicates that respondents claimed higher certainty on images they classified incorrectly than on those they did classify correctly, i.e., they do not possess adequate metaknowledge.

We proceed to run a hierarchical linear regression, using our definition of meta-knowledge as the dependent variable. Again, we included a dummy variable for the treatment, control variables for demographics, the FFM traits. The results can be found in Table 4.4. The marginal effects for model E3 are shown in Table 4.5. While there is no significant treatment effect (β = -1.843, $p > 0.1$, model E1), a number of personality-based effects were significant. The effect of conscientiousness was positive and significant (β = 3.928, $p < 0.001$) while the effect of extraversion was negative and significant (β = -2.820, $p < 0.001$). The effects are also significant when considering an interaction with the treatment. These results provide an indication that conscientiousness and extraversion shape meta-knowledge in opposite ways: While higher levels of conscientiousness are associated with higher levels of meta-knowledge, making it easier for those individuals to capture complementarities with an AI, the opposite is true for individuals with higher levels of extraversion.

Additionally, neuroticism was positively and significantly associated with meta-

Table 4.4: Results of Ordinary Least Squares (OLS) linear regression on image classification metaknowledge

	Metaknowledge		
	(E1)	(E2)	(E3)
Constant	20.624	-2.536	-12.622
	(1.929)	(8.538)	(11.344)
Treatment	-1.843	-1.913	9.832
	(1.371)	(1.306)	(15.802)
Openness		1.475	4.272
		(0.938)	(1.477)
Conscientiousness		3.928	4.030
		(1.117)	(1.434)
Extraversion		-2.820	-2.306
		(0.771)	(1.077)
Agreeableness		0.606	-0.170
		(1.114)	(1.563)
Neuroticism		2.652	2.790
		(0.911)	(1.170)
Openness × Treatment			-4.857
			(1.893)
Conscientiousness × Treatment			0.242
			(2.175)
Extraversion × Treatment			-1.169
			(1.535)
Agreeableness × Treatment			1.884
			(2.200)
Neuroticism × Treatment			0.517
			(1.759)
Observations	281	281	281
Adjusted R^2	-0.014	0.087	0.102

Standard errors in parentheses.
One observation is one subject.
Control variables for demographics, i.e., age, gender, and education level are included, but not shown in table.
$^{+}$ $p < 0.10$, $^{\square}$ $p < 0.05$, $^{\square\square}$ $p < 0.01$, $^{\square\square\square}$ $p < 0.001$

knowledge ($\beta = 2.652$, $p < 0.01$), which was consistent across treatments. Finally, the interaction analysis with the treatment showed that openness to experience showed a negative interaction effect such that openness is positively associated with metaknowledge in the Control group only (dydx = 4.272, $p < 0.01$).

Mediation analysis. Metaknowledge also helps to explain, i.e., mediates the impact of key personality traits on collaborative performance. We conducted

Table 4.5: Marginal e□ects of the respective personality trait on metaknowledge, model E3

	Control group		Treatment group	
	dy/dx	p-value	dy/dx	p-value
Openness	4.272	0.004	-0.585	0.627
Conscientiousness	4.030	0.005	4.272	0.013
Extraversion	-2.306	0.033	-3.475	0.002
Agreeableness	-0.170	0.914	1.715	0.272
Neuroticism	2.790	0.018	3.308	0.017

a causal mediation analysis to better understand the relationship between the respondents’ personality traits, metaknowledge, and collaborative performance. For this analysis, we are calculating bias-corrected and accelerated confidence intervals with 500 simulations. We report the results based on Model B2, focusing on conscientiousness and extraversion. Metaknowledge partially mediates the effect of conscientiousness (47%; Average Causal Mediated Effect = 1.590, $p < 0.01$; Average Direct Effect = 1.822, $p < 0.05$) and substantially mediates the effect of extraversion (54%; Average Causal Mediated Effect = -1.142, $p < 0.001$; Average Direct Effect = -0.985, $p < 0.05$) on the collaborative performance of humans and the AI. In other words, higher levels of conscientiousness and lower levels of extraversion are manifesting themselves in higher levels of metaknowledge, causing performance to improve. Next to this mediated effect, there are significant direct effects of conscientiousness and extraversion, respectively, on performance.

Deep-dive on user-provided certainty. Finally, to better understand metaknowledge, we also complete analyses at the image-level, using user-provided certainty in their initial classification choice as the dependent variable. This allowed us to differentiate between subjects’ ability to quantify their metaknowledge on those images which they classified correctly and those which they did not. We used a fixed effects regression on the image level for this purpose as an F test for individual effects and a Hausman test indicated this would be the appropriate way to model image-specific effects (Wooldridge 2013). In this way, we could isolate the

Table 4.6: Results of linear regression on user-provided certainty with fixed effects at the image-level

	User-Indicated Certainty	
	(1)	(2)
Constant	55.749	75.294
	(2.156)	(3.600)
Image sequence	-0.010	-0.010
	(0.004)	(0.004)
Treatment	1.350	1.326
	(0.242)	(0.243)
User choice correct	8.803	-6.806^{+}
	(0.868)	(3.587)
Openness	0.439	-0.013
	(0.153)	(0.473)
Conscientiousness	1.786	-0.692
	(0.254)	(0.577)
Extraversion	0.246	2.538
	(0.169)	(0.374)
Agreeableness	-0.553	-1.423
	(0.162)	(0.448)
Neuroticism	0.622	-0.504
	(0.195)	(0.421)
Openness × User choice correct		0.643
		(0.522)
Conscientiousness × User choice correct		3.474
		(0.591)
Extraversion × User choice correct		-3.102
		(0.387)
Agreeableness × User choice correct		1.240
		(0.491)
Neuroticism × User choice correct		1.533
		(0.433)
Observations	28,100	28,100
Clusters (images)	100	100
Subjects	281	281
Adjusted R^2	0.046	0.053

Robust standard errors in parentheses.
One observation is one image classified by one subject.
Control variables for demographics, i.e., age, gender, and education level are included, but not shown in table.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

specific effects arising from the fact that some images are more difficult for humans to classify correctly than others. The regressions included control variables for the Treatment, demographic characteristics (identical to prior models), and the sequence number (i.e., how many images the user had previously classified). They also include variables for all FFM traits and a binary variable indicating whether the class selected by the user (for which they indicated their certainty) was correct. Additionally, we included 2-way interaction effects between all five FFM traits and the dummy variable indicating whether the initial choice was correct or not. We report the results of these models in Table 4.6.

The marginal effects of this model are visualized graphically in the left-side of the panels in Figure 4.1. Additionally, we also ran a model with 3-way interaction effects with the dummy variable indicating treatment to discern specific effects of the manipulation. These are shown in the right part of the panels in Figure 4.1 and labeled as “Control” and “Treatment”, respectively. We show the predictive margins in Figure 4.2.

In support of our earlier findings that extraversion is inversely related with metaknowledge, that is, individuals with higher levels of extraversion have less metaknowledge, an analysis of the marginal effects shows that extraversion is positively and significantly associated with respondent-provided certainty for images they classified incorrectly (dydx = 2.538, $p < 0.001$), and negatively associated for images classified correctly (dydx = -0.563, $p < 0.001$). Put differently, for individuals with higher levels of extraversion, confidence is inversely associated with performance. Interestingly, the marginal effects of the three-way interaction indicate that providing the certainty of the AI does not help to better calibrate user certainty, but instead simply strengthens the effect of extraversion on confidence overall. This aggravates the problem for incorrectly classified images such that extraverts become more confident in their incorrect solutions.

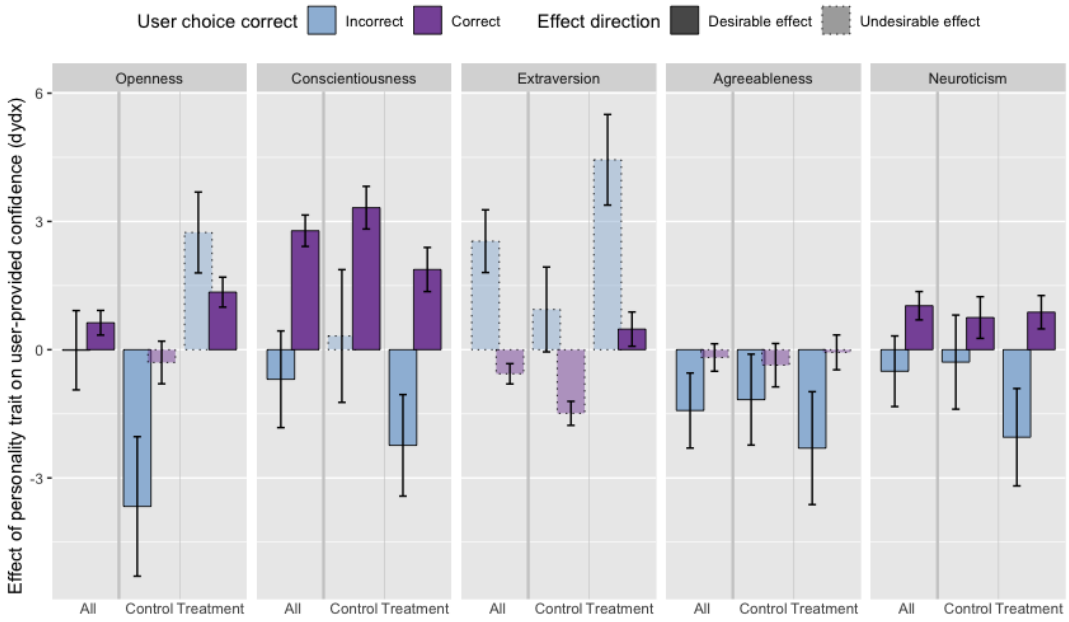

Figure 4.1: Marginal e□ects of personality traits on user-provided certainty, including 95% confidence interval. Models include the two-way interaction e□ects between personality traits and whether the user’s initial choice was correct (column “All”), as well as the three-way interaction between the personality traits, whether the respondent’s choice was correct, and the Treatment (columns “Control” and “Treatment”). N = 28,100

Concretely, the expected confidence of an individual with a high level of extraversion (i.e., one standard deviation above the mean) was 74% for incorrect and 82% for correct initial classifications. In the treatment condition, certainty for incorrect and correct classifications was 81% and 85%, respectively. In other words, not only does the confidence of extraverts in incorrect solutions increase, but also their ability to adequately distinguish what they know from what they do not (i.e., the delta) decreases. These findings clearly show that providing information about the confidence of an AI harms extraverts’ metaknowledge.

