

OUTCOMES OF INTERPERSONAL FELT DISTRUST IN THE WORKPLACE:
AN IDENTITY THREAT PERSPECTIVE

IRYNA KALYNYCHENKO

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN
HUMAN RESOURCE MANAGEMENT
YORK UNIVERSITY
TORONTO, ONTARIO

August 2025

© Iryna Kalynychenko, 2025

Abstract

While there has been an undeniable upsurge in the study of distrust in organizations, surprisingly, there remains a dearth of research on the outcomes of felt distrust in the workplace. More specifically, little attention has been paid to examining how individuals respond to feeling distrusted by others at work. This dissertation aims to address this gap by developing and empirically examining a model of the emotional and behavioral impact of felt distrust by a supervisor on their subordinates. Drawing on identity theory, I suggest that felt distrust triggers a threat to one's work-related identity, which is more pronounced for individuals high in moral identity centrality. Further, interpersonal felt distrust evokes negative moral emotions -- anger and shame -- followed by engagement in counterproductive organizational behaviors, psychological withdrawal and impression management behavior.

I empirically test this conceptual model through an experiment in Study 1 and a survey following a critical incident technique in Study 2. Study 1 tested the causal effect of interpersonal felt distrust by a supervisor on an individual's work-related identity threat and subsequent emotional reactions, and the moderating role of moral identity centrality. Study 2 replicated Study 1 and also tested behavioral reactions to interpersonal felt distrust.

The findings of this dissertation show that work-related identity threat acts as a core psychological mechanism that explains individual reactions to feeling distrusted by a supervisor at work. This threat becomes even more salient for individuals who place a high value on their moral identity. These results are supported in both studies. In Study 2, the hypotheses linking emotional and behavioral outcomes were not supported, and the analysis of the overall structural equation model reveals some interesting and unexpected findings.

The theoretical contributions of this novel perspective enrich the organizational literature on distrust, work-related identity threat and emotions in the workplace. From a practical perspective, a better understanding of how and why individuals react to feeling distrusted underscores the importance for leaders and HR practitioners to be mindful of the communication of distrust to employees given its profound impact.

Dedication

To my husband Alyosha (Oleksiy), for teaching me the true value of trust.

Acknowledgements

Behind any monumental work is a team that made it possible; oftentimes, despite the circumstances around. This dissertation is the result of seven years of my dedicated work towards attaining the doctoral degree. And it is in this section and at this time, when this work is completed and dissertation is defended, that I want to acknowledge my team – individuals who guided, mentored and supported me along the way, and my soul companions who quietly and patiently comforted me through occasional struggles and confusion. Without their enduring support and active presence, I would not have reached the level at which I am now, as a scholar, a learner, and a person.

First and foremost, I would like to thank my dissertation supervisor, Professor Sabrina Deutsch Salamon, for lighting the spark of my interest in academic research since our very first doctoral seminar, mentoring me all along during my doctoral program, making herself available to support me through my dissertation journey, setting high expectations and gently guiding me on my way to achieving them. Throughout my time working on the dissertation, Dr. Deutsch Salamon has been providing me with resources and opportunities to succeed, sharing her expertise and working through challenging aspects together in the process of idea development, research and writing, and teaching me about balance of work-life, self-compassion and trusting myself more in the process. These experiences and skills have really laid the foundation for my positive identity as a researcher, a professor and a faculty member.

I would also like to thank my dissertation committee members, Professors Parbudyal Singh and Mark Podolsky, for their time, expertise and feedback on the many drafts of my dissertation proposal, methodological designs of the studies, data analysis and writing of the final manuscript. Their input has been very helpful in improving different aspects of my dissertation. I

would like to thank Professor James Chowhan, my mentor and co-author, who created opportunities for me to advance my research methods skills, cheered on my progress through the doctoral program and advised me on research methods and data analysis when I stumbled. The completion of this dissertation would also not be possible without due attention and support from Hugh McCague and Michael Truong from the Statistical Consulting Service at York University. They helped me think through my research designs, get to know my data, make informed decisions, and broadened my understanding of approaches to data collection and analysis. Thanks to all these academic professionals, and to many of my fellow doctoral students at the School of HRM, I have pulled my thoughts and resources together and delivered.

This dissertation research and my doctoral program experience have also been made possible through the everlasting support of my family, close and far. I want to say special thank you to my husband Alyosha (Oleksiy) for his faith in my success, seeing the importance of my work both for science and personal development, standing by when things got tough, creating opportunities for me to rewind and relax, and taking care of our family. Thank you to our daughter Daryna for learning to understand and appreciate my work as a student and now as a faculty member and reminding me about the importance of spending time with family. And the three soul companions that helped me connect with nature and myself daily – to our dog Umka, cats Sonya and Elaine, thank you!