By contrast, conscientiousness is positively and significantly associated with user-provided certainty for images classified correctly (dydx = 2.782, $p < 0.001$), also when considering three-way interactions with the treatment. While the effect is somewhat weaker in the Treatment than in the Control group, the effect is consistently positive (and significant) in all conditions. The three-way interac-

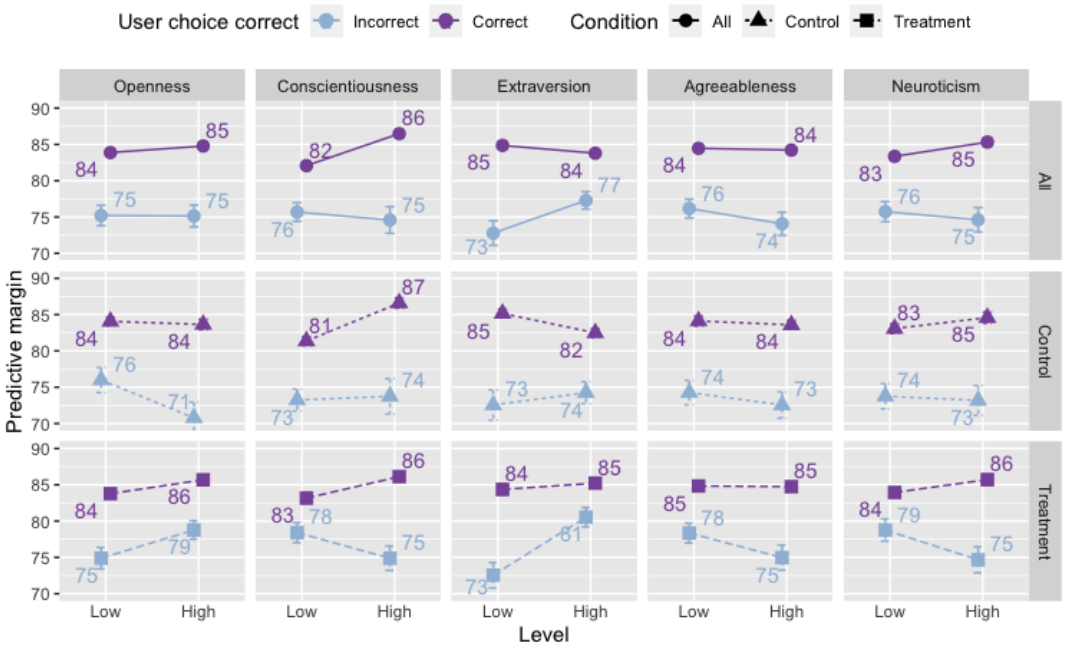

Figure 4.2: Predictive margins of personality traits on user-provided certainty, including 95% confidence interval. Models include the two-way interaction e□ects between personality traits and whether the user’s initial choice was correct (row “All”), as well as the three-way interaction between the personality traits, whether the respondent’s choice was correct, and the Treatment (rows “Control” and “Treatment”). N = 28,100

tion terms show that in the Treatment group (but not in the Control group), conscientiousness is negatively (i.e., desirably) linked to certainty for incorrectly classified images. A glance at Figure 4.2 translated this into tangible terms. In the treatment, the predicted certainty for individuals with low conscientiousness (i.e., one standard deviation below the mean) for correct and incorrect answers was 83% and 78%, respectively. For individuals with high levels of conscientiousness, these were 86% and 75%, respectively. These numbers clearly show that conscientiousness positively contributes to metaknowledge and that the Treatment overall helped conscientious individuals to improve their metaknowledge as expressed through their self-indicated confidence.

Results for individuals with low and high levels of neuroticism were similar to those for conscientiousness. Neuroticism is positively associated with confidence for images which respondents classified correctly (dydx = 1.028, $p < 0.001$).

While there was no significant effect for incorrect images in the Control group, the effect on certainty was negative (i.e., desirable) for incorrect images in the Treatment group (dydx = -2.047, $p < 0.001$). As Figure 4.2 illustrates, neuroticism — in particular in connection with the Treatment — appears to be beneficial for respondents’ metaknowledge. Although some prior research suggests that neuroticism is associated with more negative attitudes towards AI (Park and Woo 2022), our findings suggest that this personality trait is positively associated with metaknowledge.

Finally, we briefly comment on the findings with respect to Openness to Experience. Interestingly, the Treatment significantly harmed the ability of individuals who score highly on this dimension to quantify their certainty on images which they find difficult. Concretely, while an individual who scores highly on this trait (i.e., one standard deviation above the mean) provided an expected certainty of 71% for an image they classified incorrectly in the Control group, the certainty in the Treatment was 79%, thus greatly reducing the margin against those images classified correctly (see Figure 4.2 for details). While this finding clearly warrants further research, one explanation of this outcome may be that the mechanical nature of the Treatment condition ran counter to the creative and intellectually curious nature of individuals who score highly on this dimension (e.g., McCrae and Costa 1997, Sutin 2015)

4.5 Discussion

4.5.1 Implications

We study how differences in personality influence metaknowledge and human’s delegation behavior towards AI in an exploratory analyses of data from a controlled

onlinr experiment. Our findings demonstrate that differences in personality have a substantial impact on how humans work collaboratively with AI. Using the personality traits of the Five Factor Model (FFM), we demonstrate that personality impacts collaborative human-AI performance, specific delegation behavior and metaknowledge. Additionally, we show that personality can help explain the heterogeneity in how humans respond to receiving additional information about the AI, such as the algorithm’s certainty in it’s prediction.

We find particularly strong and consistent effects for conscientiousness and extraversion, which are two personality traits in the Five-Factor (FFM) model. While we detail our key findings with respect to these two traits below, it is important to note that we also find a number of interesting effects of the remaining three personality traits, in particular of neuroticism and openness to experience.

We find that higher levels of conscientiousness are associated with higher collaborative performance, an effect that can partially be explained through the higher levels of metaknowledge these individuals possess. While they do not delegate more overall, they take more rule-compliant delegation decisions, particularly not to delegate when they are more certain the AI is. This behavior appears consistent with the desire of conscientious individuals to stay in control (Jackson and Roberts 2015). However, they contradict and complement earlier findings by Tang et al. (2021), who suggested that conscientious individuals may benefit less from working with AI than others. Our analysis highlights a specific setting in which these individuals can in fact benefit from working with an AI more than others can. Understanding whether conscientiousness is in fact beneficial for humans working with AI in line with how it has traditionally been linked positively with various work-related outcomes (e.g. Siebert and DeGeest 2015) is a particularly topical question.

In our analyses, extraversion negatively impacted participants’ ability to work

in ways that leverage complementarities with the AI. In large parts, this can be explained by the lower levels of metaknowledge these individuals exhibit. While we do not explicitly capture participants overconfidence, it is intuitive that this effect is indeed an artifact of extraverts’ propensity for overconfidence (Schaefer et al. 2004, Wilt and Revelle 2016). Concerningly, higher levels of extraversion are not only associated with greater confidence overall but greater confidence in incorrect decisions in particular. The fact that this effect that is only strengthened by providing information about the certainty of the AI (i.e., our Treatment) suggests the need to more closely investigate the effects that providing certainty has on humans’ mental models (Bansal et al. 2019).

These findings have important implications for theory and practice. Firstly, they indicate that some personality types may be able to more effectively capture complementarities between themselves and the AI than others. In the absence of successful interventions to address negative impacts of personality traits like extraversion on metaknowledge and collaborative performance, this may make some people less, and others more suitable to work with an AI system in a human-to-AI delegation setting. Secondly, they indicate the importance of metaknowledge to successful collaboration between humans and AI. This calls for researchers and practitioners to find ways to refine ways of measuring metaknowledge, and, above all, to find ways to increase metaknowledge of those individuals working with AI.

4.5.2 Limitations and Avenues for Future Research

The findings of our study should be viewed in the context of its limitations, which present numerous opportunities for future research. A clear limitation is that we our study is exploratory in nature, based on data from an existing online experiment. Our analyses generates a number of strong and interesting findings which researchers may want to replicate in additional settings in which

humans and AI work together collaboratively. This would also allow researchers to complement and validate the prior findings by Fügener et al. (2021, 2022) and others in a context that is not image classification-based.

Furthermore, our findings open up a number of avenues for future research. Perhaps most obvious is the opportunity — and indeed the need — to understand whether, and how, interactions with AI may be tailored to different personality traits and constellations of personality traits. Here, researchers will not only want to find ways to illicit positive effects arising from personality traits like conscientiousness (as our Treatment condition partially achieved), but also to prevent or dampen the negative effects on metaknowledge and performance arising from personality traits like extraversion. Clearly, given that personality is a unique aspect that makes us humans, differences arising from personality should be considered when working with machines.

Appendix of Chapter 4

4.A Experimental Design

The figures below show the two steps completed for each of the 100 images in the experiment. In the first step (shown in Figure 4.3), participants selected the class they think best describes the image. Additionally, they indicate their certainty between 0 and 100 with a slider at the bottom of the page.