My progress on this dissertation is also very much due to the support and faith of our family in Ukraine. They all found time to learn about my newest achievements and struggles, and cheered for me every day, despite the horrific realities of war they live through on a daily basis. Special thanks to my dearest parents, Zinaida and Oleksandr Sydorenko and my brother Oleksiy Sydorenko, for their support of my studies, interest in my dissertation research, suggestions on

how to move forward during difficult times, and their vision of my success. I also want to thank my husband's side of the family, my mother- and father-in-law Larysa and Vladimir Kalynychenko, sisters-in-law Elina and Vasilisa, and Vasilisa's family. They have always shown their strong faith and confidence in my ability to achieve my goals and supported our family with cheer and laughter as we braced ourselves through interesting but difficult time of our integration in Canadian society and my pursuit of academic career. To all of them, I am deeply thankful and recognize that this work is possible because they have all been a part of my life.

Table of Contents

Abstract.....	ii
Dedication.....	iv
Acknowledgements.....	v
Table of contents.....	viii
List of tables.....	xi
List of figures.....	xiii
 Chapter 1 – Introduction	 1
1.1 Overview.....	1
1.2 Research contributions.....	5
 Chapter 2 – Literature review.....	 6
2.1 Distrust as a unique and separate concept from trust.....	6
2.2 Existing research on outcomes of felt distrust.....	10
2.2.1 References to felt distrust in extant research.....	11
2.2.2 Empirical research on impersonal felt distrust.....	11
2.2.3 Research on interpersonal felt distrust in organizational setting.....	14
2.2.4 Existing research on the outcomes of felt trust.....	15
2.2.5 Summary of research on the outcomes of felt distrust	17
2.3 Work-related identity.....	19
2.4 Moral identity centrality.....	21
2.5 Work-related identity threat.....	24
 Chapter 3 – Theory and hypotheses.....	 26
3.1 Felt distrust as a trigger of work-related identity threat.....	27
3.2 Emotional reactions to a threatened work-related identity.....	29
3.3 Behaviors following emotional reactions to threatened work-related identity.....	33
3.4 Conclusion.....	38
 Chapter 4 – Methods and results.....	 40
4.1 Overarching approach to research design.....	40
4.2 Study 1 description	41
4.2.1 Sample and procedure.....	42
4.2.1.1 Sample size determination.....	42
4.2.1.2 Sample characteristics.....	42
4.2.1.3 Survey design.....	43
4.2.1.4 Survey procedure.....	44
4.2.1.5 Testing the survey	45
4.2.1.6 Launching the survey	45
4.2.2 Ensuring data quality.....	46
4.2.3 Measures.....	48
4.2.3.1 Dependent variables.....	48
4.2.3.2 Mediating variable.....	49
4.2.3.3 Moderating variable.....	50
4.2.3.4 Independent variable.....	50

4.2.3.5 Control variables.....	51
4.3 Study 1 data analysis strategy.....	54
4.4 Study 1 results.....	55
4.4.1 Assessing data quality.....	55
4.4.2 Assessing data normality.....	56
4.4.2.1 Testing for univariate normality.....	56
4.4.2.2 Testing for multivariate normality.....	56
4.4.3 Multivariate outlier detection and management.....	57
4.4.4 Descriptives.....	58
4.4.5 Validity and reliability of measures.....	61
4.4.6 Correlational analysis.....	64
4.4.7 Group comparison analyses.....	66
4.4.8 Structural Equation Modeling (SEM) analysis.....	69
4.4.8.1 Assessing the measurement model.....	70
4.4.8.2 Testing for common method variance.....	71
4.4.8.3 Double-mean centering the latent interaction (moderation).....	71
4.4.8.4 Testing multicollinearity.....	72
4.4.8.5 Hypotheses testing.....	73
4.4.8.6 Simple slopes analysis.....	77
4.4.8.7 Mediation analysis.....	79
4.4.8.8 Conclusion.....	80
4.4.9 Complementary analysis: Multiple Linear Regression.....	81
4.4.9.1 Hypothesis 1.....	81
4.4.9.2 Hypothesis 2.....	83
4.4.9.3 Hypothesis 3a and 3b.....	87
4.4.9.4 Hypothesis 4a and 4b.....	92
4.5 Study 1 discussion and conclusion.....	97
Chapter 5 – Study 2 methods and results.....	100
5.1 Study 2 description	100
5.1.1 Sample and procedure.....	101
5.1.1.1 Sample size determination.....	101
5.1.1.2 Sample characteristics.....	102
5.1.1.3 Survey design.....	104
5.1.1.4 Survey procedure.....	105
5.1.2 Ensuring data quality.....	106
5.1.3 Measures.....	107
5.1.3.1 Dependent variables.....	108
5.1.3.2 Mediating variable.....	109
5.1.3.3 Moderating variable.....	110
5.1.3.4 Independent variable.....	110
5.1.3.5 Control variables.....	110
5.2 Study 2 data analysis strategy.....	112
5.3 Study 2 results.....	112
5.3.1 Assessing data quality.....	112
5.3.2 Assessing data for normality.....	113