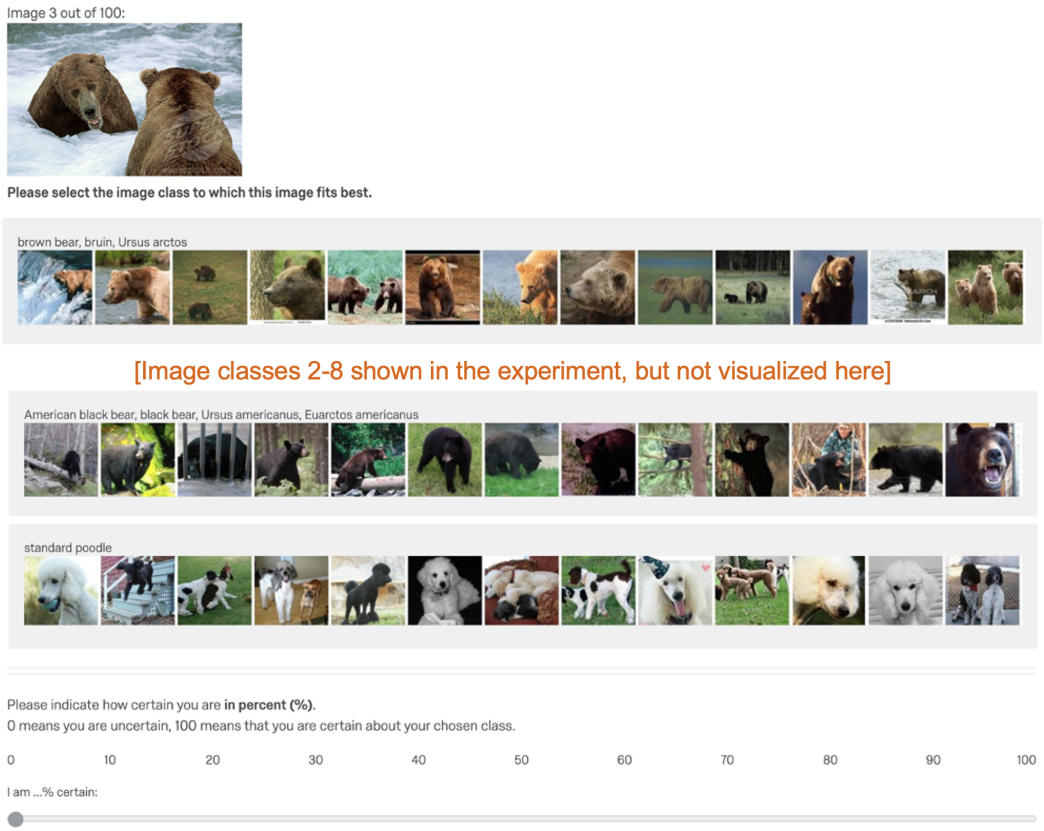

Figure 4.3: Screenshot of the first of two screens shown for each of the 100 images classified by respondents.

The second step was varied based on whether participants were randomly selected into the Control or Treatment groups. Figure 4.4 shows the screen shown to the Control group, Figure 4.4 was shown to participants in the Treatment group. The only difference between the screens is that the AIs certainty, as well as the certainty the user indicated themselves in step 1, are shown in the screen.

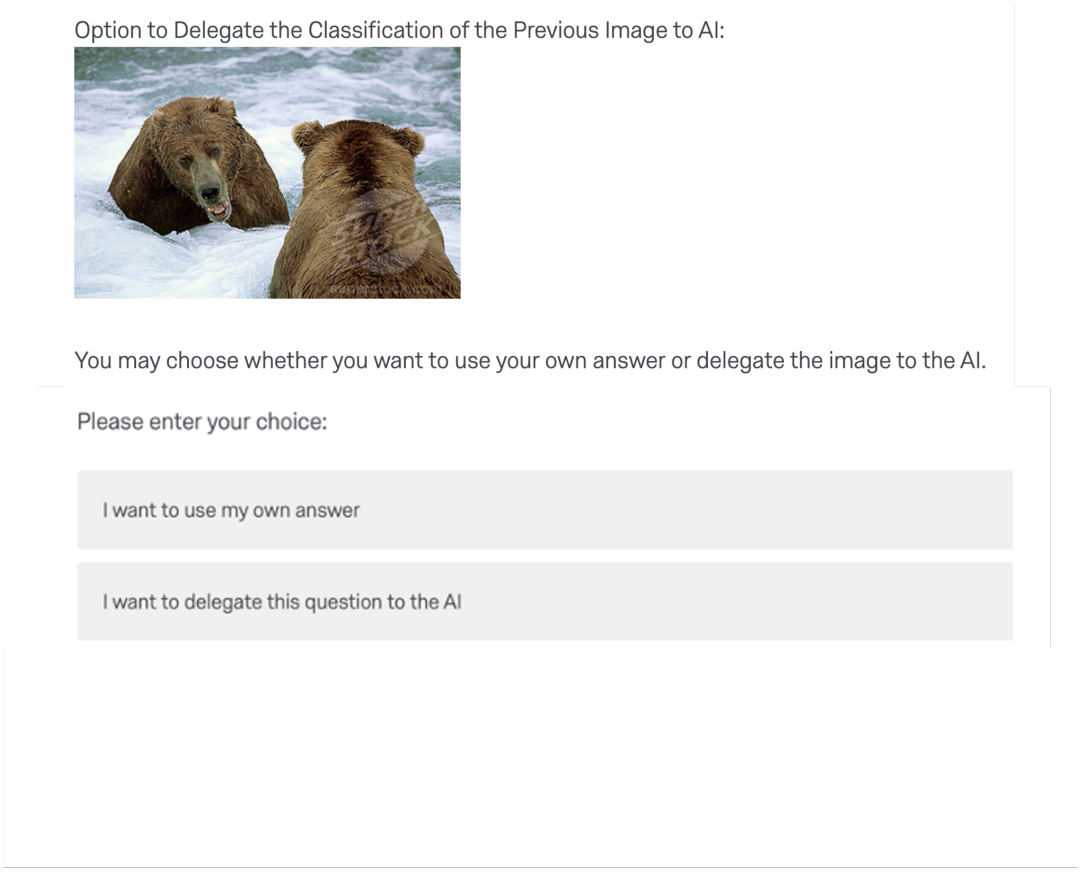

Figure 4.4: Screenshot of the second of two screens shown for each of the 100 images classified by respondents for any respondents randomly selected into the Control group.

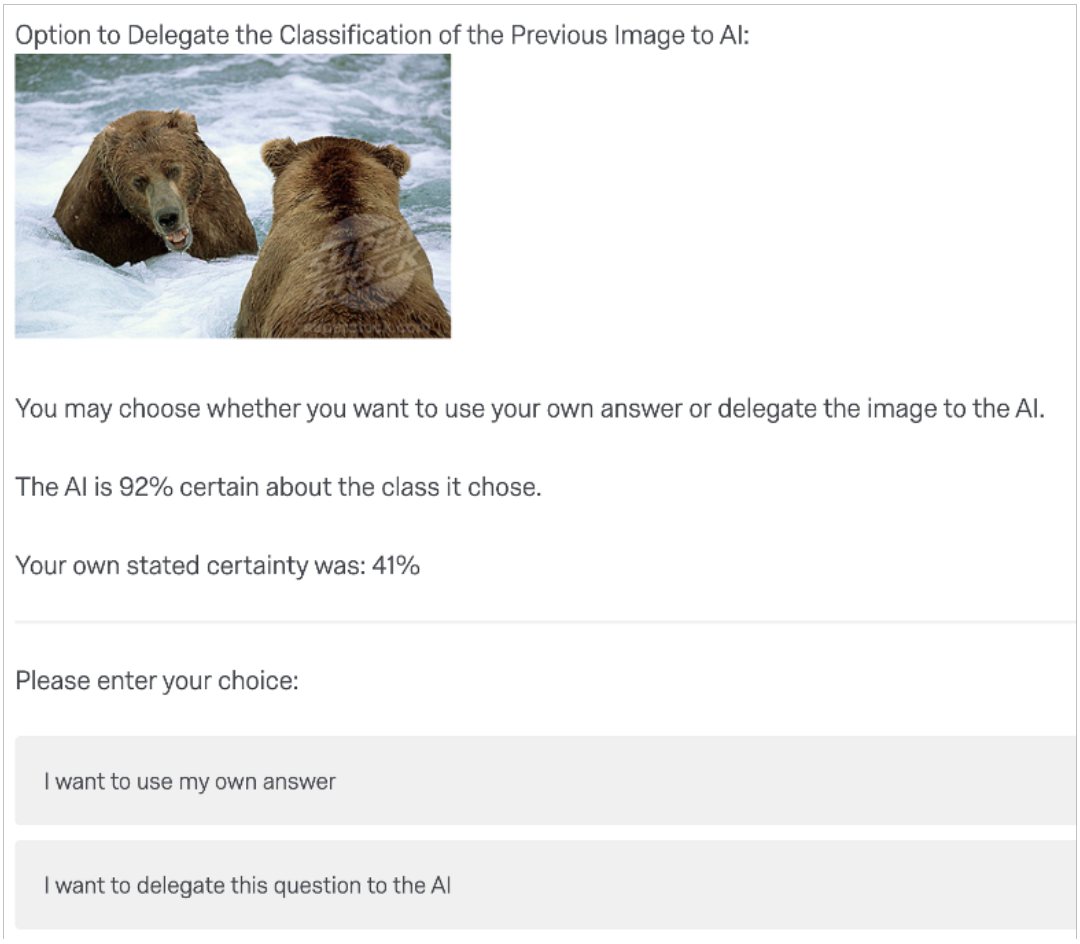

Figure 4.5: Screenshot of the second of two screens shown for each of the 100 images classified by respondents for any respondents randomly selected into the Treatment group.

Chapter 5

Conclusion

This dissertation aims at improving our understanding of the role and impact of personality in operations management. To do so, we take a specific focus on product recalls, which is an important aspect relating to firms' operations. In three research projects, we study how the individual differences arising from personality affect important outcomes in operations management, answering calls to "incorporate individual differences into OM research" (Bendoly et al. 2006, p. 741). In Chapter 2, we investigate how CEO regulatory focus affects the number of product recalls and product innovations. We build on findings from this work in Chapter 3 and investigate the effect of CEO neuroticism, an important personality trait, on handling of product recall timing. Finally, in Chapter 4, we investigate the impact of personality traits on behaviour within a different context which is highly topical for the operations management literature.

Our research shows that aspects of individual's personality such as their regulatory focus or neuroticism are very relevant predictors of their behaviour, which shapes important outcomes in the field of operations management broadly, and product recalls specifically. Our research also shows this impact at different levels. On the one hand, our work demonstrates effects of personality on the behaviour of individuals (e.g., Chapter 4). On the other hand, our work on personality of CEOs in Chapters 2 and 3 shows that personality is such a powerful factor that the personality of one individual — albeit the personality of important strategic leader — can have wide-ranging impacts on firm-level outcomes. While the wide-ranging impact of personality is well established in the psychology literature, where a large body of research demonstrates that an individual's personality is a dominant predictor of their cognitions and behaviors (Schneider and Smith 2004), these insights are largely new in the operations management literature.

The remainder of this chapter is structured as follows. In Section 5.1, we summarize the key results of the three research projects. In Section 5.2, we critically review our findings and provide an outlook for future research.

5.1 Key Results

In **Chapter 2**, we provided insights on the role of CEOs' regulatory focus, an important aspect relating to individuals' personality, on product innovation and product safety problems. Using a panel of healthcare firms listed in the S&P 1500 health care index, we found that prevention and promotion focus have opposing effects on product innovations and product recalls. Interestingly, our results suggested that the achievement orientation emblematic of promotion focus simultaneously has desirable and undesirable effects for firms. On the one hand, this focus was associated with more successful product innovation output, but at the same time also linked to more product safety problems. While prevention focus also had counter-acting effects, this focus (and its associated focus on safety, protection and vigilance) acted in the opposite way. Concretely, we found that higher CEO prevention focus was linked with fewer product safety problems, but also — at least in some operationalizations — linked with lower output of product innovations. The results indicated that personality traits of strategic leaders like CEOs, which have so far been particularly overlooked in research on product recalls, can help to explain firms' safety record. Finally, we found that the situation where CEOs also hold the title of COO, which has largely been left unexplored by research, is particularly important to understand the impact of CEO prevention focus on product recalls.

In **Chapter 3**, we built on the findings of Chapter 2 and conducted a deep-dive on the impact of CEO personality on product recalls. Concretely, this project

investigated the handling of individual recall events, rather than the total number of product recalls a firm experiences. We focused specifically on recall timing, or time-to-recall, which describes the amount of time that elapsed between the moment that a firm became aware of a safety problem and the point in time at which they issued a recall for the affected product. Clearly, this aspect of recall timing is critically important for firms and their stakeholders for a variety of reasons. We built on a large body of research around the Five-Factor Model (FFM) of personality traits to hypothesize a relationship between neuroticism and time-to-recall. In line with our theoretical predictions that more neurotic CEOs (i.e., those CEOs with higher levels of neuroticism, which is a key personality trait) are more vigilant when they face uncertainty, as they do when making recall decisions, and strive to avoid the potential threat of social disapproval from recalling late, we found that CEO neuroticism is negatively associated with time-to-recall. Put differently, the personality dispositions of neurotic CEOs means they follow a "better safe than sorry" strategy when it comes to handling product recalls. We contextualized our argumentation and findings by considering two moderators. We found that CEO agreeableness, another FFM personality trait, positively moderates the main relationship. We hypothesized that this is because agreeableness increases the salience of the threat of social disapproval. Additionally, we found that situational factors, specifically more negative recent media coverage of a CEO, also positively moderate the main relationship.

Finally, in **Chapter 4**, we built on our findings on the importance of personality and investigated the role of the FFM personality traits within a different context. We answered calls to investigate the "the interaction between humans and machines" (Samson and Kalchschmidt 2019, p. 1), a highly relevant setting given that advancements in AI presents a "major development for operations management" (Loske and Klumpp 2021, p.1) and poses a number of unanswered questions on how humans and algorithms can effectively collaborate. Using data

from an online experiment on image classification, we found that higher levels of conscientiousness improve, while higher levels extraversion harm humans' ability to work collaboratively with AI. Our analyses showed that differences in meta-knowledge, which is an individual's ability to assess what they know (and do not know), help to explain this relationship. Furthermore, we explored the effect of providing humans with information on the AI's certainty. We were able to use individuals personality traits — especially conscientiousness and extraversion — to explain heterogeneity in the effect that this intervention had on metaknowledge.

5.2 Critical Review and Future Research

The research presented in this thesis contributes to the literature on operations management and highlights the role of individuals' personality in explaining important outcomes in operations management, focusing specifically on product recalls. While our papers used different methodology, all three research projects focused on the personality of a single individual. Clearly, this focus enabled our theorizing to be more targeted and alleviates potential difficulties in data gathering. However, in reality, individuals in OM (like in many other fields) rarely work alone. This is the case no matter whether the individual is a CEO — who will work with other strategic leaders like the Chief Operating Officer (COO), Chief Financial Officer (CFO), and Chief Supply Chain Officer (CSCO) — or an employee at other hierarchical levels of an organization. Therefore, it could be promising for future research to investigate how combinations or constellations of personalities of several individuals, such as CEO-COO duos (Hambrick and Cannella 2004), may have different impact than the personalities of individuals have in isolation.

In order to robustly estimate the effects of personality traits and reduce the

potential for biases, we conducted our analyses within specific contexts. In Chapters 2 and 3, this was the healthcare industry, which provides a relatively homogeneous regulatory and operating environment. In Chapter 4, we focused on image classification. While this focus is not without precedent (e.g., Wowak et al. 2021, Fügener et al. 2021), our focus also presents an opportunity for future research to validate our findings in other contexts, such as different industries or different cognitive tasks.

Finally, the methods used in the different chapters of this dissertation each come with some limitations that present opportunities for complementary future research which uses different methodological approaches. In Chapters 2 and 3, we explained outcomes at the firm-level by recourse to the personality of strategic leaders. While we use a number of well-tested empirical tools (e.g., time lags, fixed effects, focusing on recalls which CEOs oversee, etc.) to try to ensure that we are in fact measuring the effect of CEO personality, such an analysis cannot account for all possible confounding factors. Also, it cannot detail out the exact means through which personality in fact changes the outcomes we observed. Accordingly, it may be interesting for researchers to complement our findings by studying the impact of CEO personality on recalls using qualitative and ethnographic research. In contrast, Chapter 4 presents different opportunities. Given that this is an experimental online study, the impacts we observed can more readily be explained by personality and the Treatment condition. However, because multiple environmental and interpersonal influences are present in real life, researchers may want to use field experiments to provide complementary insights on the role of personality.

References

Aktan, A. C., Castellucci, F., and Menz, M. (2021). Chief innovation o□cers: An empirical study of behavioral and institutional antecedents. In *Academy of Management Proceedings*, volume 2021, page 10577. Academy of Management.

Allik, J. and Realo, A. (2015). Universal and Specific in the Five Factor Model of Personality. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*. Oxford University Press.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609.

American Psychological Association (2022). Psychology Topics: Personality. https://www.apa.org/topics/personality.

Améndola, L., Weary, D., and Zobel, G. (2022). E□ects of personality on assessments of anxiety and cognition. *Neuroscience & Biobehavioral Reviews*, 141:104827.

Anand, G. and Gray, J. V. (2017). Strategy and organization research in operations management. *Journal of Operations Management*, 53-56(1):1–8.

Andrei, A. G., van Oosterhout, J. H., and Sauerwald, S. (2022). Symbolic Shareholder Democracy: Toward a Behavioral Understanding of the Role of Shareholder Voting in CEO Dismissals. *Organization Science*, Articles In Advance:1–23.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Bainbridge, T. F., Ludeke, S. G., and Smillie, L. D. (2022). Evaluating the Big Five as an organizing framework for commonly used psychological trait scales. *Journal of Personality and Social Psychology*, 122(4):749–777.

Baird, A. and Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, 45(1):315–341.

Balkin, D. B., Markman, G. D., and Gomez-Mejia, L. R. (2000). Is CEO pay in high-technology firms related to innovation? *Academy of Management Journal*, 43(6):1118–1129.

Ball, G., Macher, J. T., and Stern, A. D. (2018a). Recalls, Innovation, and Competitor Response: Evidence from Medical Device Firms. *Harvard Business School Working Paper*, No. 19-028.

Ball, G. P., Shah, R., and Donohue, K. (2018b). The decision to recall: A behavioral investigation in the medical device industry. *Journal of Operations Management*, 62:1–15.

Ball, G. P., Shah, R., and Wowak, K. D. (2018c). Product competition, managerial discretion, and manufacturing recalls in the U.S. pharmaceutical industry. *Journal of Operations Management*, 58-59(1):59–72.

Ballinger, G. A. (2004). Using Generalized Estimating Equations for Longitudinal Data Analysis. *Organizational Research Methods*, 7(2):127–150.