5.3.3 Assessing data for linearity.....	113
5.3.4 Assessing data for outliers and influential cases.....	114
5.3.5 Descriptives.....	116
5.3.6 Validity and reliability of measures.....	119
5.3.7 Correlational analysis.....	120
5.3.8 Path analysis within Structural Equation Modeling (SEM)	123
5.3.8.1 Assessing common method variance.....	124
5.3.8.2 Hypotheses testing results (original SEM model).....	125
5.3.8.3 Additional findings (original SEM model).....	129
5.3.8.4 Post-hoc analysis (original SEM model).....	130
5.3.9 SEM Sensitivity analysis.....	131
5.3.9.1 Accounting for range restriction.....	131
5.3.9.2 Testing hypotheses results (corrected SEM model).....	132
5.3.9.3 Additional findings (corrected SEM model).....	135
5.3.9.4 Post-hoc analysis (corrected SEM model).....	137
5.4 Study 2 discussion.....	138
Chapter 6 – General discussion.....	141
6.1 Overview.....	141
6.2 Interpersonal felt distrust as a cause of work-related identity threat.....	142
6.3 The impact of interpersonal felt distrust on emotions	144
6.4 Interpersonal felt distrust by a supervisor	145
6.5 Moral identity centrality as a moderator in the relationship between interpersonal felt distrust and work-related identity threat	146
6.6 Behavioral reactions to interpersonal felt distrust	147
6.7 Practical implications.....	150
Chapter 7 – Limitations and future research.....	153
7.1 Limitations.....	153
7.2 Future research.....	159
References.....	164
Appendices.....	202
Appendix A: Study 1 experimental vignettes.....	202
Appendix B: Study 1 results of Confirmatory Factor Analyses.....	203
Appendix C: Report on hypotheses testing	205
Appendix D: Study 2 results of Confirmatory Factory Analyses.....	206

List of Tables

Table 1. Study 1. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables	59
Table 2. Study 1. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables for the Neutral Condition.....	60
Table 3. Study 1. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables for the Distrust Condition.....	61
Table 4. Study 1. Intercorrelations of Continuous Variables.....	65
Table 5. Study 1. SEM: Regression Path Estimates (Unstandardized and Standardized), Confidence Intervals for Unstandardized Coefficients, Standard Errors and P-Values.....	76
Table 6. Study 1. SEM Results: Direct, Indirect, Total and Proportional Indirect Effects in the Mediation.....	80
Table 7. Study 1. Linear Multiple Regression Predicting Identity Threat from Interpersonal Felt Distrust (Controlling for Sex, Race and Age) (Hypothesis 1).....	82
Table 8. Study 1. Linear Multiple Regression Predicting Identity Threat from Interaction of Interpersonal Felt Distrust and Moral Identity Centrality mean-centered (Hypothesis 2).....	84
Table 9. Study 1. Mediation Testing. Results of the Three Regression Analyses (Hypothesis 3).....	91
Table 10. Study 1. Causal Mediation Analysis, Interpersonal Felt Distrust Predicting Anger-Experience Through Identity Threat (Hypothesis 3b).....	92
Table 11. Study 1. Mediation Testing. Results of the Three Regression Analyses. (Hypothesis 4).....	96
Table 12. Study 1. Causal Mediation Analysis, Interpersonal Felt Distrust Predicting Shame-Experience Through Identity Threat (Hypothesis 4b).....	97
Table 13. Study 1. Comparison of Effects Between SEM and Multiple Regression Analyses.....	99
Table 14. Study 2. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables	118
Table 15. Study 2. Intercorrelations of Continuous variables.....	121

Table 16. Study 2. Comparison Table: Observed vs. Range-Restricted–Corrected Correlations.....	129
Table 17. Study 2 Hypotheses Testing Results (Original SEM model).....	127
Table 18. Study 2. SEM: Direct, Indirect, Total and Proportional Indirect Effects in the Mediation Analysis (Original SEM Model).....	130
Table 19. Study 2 Hypotheses Testing Results (Corrected SEM Model).....	134
Table 20. Study 2. SEM: Direct, Indirect, Total and Proportional Indirect Effects in the Mediation Analysis (Corrected SEM Model).....	137