Bansal, G., Nushi, B., Kamar, E., Weld, D. S., Lasecki, W. S., and Horvitz, E. (2019). Updates in Human-AI Teams: Understanding and Addressing the Performance/Compatibility Tradeo□. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:2429–2437.

Bednar, M. K. (2012). Watchdog or Lapdog? A Behavioral View of the Media as a Corporate Governance Mechanism. *Academy of Management Journal*, 55(1):131–150.

Bendersky, C. and Shah, N. P. (2013). The Downfall of Extraverts and Rise of Neurotics: The Dynamic Process of Status Allocation in Task Groups. *Academy of Management Journal*, 56(2):387–406.

Bendoly, E., Donohue, K., and Schultz, K. L. (2006). Behavior in operations management: Assessing recent findings and revisiting old assumptions. *Journal of Operations Management*, 24(6):737–752.

Benischke, M. H., Martin, G. P., and Glaser, L. (2019). CEO equity risk bearing and strategic risk taking: The moderating e□ect of CEO personality. *Strategic Management Journal*, 40(1):153–177.

Berenbaum, H., Bredemeier, K., and Thompson, R. J. (2008). Intolerance of uncertainty:

Exploring its dimensionality and associations with need for cognitive closure, psychopathology, and personality. *Journal of Anxiety Disorders*, 22(1):117–125.

Berman, B. (1999). Planning for the inevitable product recall. *Business Horizons*, 42(2):69–78.

Bono, J. E. and Judge, T. A. (2004). Personality and Transformational and Transactional Leadership: A Meta-Analysis. *Journal of Applied Psychology*, 89(5):901–910.

Bowers, K. S. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review*, 80(5):307–336.

Brockner, J. and Higgins, E. (2001). Regulatory Focus Theory: Implications for the Study of Emotions at Work. *Organizational Behavior and Human Decision Processes*, 86(1):35–66.

Brockner, J., Higgins, E. T., and Low, M. B. (2004). Regulatory focus theory and the entrepreneurial process. *Journal of Business Venturing*, 19(2):203–220.

Busenbark, J. R., Frank, K. A., Maroulis, S. J., Xu, R., and Lin, Q. (2021). Quantifying the Robustness of Empirical Inferences in Strategic Management: The Impact Threshold of a Confounding Variable and Robustness of Inference to Replacement. In Hill, A. D., Lê, J. K., McKenny, A. F., O'Kane, P., Paroutis, S., and Smith, A. D., editors, *Research Methodology in Strategy and Management*, pages 123–150. Emerald Publishing Limited.

Busenbark, J. R., Yoon, H. E., Gamache, D. L., and Withers, M. C. (2022). Omitted Variable Bias: Examining Management Research With the Impact Threshold of a Confounding Variable (ITCV). *Journal of Management*, 48(1):17–48.

Cable, D. M. and Judge, T. A. (2003). Managers' upward influence tactic strategies: the role of manager personality and supervisor leadership style. *Journal of Organizational Behavior*, 24(2):197–214.

Certo, S. T., Withers, M. C., and Semadeni, M. (2017). A tale of two e□ects: Using longitudinal data to compare within- and between-firm e□ects: Longitudinal Relationships in Strategy Research. *Strategic Management Journal*, 38(7):1536–1556.

Chatterjee, A. and Hambrick, D. C. (2007). It's All about Me: Narcissistic Chief Executive O□cers and Their E□ects on Company Strategy and Performance. *Administrative Science Quarterly*, 52(3):351–386.

Chatterji, A. K. and Fabrizio, K. R. (2014). Using users: When does external knowledge enhance corporate product innovation? *Strategic Management Journal*, 35(10):1427–1445.

Chen, G., Meyer-Doyle, P., and Shi, W. (2018). How CEO and CFO Regulatory Focus interact to shape the Firm's Corporate Strategy. In *Behavioral Strategy for Competitive Advantage*, pages 37–73. Information Age Publishing, Charlotte, NC.

Cohen, M., Cui, S., Doetsch, S., Ernst, R., Huchzermeier, A., Kouvelis, P., Lee, H., Matsuo, H., and Tsay, A. A. (2022). Bespoke supply-chain resilience: The gap between theory and practice. *Journal of Operations Management*, 68(5):515–531.

Colbert, A. E., Barrick, M. R., and Bradley, B. H. (2014). Personality And Leadership Composition in Top Management Teams: Implications For Organizational E□ectiveness. *Personnel Psychology*, 67(2):351–387.

Costa, P. T. and McCrae, R. R. (1985). *The NEO personality inventory*. Psychological assessment resources, Odessa, FL.

Costa, P. T. and McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). In *The SAGE Handbook of Personality Theory and Assessment: Volume 2 — Personality Measurement and Testing*, pages 179–198. SAGE Publications Ltd, London.

Costa, P. T., McCrae, R. R., and Dye, D. A. (1991). Facet Scales for Agreeableness and Conscientiousness: A Revision of the NEO Personality Inventory. *Personality and Individual Di erences*, 12(9):887–898.

Cronqvist, H., Makhija, A. K., and Yonker, S. E. (2012). Behavioral consistency in corporate finance: CEO personal and corporate leverage. *Journal of Financial Economics*, 103(1):20–40.

Crowe, E. and Higgins, E. T. (1997). Regulatory Focus and Strategic Inclinations:

Promotion and Prevention in Decision-Making. *Organizational Behavior and Human Decision Processes*, 69(2):117–132.

Cunningham, W. A., Raye, C. L., and Johnson, M. K. (2005). Neural correlates of evaluation associated with promotion and prevention regulatory focus. *Cognitive, A ective, & Behavioral Neuroscience*, 5(2):202–211.

de Bruin, G. O., Rassin, E., and Muris, P. (2007). The Prediction of Worry in Non-Clinical Individuals: The Role of Intolerance of Uncertainty, Meta-Worry, and Neuroticism. *Journal of Psychopathology and Behavioral Assessment*, 29(2):93–100.

Deephouse, D. L. and Suchman, M. (2008). Legitimacy in Organizational Institutionalism. In *The SAGE Handbook of Organizational Institutionalism*, pages 49–77. SAGE Publications Ltd, London.

Denissen, J. J. A. and Penke, L. (2008a). Motivational individual reaction norms underlying the Five-Factor model of personality: First steps towards a theory-based conceptual framework. *Journal of Research in Personality*, 42(5):1285–1302.

Denissen, J. J. A. and Penke, L. (2008b). Neuroticism predicts reactions to cues of social inclusion. *European Journal of Personality*, 22(6):497–517.

DesJardine, M. R. and Shi, W. (2021). How Temporal Focus Shapes the Influence of Executive Compensation on Risk Taking. *Academy of Management Journal*, 64(1):265–292.

DeYoung, C. G., Quilty, L. C., and Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93(5):880–896.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *The Academy of Management Review*, 14(1):57.

Elliot, A. J. and Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, 82(5):804–818.

Evans, J. A. and Foster, J. G. (2011). Metaknowledge. *Science*, 331(6018):721–725.

Exhibits (2023). Title 17, Chapter II, Part 229. *Code of Federal Regulations*, page 17 CFR § 229.601.

FDA (2010). FDA 101: Product Recalls. http://www.fda.gov/consumers/consumer-updates/fda-101-product-recalls.

FDA (2019). Premarket Approval (PMA). https://www.fda.gov/medical-devices/premarket-submissions-selecting-and-preparing-correct-submission/premarket-approval-pma.

FDA (2020). Regulatory Procedures Manual. https://www.fda.gov/media/71814/download.

FDA (2022). How to Study and Market Your Device. https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/how-study-and-market-your-device.

Finkelstein, S. and Hambrick, D. C. (1990). Top-Management-Team Tenure and Organizational Outcomes: The Moderating Role of Managerial Discretion. *Administrative Science Quarterly*, 35(3):484–503.

Finkelstein, S., Hambrick, D. C., and Cannella, A. A. (2009). *Strategic leadership: theory and research on executives, top management teams, and boards*. Strategic management series. Oxford University Press, New York.

Florack, A. and Hartmann, J. (2007). Regulatory focus and investment decisions in small groups. *Journal of Experimental Social Psychology*, 43(4):626–632.

Förster, J., Higgins, E., and Bianco, A. T. (2003). Speed/accuracy decisions in task performance: Built-in trade-o or separate strategic concerns? *Organizational Behavior and Human Decision Processes*, 90(1):148–164.

Frank, K. A. (2000). Impact of a Confounding Variable on a Regression Coecient. *Sociological Methods & Research*, 29(2):147–194.

Frank, K. A., Lin, Q., Xu, R., Maroulis, S., and Mueller, A. (2023). Quantifying the robustness of causal inferences: Sensitivity analysis for pragmatic social science. *Social Science Research*, 110:102815.

Frank, K. A., Maroulis, S. J., Duong, M. Q., and Kelcey, B. M. (2013). What Would It Take to Change an Inference? Using Rubin's Causal Model to Interpret the

Robustness of Causal Inferences. *Educational Evaluation and Policy Analysis*, 35(4):437–460.

Fügener, A., Grahl, J., Gupta, A., and Ketter, W. (2021). Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI. *MIS Quarterly*, 45(3):1527–1556.

Fügener, A., Grahl, J., Gupta, A., and Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, 33(2):678–696.

Fung, H. and Cucho, A. (2018). Everything You Need to Know About the Implant Files. *The International Consortium of Investigative Journalists*.

Gamache, D. L. and McNamara, G. (2019). Responding to Bad Press: How CEO Temporal Focus Influences the Sensitivity to Negative Media Coverage of Acquisitions. *Academy of Management Journal*, 62(3):918–943.

Gamache, D. L., McNamara, G., Mannor, M. J., and Johnson, R. E. (2015). Motivated to Acquire? The Impact of CEO Regulatory Focus on Firm Acquisitions. *Academy of Management Journal*, 58(4):1261–1282.

Gamache, D. L., Neville, F., Bundy, J., and Short, C. E. (2020). Serving di□erently: CEO regulatory focus and firm stakeholder strategy. *Strategic Management Journal*, 41(7):1305–1335.

Gass, D. F. (2023). Exploring personality-based heterogeneity in metaknowledge and human-AI collaboration. In *ICIS 2023 Proceedings*. Association for Information Systems.

Gass, D. F., Fügener, A., and Graf-Vlachy, L. (2022). Better Safe Than Sorry: How CEO Neuroticism A□ects Time-To-Recall in Product Recalls. *Academy of Management Proceedings*, 2022(1):11159.

Gleason, K. A., Jensen-Campbell, L. A., and South Richardson, D. (2004). Agreeableness as a predictor of aggression in adolescence. *Aggressive Behavior*, 30(1):43–61.

Goebel, P., Reuter, C., Pibernik, R., Sichtmann, C., and Bals, L. (2018). Purchasing

managers' willingness to pay for attributes that constitute sustainability. *Journal of Operations Management*, 62(1):44–58.

Goldbach, C., Kayar, D., Pitz, T., and Sickmann, J. (2019). Transferring decisions to an algorithm: A simple route choice experiment. *Transportation Research Part F: Tra c Psychology and Behaviour*, 65:402–417.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., and Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1):84–96.

Graf-Vlachy, L., Bundy, J., and Hambrick, D. C. (2020). E□ects of an Advancing Tenure on CEO Cognitive Complexity. *Organization Science*, 31(4):797–1051.

Graziano, W. G. and Eisenberg, N. (1997). Agreeableness: A dimension of personality. In Hogan, R., Johnson, J. A., and Briggs, S. R., editors, *Handbook of Personality Psychology*, pages 795–824. Elsevier.

Graziano, W. G. and Tobin, R. M. (2009). Agreeableness. In Leary, M. R. and Hoyle, R. H., editors, *Handbook of individual di erences in social behavior*, pages 46–61. Guilford Press, New York, NY.

Graziano, W. G. and Tobin, R. M. (2016). Agreeableness and the Five Factor Model. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*. Oxford University Press.

Greising, D. (2001). Baxter gets grip on crisis, responsibility. *Chicago Tribune*, pages (November 7), Business.

Greising, D. and Japsen, B. (2008). Flood of drugs, little oversight. *Chicago Tribune*, pages (March 03), Business.

Gupta, A., Nadkarni, S., and Mariam, M. (2019). Dispositional Sources of Managerial Discretion: CEO Ideology, CEO Personality, and Firm Strategies. *Administrative Science Quarterly*, 64(4):855–893.

Hambrick, D. C. (2007). Upper Echelons Theory: An Update. *Academy of Management Review*, 32(2):334–343.

Hambrick, D. C. and Cannella, A. A. (2004). CEOs who have COOs: contingency analysis of an unexplored structural form. *Strategic Management Journal*, 25(10):959–979.

Hambrick, D. C. and Mason, P. A. (1984). Upper Echelons: The Organization as a Reflection of Its Top Managers. *Academy of Management Review*, 9(2):193–206.

Hambrick, D. C. and Quigley, T. J. (2014). Toward more accurate contextualization of the CEO e□ect on firm performance: Toward More Accurate Contextualization of the CEO E□ect. *Strategic Management Journal*, 35(4):473–491.

Harrison, J., Thurgood, G. R., Boivie, S., and Pfarrer, M. (2020). Perception Is Reality: How CEOs' Observed Personality Influences Market Perceptions of Firm Risk and Shareholder Returns. *Academy of Management Journal*, 63(4):1166–1195.

Harrison, J. S., Boivie, S., Sharp, N. Y., and Gentry, R. J. (2018). Saving Face: How Exit in Response to Negative Press and Star Analyst Downgrades Reflects Reputation Maintenance by Directors. *Academy of Management Journal*, 61(3):1131–1157.

Harrison, J. S., Boivie, S., and Withers, M. C. (2022). Executives' Prior Employment Ties to Interlocking Directors and Interfirm Mobility. *Organization Science*, Articles in Advance:1–24.

Harrison, J. S., Thurgood, G. R., Boivie, S., and Pfarrer, M. D. (2019). Measuring CEO personality: Developing, validating, and testing a linguistic tool. *Strategic Management Journal*, 40(8):1316–1330.

Haunschild, P. R. and Rhee, M. (2004). The Role of Volition in Organizational Learning: The Case of Automotive Product Recalls. *Management Science*, 50(11):1545–1560.

Hayward, M. L. A., Rindova, V. P., and Pollock, T. G. (2004). Believing one's own press: the causes and consequences of CEO celebrity. *Strategic Management Journal*, 25(7):637–653.

Hensley, S. (2010). FDA: Johnson & Johnson Concealed Motrin Recall. *NPR*.

Herrmann, P. and Nadkarni, S. (2014). Managing strategic change: The duality of CEO personality. *Strategic Management Journal*, 35(9):1318–1342.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52(12):1280–1300.

Higgins, E. T. (1998). Promotion and Prevention: Regulatory Focus as A Motivational Principle. In Zanna, M. P., editor, *Advances in Experimental Social Psychology*, volume 30, pages 1–46. Academic Press.

Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist*, 55(11):1217–1230.

Higgins, E. T. (2014). Promotion and prevention: How "0" can create dual motivational forces. In Sherman, J. W., Gawronski, B., and Trope, Y., editors, *Dual-Process Theories of the Social Mind*, pages 423–435. The Guilford Press, New York.

Higgins, E. T. (2015). Regulatory Focus Theory. In Scott, R. A. and Kosslyn, S. M., editors, *Emerging Trends in the Social and Behavioral Sciences*, pages 1–18. Wiley, first edition.

Higgins, E. T. and Cornwell, J. F. (2016). Securing foundations and advancing frontiers: Prevention and promotion e□ects on judgment & decision making. *Organizational Behavior and Human Decision Processes*, 136:56–67.

Higgins, E. T., Idson, L. C., Freitas, A. L., Spiegel, S., and Molden, D. C. (2003). Transfer of value from fit. *Journal of Personality and Social Psychology*, 84(6):1140–1153.

Higgins, E. T., Shah, J., and Friedman, R. (1997). Emotional responses to goal attainment: Strength of regulatory focus as moderator. *Journal of Personality and Social Psychology*, 72(3):515–525.

Hill, A. D., Petrenko, O. V., Ridge, J. W., and Aime, F. (2019). Videometric Measurement of Individual Characteristics in Di□cult to Access Subject Pools: Demonstrating with Ceos. In Boyd, B., Crook, T. R., Lê, J. K., and Smith, A. D., editors, *Research Methodology in Strategy and Management*, volume 11, pages 39–61. Emerald Publishing Limited.

Hills, P. and Argyle, M. (2001). Emotional stability as a major dimension of happiness. *Personality and Individual Di erences*, 31(8):1357–1364.

Hirsh, J. B. and Inzlicht, M. (2008). The Devil You Know: Neuroticism Predicts Neural Response to Uncertainty. *Psychological Science*, 19(10):962–967.

Hirshleifer, D., Low, A., and Teoh, S. H. (2012). Are Overconfident CEOs Better Innovators? *The Journal of Finance*, 67(4):1457–1498.

Hofstee, W. K., de Raad, B., and Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63(1):146–163.

Hogan, R., Hogan, J., and Roberts, B. W. (1996). Personality measurement and employment decisions: Questions and answers. *American Psychologist*, 51(5):469–477.

Hora, M., Bapujib, H., and Rothc, A. V. (2011). Safety hazard and time to recall: The role of recall strategy, product defect type, and supply chain player in the U.S. toy industry. *Journal of Operations Management*, 29(1):766–777.

Jackson, J. J. and Roberts, B. W. (2015). Conscientiousness. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*, volume 1. Oxford University Press.

Janis, I. L. and Fadner, R. H. (1943). A coe□cient of imbalance for content analysis. *Psychometrika*, 8(2):105–119.

John, O. P., Donahue, E., and Kentle, R. (1991). *The Big Five Inventory–versions 4a and 54*. University of California, Berkeley, Institute of Personality and Social Research, Berkeley, CA.

John, O. P. and Srivastava, P. (1991). *The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspective*, pages 102–138. Guilford Press, New York.

Johnson, R. E., King, D. D., (Joanna) Lin, S.-H., Scott, B. A., Jackson Walker, E. M., and Wang, M. (2017). Regulatory focus trickle-down: How leader regulatory focus and behavior shape follower regulatory focus. *Organizational Behavior and Human Decision Processes*, 140:29–45.

Judge, T. A., Bono, J. E., Ilies, R., and Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4):765–780.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., and Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology*, 98(6):875–925.

Juran, D. C. and Schruben, L. W. (2004). Using worker personality and demographic information to improve system performance prediction. *Journal of Operations Management*, 22(4):355–367.

Kammerlander, N., Burger, D., Fust, A., and Fueglistaller, U. (2015). Exploration and exploitation in established small and medium-sized enterprises: The e□ect of CEOs' regulatory focus. *Journal of Business Venturing*, 30(4):582–602.

Kashmiri, S., Gala, P., and Nicol, C. D. (2019). Seeking pleasure or avoiding pain: Influence of CEO regulatory focus on firms' advertising, R&D, and marketing controversies. *Journal of Business Research*, 105:227–242.

Khurana, R. (2002). The Curse of the Superstar CEO. *Harvard Business Review*, 80(9):60–66.

Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W., editors (2005). *Applied linear statistical models*. The McGraw-Hill/Irwin series operations and decision sciences. McGraw-Hill Irwin, Boston, 5th ed edition.