List of Figures

Figure 1. Conceptual Model of the Outcomes of Interpersonal Felt Distrust in the Workplace.....	27
Figure 2. The Core of the Conceptual Model Analyzed in Study 1.....	42
Figure 3. Study 1. SEM Hypothesis Testing Results.....	77
Figure 4. Study 1. Simple Slopes Plot Depicting Conditional Effects of Scenario (Distrust vs. Neutral) by Moral Identity Centrality.....	78
Figure 5. Study 1. Predicted Work-Related Identity Threat Across MIC Levels by Scenario Condition.....	79
Figure 6. Study 1. Interaction of Scenario and Moral Identity Centrality (MIC) on Work-Related Identity Threat.....	85
Figure 7. Study 1. Effect of the Interaction Between Interpersonal Felt Distrust (Neutral vs. Distrust) and MIC (Mean-Centered) on Work-Related Identity Threat	86
Figure 8. Study 1. Confidence Intervals of the Interaction Effect (Scenario Distrust x MIC) on Identity Threat at Different Levels of MIC.....	87
Figure 9. Conceptual Model Analyzed in Study 2.....	101
Figure 10. Study 2. SEM Hypothesis Testing Results (Original SEM Model).....	127
Figure 11. Study 2. SEM Hypothesis Testing Results (Corrected SEM Model).....	133

Chapter 1: Introduction

1.1 Overview

There has been an undeniable upsurge in distrust in and among organizations and societies in the past few decades (Fukuyama, 1995; Hardin, 2004; Kramer, 2004), which remains on the rise (Gunter, 2024; Lewicki, 2021; Norquay, 2022). Scholars have noted our lack of understanding of the effects of distrust on individual behaviors and on important organizational indicators (C. P. Long & Sitkin, 2018). Surprisingly, however, little research to date has been devoted to theoretical explanations and empirical examinations of the outcomes of interpersonal felt distrust in the workplace (i.e., how individuals respond to feeling distrusted by others at work). Given today's increasingly complex and relationally fragile work environments, examining how individuals respond to interpersonal felt distrust is both timely and essential for identifying its outcomes and uncovering the mechanisms that underlie them.

While the concept of felt distrust has been mentioned across various literatures, including economics (Falk & Kosfeld, 2006; Frey, 1993; Rudolf et al., 2018), social psychology (Enzle & Anderson, 1993; Kruglanski, 1970; Mooijman et al., 2017), sociology (Kerekes et al., 2013), organization science (Belot & Schröder, 2016; Cha & Wong-On-Wing, 2020; Lanaj et al., 2018; Mooijman & Graham, 2018; Schweitzer et al., 2018), accounting (Christ, 2013), and public administration (Lee & Dodge, 2019; Mendoza & Wielhouwer, 2024; Weinberg, 2023), only a handful of studies to date have been specifically designed to empirically examine the impact of felt distrust (Cha & Wong-On-Wing, 2020; Lanaj et al., 2018; Liu et al., 2024; Mendoza & Wielhouwer, 2024; Mooijman et al., 2017; Rudolf et al., 2018; Weinberg, 2023). This existing research, for the most part, refers to responses to the impersonal communication of distrust, which

refers to distrust signaled by entities such as organizations through formal practices, controls, or surveillance systems, rather than to responses to interpersonal distrust (specifically, Cha & Wong-On-Wing, 2020; Mendoza & Wielhouwer, 2024; Mooijman et al., 2017; Rudolf et al., 2018; Weinberg, 2023). The only two studies that examined the effects of interpersonal felt distrust have been limited to felt distrust by colleagues (Lanaj et al., 2018) and feeling ability-distrusted by supervisors (Liu et al., 2024). As a result, a critical void remains in our understanding of the impact of interpersonal felt distrust in organizations—that is, the response to distrust that is perceived to be conveyed by another person. Specifically, we do not know how employees react to feeling distrusted by their supervisors in a relationship wherein the source of the distrust comes from a person who holds a position of authority over employees and therefore wields power and influence over their standings and careers in the organization.

This dissertation aims to address this gap in the literature by empirically examining individual emotional and behavioral reactions to feeling distrusted by a supervisor. Importantly, it builds on the views of distrust as a separate and unique concept from trust (Bijlsma-Frankema et al., 2015; Casado-Aranda et al., 2019; J. Cho, 2006; Dimoka, 2010; Korsgaard, 2018; Lee & Lee, 2024; Lewicki et al., 1998; Reimann et al., 2017; Saunders et al., 2014; Saunders & Thornhill, 2004; Sitkin & Roth, 1993), and felt distrust as distinct from low felt trust. Recent research in organizational science finds that leaders often distrust others (Inesi et al., 2012; Mooijman et al., 2015, 2017, 2019; Mooijman & Graham, 2018). Specifically, power motivation, imbalance, and instability in social hierarchies create room for interpersonal distrust because individuals in power (e.g., managers, supervisors, leaders) face concerns about loss of power or position as well as potential loss of valuable resources (Inesi et al., 2012; Mooijman et al., 2019). Power motivation alone often predisposes leaders to be distrustful of others' underlying intentions and to view them

as untrustworthy, exploitative, and corrupt (Mooijman et al., 2015). Because power motivation and concern over power loss often drive leaders to be distrustful of others, they may engage in behaviors that signal distrust to their subordinates. This dissertation aims to enhance our understanding of how and why employees may feel and be inclined to behave in the workplace as a response to feeling distrusted by their supervisors.