Lai, V. and Tan, C. (2019). On Human Predictions with Explanations and Predictions of Machine Learning Models: A Case Study on Deception Detection. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 29–38, Atlanta, GA. ACM.

Lanaj, K., Chang, C.-H. D., and Johnson, R. E. (2012). Regulatory focus and work-related outcomes: A review and meta-analysis. *Psychological Bulletin*, 138:998–1034.

Larcker, D. F. and Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, 49(3):186–205.

Larsen, R. J. and Buss, D. M. (2010). *Personality Psychology: Domains of Knowledge about Human Nature*. McGraw-Hill Higher Education, New York, 4th ed edition.

Larsen, R. J. and Ketelaar, T. (1989). Extraversion, neuroticism and susceptibility to positive and negative mood induction procedures. *Personality and Individual Di erences*, 10(12):1221–1228.

Lerner, J. S. and Tetlock, P. E. (1999). Accounting for the e ects of accountability. *Psychological Bulletin*, 125(2):255–275.

Li, A., Chiu, S.-c. S., Kong, D. T., Cropanzano, R., and Ho, C.-W. (2021). How CEOs respond to mortality salience during the COVID-19 pandemic: Integrating terror management theory with regulatory focus theory. *Journal of Applied Psychology*, 106(8):1188–1201.

Li, H., Lam, H. K. S., Ho, W., and Yeung, A. C. L. (2022). The impact of chief risk o cer appointments on firm risk and operational e ciency. *Journal of Operations Management*, 68(3):241–269.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22.

Lindner, T., Puck, J., and Verbeke, A. (2020). Misconceptions about multicollinearity in international business research: Identification, consequences, and remedies. *Journal of International Business Studies*, 51(3):283–298.

Liu, A. X., Liu, Y., and Luo, T. (2016). What Drives a Firm's Choice of Product Recall Remedy? The Impact of Remedy Cost, Product Hazard, and the CEO. *Journal of Marketing*, 80(3):79–95.

Liu, Y., Shankar, V., and Yun, W. (2017). Crisis Management Strategies and the Long-Term E ects of Product Recalls on Firm Value. *Journal of Marketing*, 81(5):30–48.

Lommen, M. J., Engelhard, I. M., and van den Hout, M. A. (2010). Neuroticism and avoidance of ambiguous stimuli: Better safe than sorry? *Personality and Individual Di erences*, 49(8):1001–1006.

Loske, D. and Klumpp, M. (2021). Human-AI collaboration in route planning: An empirical e ciency-based analysis in retail logistics. *International Journal of Production Economics*, 241:108236.

Love, E. G., Lim, J., and Bednar, M. K. (2017). The Face of the Firm: The Influence of CEOs on Corporate Reputation. *Academy of Management Journal*, 60(4):1462–1481.

Lovelace, J. B., Bundy, J., Hambrick, D. C., and Pollock, T. G. (2018). The Shackles of CEO Celebrity: Sociocognitive and Behavioral Role Constraints on "Star" Leaders. *Academy of Management Review*, 43(3):419–444.

Malhotra, S., Reus, T. H., Zhu, P., and Roelofsen, E. M. (2018). The Acquisitive Nature of Extraverted CEOs. *Administrative Science Quarterly*, 63(2):370–408.

Matsumoto, D., Pronk, M., and Roelofsen, E. (2011). What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions. *The Accounting Review*, 86(4):1383–1414.

Mayo, K., Ball, G., and Mills, A. (2022). CEO Tenure and Recall Risk Management in the Consumer Products Industry. *Production and Operations Management*, 31(2):743–763.

McCrae, R. R. and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1):81–90.

McCrae, R. R. and Costa, P. T. (1997). Chapter 31 - Conceptions and Correlates of Openness to Experience. In Hogan, R., Johnson, J., and Briggs, S., editors, *Handbook of Personality Psychology*, pages 825–847. Academic Press, San Diego.

McMullen, J. S. and Kier, A. S. (2016). Trapped by the entrepreneurial mindset: Opportunity seeking and escalation of commitment in the Mount Everest disaster. *Journal of Business Venturing*, 31(6):663–686.

Medical Device Classification Procedures (2023). Title 21, Chapter I, Subchapter H, Part 860. *Code of Federal Regulations*, page 21 CFR § 860.3.

Medical Device Recall Authority (1996). 21 CFR Part 810 [Docket No. 93N–0260]. *Federal Register*, 61(225):59004–59022.

Meindl, J. R., Ehrlich, S. B., and Dukerich, J. M. (1985). The Romance of Leadership. *Administrative Science Quarterly*, 30(1):78–102.

Menon, T. and Pfe□er, J. (2003). Valuing Internal vs. External Knowledge: Explaining the Preference for Outsiders. *Management Science*, 49(4):497–513.

Mintzberg, H. (1973). *The nature of managerial work*. Harper & Row, New York, NY.

Monat, A., Averill, J. R., and Lazarus, R. S. (1972). Anticipatory stress and coping reactions under various conditions of uncertainty. *Journal of Personality and Social Psychology*, 24(2):237–253.

Mount, M. P. and Baer, M. (2022). CEOs' Regulatory Focus and Risk-Taking When Firms Perform Below and Above the Bar. *Journal of Management*, 48(7):1980–2008.

Mukherjee, U. K., Ball, G. P., Wowak, K. D., Natarajan, K. V., and Miller, J. W. (2022). Hiding in the Herd: The Product Recall Clustering Phenomenon. *Manufacturing & Service Operations Management*, 24(1):392–410.

Mukherjee, U. K. and Sinha, K. K. (2018). Product Recall Decisions in Medical Device Supply Chains: A Big Data Analytic Approach to Evaluating Judgment Bias. *Production and Operations Management*, 27(10):1816–1833.

Muralidharan, E., Bapuji, H., and Hora, M. (2019). The more I err, the less I pay: E□ect of firm recall experience, firm type and recall severity on remedy to consumers. *European Journal of Marketing*, 53(5):916–943.

Nadkarni, S. and Herrmann, P. (2010). CEO Personality, Strategic Flexibility, and Firm Performance: The Case of the Indian Business Process Outsourcing Industry. *Academy of Management Journal*, 53(5):1050–1073.

Neely, B. H., Lovelace, J. B., Cowen, A. P., and Hiller, N. J. (2020). Metacritiques of Upper Echelons Theory: Verdicts and Recommendations for Future Research. *Journal of Management*, 46(6):1029–1062.

Nettle, D. (2006). The evolution of personality variation in humans and other animals. *American Psychologist*, 61(6):622–631.

Nettle, D. and Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *European Journal of Personality*, 22(4):323–335.

Ni, J. and Huang, X. (2018). Discovery-to-Recall in the Automotive Industry: A

Problem-Solving Perspective on Investigation of Quality Failures. *Journal of Supply Chain Management*, 54(2):71–95.

Ni, J. Z., Flynn, B. B., and Jacobs, F. R. (2014). Impact of product recall announcements on retailers' financial value. *International Journal of Production Economics*, 153:309–322.

Ormiston, M. E., Wong, E. M., and Ha, J. (2022). The role of CEO emotional stability and team heterogeneity in shaping the top management team a□ective tone and firm performance relationship. *The Leadership Quarterly*, 33(3):101543.

Park, J. and Woo, S. E. (2022). Who Likes Artificial Intelligence? Personality Predictors of Attitudes toward Artificial Intelligence. *The Journal of Psychology*, 156(1):68–94.

Patel, B. N., Rosenberg, L., Willcox, G., Baltaxe, D., Lyons, M., Irvin, J., Rajpurkar, P., Amrhein, T., Gupta, R., Halabi, S., Langlotz, C., Lo, E., Mammarappallil, J., Mariano, A. J., Riley, G., Seekins, J., Shen, L., Zucker, E., and Lungren, M. P. (2019). Human–machine partnership with artificial intelligence for chest radiograph diagnosis. *npj Digital Medicine*, 2(1):1–10.

Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). *Linguistic inquiry and word count*. Lawerence Erlbaum Associates, Mahwah, NJ.

Peterson, R. S., Smith, D. B., Martorana, P. V., and Owens, P. D. (2003). The impact of chief executive o□cer personality on top management team dynamics: One mechanism by which leadership a□ects organizational performance. *Journal of Applied Psychology*, 88(5):795–808.

Pfarrer, M. D., Pollock, T. G., and Rindova, V. P. (2010). A Tale of Two Assets: The E□ects of Firm Reputation and Celebrity on Earnings Surprises and Investors' Reactions. *Academy of Management Journal*, 53(5):1131–1152.

Pollock, T. G., Ragozzino, R., and P. Blevins, D. (2023). Not Like the Rest of Us? How CEO Celebrity A□ects Quarterly Earnings Call Language. *Journal of Management*, page 014920632211506.

Priem, R. L., Lyon, D. W., and Dess, G. G. (2016). Inherent Limitations of Demo-

graphic Proxies in Top Management Team Heterogeneity Research:. *Journal of Management*, 25(6):935–953.

Qian, C., Balaji, P., Crilly, D., and Liu, Y. (2023). Better Safe Than Sorry: CEO Regulatory Focus and Workplace Safety. *Journal of Management*, Articles in Advance:1–35.

Quigley, T. J. and Hambrick, D. C. (2015). Has the "CEO e□ect" increased in recent decades? A new explanation for the great rise in America's attention to corporate leaders. *Strategic Management Journal*, 36(6):821–830.

Quigley, T. J., Hubbard, T. D., Ward, A., and Gra□n, S. D. (2020). Unintended Consequences: Information Releases and CEO Stock Option Grants. *Academy of Management Journal*, 63(1):155–180.

Raithel, S. and Hock, S. J. (2021). The crisis-response match: An empirical investigation. *Strategic Management Journal*, 42(1):170–184.