Specifically, in this dissertation, I aim to answer the following research questions: *How do employees react to feeling distrusted by supervisors, and why? What factors affect their response?* To address these questions, I offer a novel identity-focused perspective to explain the emotional and behavioral outcomes of interpersonal felt distrust. Because the workplace is a salient source for the development and maintenance of individuals' identities (Daniels & Robinson, 2019) and because people are motivated to maintain positive work-related identities (Dutton, Roberts, & Bednar, 2010), I argue that being perceived as possessing intentions to inflict harm on another person at work threatens one's self-conception as a moral being. By illuminating how felt distrust may threaten core aspects of work-related identity, this dissertation contributes to a deeper understanding of the psychological consequences of distrust in organizations and offers new insight into the relational dynamics that shape supervisor-employee relations and employee emotional and behavioral outcomes at work.

I draw on identity theory (Stryker & Burke, 2000) and established social psychological perspectives on identity (Breakwell, 1986; Erez & Early, 1993; McCall & Simmons, 1978) to suggest that felt distrust triggers an identity threat (Aquino & Douglas, 2003; Breakwell, 1986; Petriglieri, 2011) to one's work-related identity (Miscenko & Day, 2016; Selenko et al., 2018), and that this threat will be more pronounced for individuals who hold their moral identity (Hart, 2005) as central to their self-conception (Aquino et al., 2009; Aquino & Reed, 2002). I further incorporate

the literature on moral emotions (de Hooge et al., 2010; Douglas et al., 2008) to suggest that by threatening one's work-related identity, interpersonal felt distrust evokes negative moral emotions (e.g., anger and shame), followed by behavioral outcomes such as counterproductive organizational behaviors against the organization, psychologically withdrawing from the workplace, and engaging in prosocial impression management behavior.

I empirically test this conceptual model by undertaking a multimethod research design (Wellman et al., 2023) with the aim of providing methodological triangulation for hypothesis testing and thereby strengthening the validity of the research results. To achieve this aim, I combine two complementary deductive methods, namely an experiment in Study 1 and a survey following a critical incident technique (Flanagan, 1954) in Study 2. Both studies are set up on Qualtrics, utilize convenience non-representative samples and draw respondents from an online panel of employees on Prolific. Study 1 sample is of the working population of the United States. Study 2 sample, represents a global working pool and consists of employees who felt distrusted by their supervisor at work. Study 1 enhances internal validity by experimentally manipulating interpersonal felt distrust and randomly assigning participants to two conditions. Study 2, in turn, contributes to enhanced external validity by testing the same core model, along with behavioral outcomes, in a different context, i.e., employees with real experiences of felt distrust. It utilizes a well-established methodological approach, a critical incident technique combined with a survey design (Hershcovis, 2011; O'Driscoll & Cooper, 1994), to effectively capture individual reactions to discrete workplace events. The multimethod research approach (Wellman et al., 2023) to examining the research questions in this dissertation aims to enhance the rigor and validity of the research findings.

1.2 Research Contributions

Understanding the consequences of interpersonal felt distrust in the workplace is essential both for theory and practice. This dissertation provides several important theoretical contributions to the organizational behavior and human resource management literature. First, it examines felt distrust as a phenomenon that arises within the leader-subordinate dyad – a relational context that significantly impacts one’s identity. In this way, it extends prior research that has recognized the unique role of distrust in the workplace (Bijlsma-Frankema et al., 2015; Lewicki et al., 1998; Sitkin & Bijlsma-Frankema, 2018) by examining what it is like to be on the receiving end of distrust. Second, it directly responds to Long & Sitkin's (2018) call for more research on the effects on individual employees of feeling trusted or distrusted at work. Third, it addresses the calls for research that examines the effects of subordinate felt distrust by a leader on subordinates’ emotions and behavior at work (Mooijman & Graham, 2018), and more general calls for research on the relationship between emotions and trust or distrust in the organizational context (Bies et al., 2018; Schoorman et al., 2007; Searle et al., 2018).

As a practical contribution, this research provides crucial insights to organizational leaders and human resources professionals on the impact of interpersonal felt distrust in the workplace. In order to built trust in the organization, distrust has to be eradicated first (Bijlsma-Frankema et al., 2015). Being able to recognize employees feeling distrusted at work and address the issue are important prerequisites to establishing trustful and positive work environment. Moreover, it is possible that employees often experience felt distrust as an unanticipated consequence of poor managerial decision-making and communication processes (Litzky et al., 2006). Therefore, supervisors need to be trained not to signal distrust to employees inadvertently.

Chapter 2: Literature Review

The research on the outcomes of felt distrust is sparse and spread across diverse literatures. Most of this research mentions distrust that may be implicitly conveyed through particular organizational practices, such as control or surveillance. As such, felt distrust is mostly impersonal (general and systemic, rather than interpersonal). Importantly, in most of this research, felt distrust is assumed by researchers rather than examined empirically.

This chapter provides a review of the literature underpinning this dissertation, specifically drawing on distrust, work-related identity, moral identity centrality, and identity threat. It begins with a section on distrust as a unique and separate construct from trust and proceeds to an overview of the research on the outcomes of felt distrust. The analysis of the felt distrust literature covers what is known to date, and why the existing research has been limited in furthering knowledge around the emotional and behavioral outcomes of interpersonal felt distrust in the workplace. Building on the scholarly understanding that high distrust is not the same as low trust, this section also briefly covers existing research on outcomes associated with felt trust and provides arguments as to why this body of knowledge is limited in informing our understanding of the outcomes of interpersonal felt distrust. In the remainder of this chapter, I review the literatures relevant to work-related identity, moral identity centrality, and identity threat.

2.1 Distrust as a unique and separate concept from trust

Although both trust and distrust serve to reduce the complexity and uncertainty engrained in human relationships (Luhmann, 1979), the period spanning the late 20th century and first quarter of the 21st century has been marked by the unprecedented growth of distrust in societies and organizations across the globe (Grace, 2022; Gunter, 2024; Lewicki, 2021; Norquay, 2022;

Zuckerman, 2021). This upsurge of distrust has triggered increased attention on the part of organizational researchers as to the nature, antecedents, and consequences of distrust in the workplace. Scholars have documented the unique features of distrust, not merely as the opposite of trust, but as a distinctively different phenomenon from trust in its qualities and emotional load, antecedents and outcomes, and the managerial practices that are effective when it comes to harnessing distrust in the workplace (Guo et al., 2017; Hardin, 2004; Korsgaard, 2018; Lewicki et al., 1998; Lumineau, 2017; McKnight & Chervany, 2001a, 2001b; Sitkin & Bijlsma-Frankema, 2018; Sitkin & Stickel, 1996). The fact that the growing organizational presence of distrust is documented across diverse social institutions in multiple literatures, including organizational science (Guo et al., 2017; Korsgaard, 2018; Kramer, 1994, 1995; Lewicki et al., 1998), political science (Gunter, 2024; Hardin, 2004; Norquay, 2022), healthcare (Alpers, 2018; Armstrong et al., 2006; Webb Hooper et al., 2022), and policing (Andreescu & Keeling, 2012; Sharp & Johnson, 2009) implies that individuals employed in organizations adjacent to these fields are regular recipients of distrust, making it crucial to understand how feeling distrusted at work affects them.

Much of the recent scholarship on distrust and trust positions the two constructs as distinct from one another (Bies et al., 2018; Bijlsma-Frankema et al., 2015; Capiola et al., 2022; Connelly et al., 2012; D'Cruz, 2020; Guo et al., 2017; Hardin, 2004; Huang & Dastmalchian, 2006; Isaeva et al., 2019; Jason Huang & Dastmalchian, 2006; Kostis & Näsholm, 2020; Kramer, 1999; Kramer & Cook, 2004; Kujala et al., 2016; Levi, 2000; Lewicki, 2007; Lewicki et al., 1998, 2016; Lumineau, 2017; Marineau, 2017; McKnight & Chervany, 2001a, 2001b; McKnight & Choudhury, 2006; Raza-Ullah, 2019; Reimann et al., 2017; Saunders et al., 2014; Siebert & Czarniawska, 2020; Sitkin & Bijlsma-Frankema, 2018; Sitkin & Roth, 1993). Trust rests on the willingness of the principal to be vulnerable to the actions of another based on their *positive*

expectations of another party's integrity, benevolence, and ability (Mayer et al., 1995). Distrust, on the other hand, is based on *negative* expectations of another's conduct; more specifically, distrust is "an actor's assured expectation of intended harm from the other" (Lewicki et al., 1998, p. 446). It involves "a fear of, a propensity to attribute sinister intentions to, and a desire to buffer oneself from the effects of another's conduct" (Lewicki et al., 1998, p. 439).

Distrust is expressed through wariness, skepticism, and behaviors like defensiveness, watchfulness, and vigilance. It is also manifested through social constraints that represent practical responses by wary and vigilant dependent parties to perceived threats (e.g., monitoring mechanisms or bureaucratic and regulatory controls) (Lewicki et al., 1998, p. 446). Unlike trust, which motivates a type of approach behavior (i.e., a willingness to engage with the other party and take risks), distrust triggers avoidant behaviors (Korsgaard, 2018) in one's attempt to protect oneself from the harmful conduct of others (Lewicki et al., 1998). Unlike trust, which is domain-specific (p. 712), distrust is more generalized and obviates the existence of trust (Bijlsma-Frankema et al., 2015). Finally, distrust is not a mere opposite of trust in that "low trust is not the same as high distrust" (Lewicki et al., 1998). Consequently, felt trust—which refers to the extent to which one believes that another trusts them (Salamon & Robinson, 2008)—is distinct from felt distrust, which refers to the extent to which one believes another distrusts them. Therefore, the research on the theory and outcomes of felt trust (Baer et al., 2021; Campagna et al., 2019; Lau et al., 2014; Lau & Lam, 2008; Salamon & Robinson, 2008) provides little insight into how and why individuals respond to felt distrust at work.