Rauthmann, J. F., Sherman, R. A., Nave, C. S., and Funder, D. C. (2015). Personality-driven situation experience, contact, and construal: How people's personality traits predict characteristics of their situations in daily life. *Journal of Research in Personality*, 55:98–111.

Rhee, M., Kim, Y.-C., and Han, J. (2006). Confidence in Imitation: Niche-Width Strategy in the UK Automobile Industry. *Management Science*, 52(4):501–513.

Roberts, B. W. and DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1):3–25.

Roh, J., Krause, R., and Swink, M. (2016). The appointment of chief supply chain o□cers to top management teams: A contingency model of firm-level antecedents and consequences. *Journal of Operations Management*, 44(1):48–61.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3):211–252.

Rusting, C. L. (1998). Personality, mood, and cognitive processing of emotional information: Three conceptual frameworks. *Psychological Bulletin*, 124(2):165–196.

Salvador, F., Alba, C., Madiedo, J. P., Tenhiälä, A., and Bendoly, E. (2021). Project managers' breadth of experience, project complexity, and project performance. *Journal of Operations Management*, 67(6):729–754.

Samson, D. and Kalchschmidt, M. (2019). Looking forward in operations management research. *Operations Management Research*, 12(1):1–3.

Schaedler, L., Graf-Vlachy, L., and König, A. (2022). Strategic leadership in organizational crises: A review and research agenda. *Long Range Planning*, 55(2):102156.

Schaefer, P. S., Williams, C. C., Goodie, A. S., and Campbell, W. K. (2004). Overconfidence and the Big Five. *Journal of Research in Personality*, 38(5):473–480.

Schneider, B. and Smith, D. B., editors (2004). *Personality and organizations*. LEA's organization and management series. Erlbaum, Mahwah, NJ.

Scoresby, R. B., Withers, M. C., and Ireland, R. D. (2021). The e□ect of CEO regulatory focus on changes to investments in R&D. *Journal of Product Innovation Management*, 38(4):401–420.

Shah, R., Ball, G. P., and Netessine, S. (2017). Plant Operations and Product Recalls in the Automotive Industry: An Empirical Investigation. *Management Science*, 63(8):2439–2459.

Shiraev, E. (2016). *Personality Theories: A Global View*. SAGE Publications, Thousand Oaks, CA.

Shoss, M. K. and Witt, L. A. (2014). Trait Interactions and Other Configural Approaches to Personality. In Christiansen, N. and Tett, R., editors, *Handbook of Personality at Work*. Routledge, New York, 1st edition.

Siebert, S. E. and DeGeest, D. S. (2015). The Five Factor Model of Personality in Business and Industry. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*. Oxford University Press.

Singh, I. R. (2007). *Factors Influencing the Time for FDA Review of Medical Devices*. Thesis, Massachusetts Institute of Technology.

Smith, A. D., Plowman, D. A., Duchon, D., and Quinn, A. M. (2009). A qualitative study of high-reputation plant managers: Political skill and successful outcomes. *Journal of Operations Management*, 27(6):428–443.

Smith, N. C., Thomas, R. J., and Quelch, J. A. (1996). A Strategic Approach to Managing Product Recalls. *Harvard Business Review*, 74(5):102–113.

S&P Global Market Intelligence (2018). Global Industry Classification Standard (GICS). Technical report, tandard & Poor's Financial Services LLC (S&P), MSCI.

Spanjol, J., Tam, L., Qualls, W. J., and Bohlmann, J. D. (2011). New Product Team Decision Making: Regulatory Focus E□ects on Number, Type, and Timing Decisions. *Journal of Product Innovation Management*, 28(5):623–640.

Steven, A. B., Dong, Y., and Corsi, T. (2014). Global sourcing and quality recalls: An empirical study of outsourcing-supplier concentration-product recalls linkages. *Journal of Operations Management*, 32(5):241–253.

Steyvers, M., Tejeda, H., Kerrigan, G., and Smyth, P. (2022). Bayesian modeling of human–AI complementarity. *Proceedings of the National Academy of Sciences*, 119(11):e2111547119.

Strohhecker, J. and Größler, A. (2013). Do personal traits influence inventory management performance?—The case of intelligence, personality, interest and knowledge. *International Journal of Production Economics*, 142(1):37–50.

Sutin, A. R. (2015). Openness. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*. Oxford University Press.

Sutin, A. R., McCrae, R. R., and Costa, P. T. (2011). *The Neuroscience of Personality Traits: Descriptions and Prescriptions*. Oxford University Press.

Szegedy, C., Vanhoucke, V., Io□e, S., Shlens, J., and Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, Las Vegas, NV. IEEE.

Tackett, J. L. and Lahey, B. B. (2016). Neuroticism. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*. Oxford University Press.

Tamir, M. (2005). Don't worry, be happy? Neuroticism, trait-consistent a□ect regulation, and performance. *Journal of Personality and Social Psychology*, 89(3):449–461.

Tang, P. M., Koopman, J., McClean, S. T., Zhang, J. H., Li, C. H., De Cremer, D., Lu, Y., and Ng, C. T. S. (2021). When Conscientious Employees Meet Intelligent Machines: An Integrative Approach Inspired by Complementarity Theory and Role Theory. *Academy of Management Journal*, 65(3).

Taudien, A., Fügener, A., Gupta, A., and Ketter, W. (2022). Calibrating Users' Mental Models for Delegation to AI. In *ICIS 2022 Proceedings. 16.*

Tett, R. P. and Guterman, H. A. (2000). Situation Trait Relevance, Trait Expression, and Cross-Situational Consistency: Testing a Principle of Trait Activation. *Journal of Research in Personality*, 34(4):397–423.

Thirumalai, S. and Sinha, K. K. (2011). Product Recalls in the Medical Device Industry: An Empirical Exploration of the Sources and Financial Consequences. *Management Science*, 57(2):376–392.

Thompson, E. R. (2008). Development and Validation of an International English Big-Five Mini-Markers. *Personality and Individual Di erence*s, 45(6):542–548.

Tomsett, R., Preece, A., Braines, D., Cerutti, F., Chakraborty, S., Srivastava, M., Pearson, G., and Kaplan, L. (2020). Rapid Trust Calibration through Interpretable and Uncertainty-Aware AI. *Patterns*, 1(4):100049.

Tuncdogan, A., Boon, A., Mom, T., Van Den Bosch, F., and Volberda, H. (2017). Management teams' regulatory foci and organizational units' exploratory innovation: The mediating role of coordination mechanisms. *Long Range Planning*, 50(5):621–635.

van der Linden, D., te Nijenhuis, J., and Bakker, A. B. (2010). The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, 44(3):315–327.

Van Heerde, H., Helsen, K., and Dekimpe, M. G. (2007). The Impact of a Product-Harm Crisis on Marketing E□ectiveness. *Marketing Science*, 26(2):230–245.

Vergne, J., Wernicke, G., and Brenner, S. (2018). Signal Incongruence and Its Conse-

quences: A Study of Media Disapproval and CEO Overcompensation. *Organization Science*, 29(5):796–817.

Wang, X. and Yin, M. (2021). Are Explanations Helpful? A Comparative Study of the E□ects of Explanations in AI-Assisted Decision-Making. In *26th International Conference on Intelligent User Interfaces*, pages 318–328, College Station TX USA. ACM.

Wernicke, G., Sajko, M., and Boone, C. (2022). How much Influence do CEOs have on Company Actions and Outcomes? The Example of Corporate Social Responsibility. *Academy of Management Discoveries*, 8:36–55.

Widiger, T. A. (2009). Neuroticism. In Leary, M. R. and Hoyle, R. H., editors, *Handbook of individual di erences in social behavior*, pages 129–146. Guilford Press, New York, NY.

Wiesenfeld, B. M., Wurthmann, K. A., and Hambrick, D. C. (2008). The Stigmatization and Devaluation of Elites Associated with Corporate Failures: A Process Model. *Academy of Management Review*, 33(1):231–251.

Wilt, J. and Revelle, W. (2016). Extraversion. In Widiger, T. A., editor, *The Oxford Handbook of the Five Factor Model*, volume 1. Oxford University Press.

Wooldridge, J. (2002). *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, MA.

Wooldridge, J. M. (2013). *Introductory econometrics: a modern approach*. South-Western Cengage Learning, Mason, OH, 5th edition.

Wowak, A. J., Gomez-Mejia, L. R., and Steinbach, A. L. (2017). Inducements and Motives at the Top: A Holistic Perspective on the Drivers of Executive Behavior. *Academy of Management Annals*, 11(2):669–702.

Wowak, A. J., Mannor, M. J., and Wowak, K. D. (2015). Throwing caution to the wind: The e□ect of CEO stock option pay on the incidence of product safety problems: CEO Stock Option Pay and Product Safety Problems. *Strategic Management Journal*, 36(7):1082–1092.

Wowak, K. D., Ball, G. P., Post, C., and Ketchen, D. J. (2021). The Influence of Female

Directors on Product Recall Decisions. *Manufacturing & Service Operations Management*, 23(4):895–913.

Wowak, K. D. and Boone, C. A. (2015). So Many Recalls, So Little Research: A Review of the Literature and Road map for Future Research. *Journal of Supply Chain Management*, 51(4):54–72.

Yoon, Y., Sarial-Abi, G., and Gürhan-Canli, Z. (2012). E□ect of Regulatory Focus on Selective Information Processing. *Journal of Consumer Research*, 39(1):93–110.

Zhang, Y. (2006). The presence of a separate COO/president and its impact on strategic change and CEO dismissal. *Strategic Management Journal*, 27(3):283–300.

Zhang, Y., Liao, Q. V., and Bellamy, R. K. E. (2020). E□ect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 295–305, Barcelona, Spain. ACM.