Many studies provide empirical evidence concerning the unique nature of distrust as separate from trust (i.e., Bijlsma-Frankema et al., 2015; Casado-Aranda et al., 2019; Cho, 2006; Dimoka, 2010; Lee & Lee, 2024; Reimann et al., 2017; Saunders et al., 2014; Saunders &

Thornhill, 2004; Sitkin & Roth, 1993). Specifically, several neuroimaging studies (Casado-Aranda et al., 2019; Dimoka, 2010) reveal that different parts of the brain are activated when an individual distrusts compared to when they trust. A study of monozygotic and dizygotic female twins found that trust is explained by heritability while distrust, in contrast, is explained by shared socialization (Reimann et al., 2017). In an organizational setting, scholars have discovered that different managerial practices are needed to reduce distrust than those used to build trust (Saunders et al., 2014; Sitkin & Roth, 1993). Other organizational researchers have examined the self-perpetrating and trust-obviating nature of distrust in that until distrust is eradicated, there can be no room for trust to emerge (Bijlsma-Frankema et al., 2015). These findings all point to the drastic distinction between distrust and trust as human experiences and scientific phenomena. Following the lead from the organizational research domain, several related fields of science have adopted the view that distrust is a separate and distinct concept from trust and have further developed this thinking and its application in, for instance, e-commerce and consumer research (Benamati et al., 2006; Chang & Fang, 2013; J. Cho, 2006; McKnight et al., 2004; McKnight & Choudhury, 2006; Moody et al., 2014; Ou & Sia, 2010), information technology and AI research (Baba, 1999; Komiak & Benbasat, 2008; McKnight et al., 2017; Schelble et al., 2022), healthcare research (Alpers, 2018; Greenidge, 2009), and public policy and administration research (Devine et al., 2020; Harding et al., 2016; Lee, 2021; Lee & Lee, 2024; Sharp & Johnson, 2009; Six & Latusek, 2023; Van De Walle & Six, 2014; Weinberg, 2023).

While previous research has referred to the concept of felt distrust, a formal definition has not been suggested or widely adopted. Based on the notion of distrust as “a fear of, a propensity to attribute sinister intentions to, and a desire to buffer oneself from the effects of another's conduct” (Lewicki et al., 1998), I define *interpersonal felt distrust* as a social cognition that another

individual thinks one has ill intentions toward them and fears one intends to harm them. In the following subsections, I review the research on the outcomes of felt distrust and identify the specific limitations inherent therein.

2.2 Existing research on the outcomes of felt distrust

Research on the outcomes of felt distrust, also referred to as perceived distrust, is still in its nascent stage. While the concept of felt distrust has been mentioned across various literatures, including economics, organization science, and public administration, only a handful of studies to date have been specifically designed to empirically examine the impact of felt distrust. Overwhelmingly, felt distrust has been alluded to as a reason for retaliation in specific contexts, such as monitoring and control (Belot & Schröder, 2016; Cha & Wong-On-Wing, 2020; Christ, 2013; Enzle & Anderson, 1993; Falk & Kosfeld, 2006; Rudorf et al., 2018; Schweitzer et al., 2018), compliance (Mendoza & Wielhouwer, 2024; Mooijman et al., 2017), risky decision-making among politicians (Weinberg, 2023). Existing research in these contexts has predominantly focused on impersonal distrust, i.e., general and systemic distrust, and very few studies have examined consequences of interpersonal felt distrust, focusing on distrust by coworkers (Lanaj et al., 2018) and feeling ability-distrusted by supervisors (Liu et al., 2024). Both studies were limited to specific trust-relevant factors in their conceptualization and measurement of felt distrust, such as ability (Liu et al., 2024) and ability and integrity (Lanaj et al., 2018).

This small number of publications spread across diverse literatures and with little empirical attention to the outcomes of interpersonal felt distrust, indicates an absence of a dedicated stream of research on felt distrust in the organizational science domain. Moreover, the limitations inherent in previous studies, such as limited empirical inquiry of the underlying psychological mechanisms explaining the impact of felt distrust, as well as the emphasis on impersonal (rather than

interpersonal) felt distrust indicate a gap in our understanding of individual emotional and behavioral reactions to interpersonal felt distrust in the workplace. The following subsections present a systematic review of what is known about felt distrust.

2.2.1 References to Felt Distrust in Extant Research

Some research in the literature on monitoring, control, and rule compliance refers to felt distrust as an unintentional consequence of imposed formal control, or as a hidden cost of control (Belot & Schröder, 2016; Christ, 2013; Enzle & Anderson, 1993; Falk & Kosfeld, 2006; Frey, 1993). Scholars note that monitoring, when paired with a controlling intention, signals distrust and limited individual agency and autonomy, thereby negatively affecting individuals' intrinsic motivation and cooperation (Christ, 2013; Enzle & Anderson, 1993; Falk & Kosfeld, 2006). Some, like Falk & Kosfeld (2006), explain this through reference to the “self-fulfilling prophecy of distrust” (Luhmann, 1968) and with Self-Determination Theory (Deci et al., 1989). Others (Belot & Schröder, 2016; Christ, 2013; Enzle & Anderson, 1993) attribute it to the notion of negative reciprocity and the “crowding out effect” of felt distrust (Frey, 1993). Importantly, in these studies, felt distrust was not examined empirically as a standalone variable in the model; instead, the authors allude to felt distrust as an explanation for the negative attitudes and behaviors of individuals who were monitored.

2.2.2 Empirical Research on Impersonal Felt Distrust

Only a limited number of studies have empirically examined the consequences of felt distrust. Those that have (Cha & Wong-On-Wing, 2020; Mendoza & Wielhouwer, 2024; Mooijman et al., 2017; Rudolf et al., 2018; Weinberg, 2023) focused on the impersonal nature of felt distrust, reflective of institutional or organizational forms of distrust, rather than on interpersonal felt

distrust, which emerges in the context of an interpersonal relationship. In these studies, individuals as a collective were found to feel distrusted by authorities (Mendoza & Wielhouwer, 2024; Mooijman et al., 2017), organizations (Cha & Wong-On-Wing, 2020), and the public (Weinberg, 2023). In contrast, interpersonal relationships center on interaction between two or more persons (Weber, 1947). They necessarily involve each participant as a distinctive individual tied by reward dependability, ascription, commitment, investment, and attachment.¹ These factors blend and run together as a “social psychological glue that cements individuals into social units [in] interpersonal relationships” (McCall & Simmons, 1978).

Studies that looked at felt distrust by authorities, organizations and the public examined very specific outcomes of interest. These typically consisted of professional behaviors expected by the institutional norms and policies of the actors subject to public entities; for example, reporting revenue to tax authorities, students expressing willingness to engage in plagiarism, employees reporting on their performance, and risky decision-making in politicians.

For example, in a series of experiments, Mooijman et al. (2017) found that when sanctions were justified as attempts to deter people from rule breaking, individuals felt distrusted by authorities and complied less with the rules. This mediating effect was attenuated when the sanctions were targeted at other people (instead of the self) and when the authority was perceived as legitimate. Mooijman et al. (2017) explained this by referring to the theory of Interpersonal Justice (Tyler & Lind, 1992), the violation of the expectation of respectful interpersonal treatment

¹ Five elements fuel interpersonal relationships: *Reward dependability* entails relying on another person as a dependable source of affirmation of one’s prominent identities and other exchange rewards; *ascription* refers to legitimization and clear definition of the relationship in the eyes of the relationship actors and others around them; *commitment* serves as a strategy for increasing and sustaining the dependability of the source of exchange rewards; *investment* refers to respecting both parties’ contribution to the relationship and expecting substantial rewards from it; and *attachment* refers to the impact the other person has on the content, legitimization and enactment of one’s identity (McCall & Simmons, 1978).

(Ellemers, 2012; Ellemers et al., 2004), and “the self-fulfilling prophesy of distrust” (Luhmann, 1968).

In a quantitative study of national politicians in the United Kingdom, Canada, and South Africa, Weinberg (2023) examined descriptive levels of felt trust and felt distrust among politicians and evaluated the impact of these feelings on the politicians’ risky decision-making. As a result, Weinberg (2023) found that the MPs’ perceptions of public trust and distrust make a difference for risky decision-making. Contrary to the expected reflection effect (i.e., a tendency for people to avoid risk when facing gain but embrace risk when facing losses (Linde & Vis, 2017)), when politicians felt distrusted, they were *more* likely to engage in risky behavior in the face of potential gains but reduce the incidence of risk-taking behavior when facing potential losses. Weinberg (2023) suggested that the public’s negative perceptions of politicians may stimulate them to upweight rather than downweigh the probability of good outcomes in particular circumstances. Another theoretical explanation given by the author is that perceived public distrust may inhibit a politician’s belief in their ability to “make a difference” and receive praise for doing so, while at the same time increasing the saliency of anticipated punishment when things do go wrong (Weinberg, 2023).

In another study, Mendoza & Wielhouwer (2024) tried to separate the effect of felt distrust from the factors with which it has been commonly associated, such as control, monitoring, punishment, and examined felt distrust as an independent variable in the model. The authors sought to answer the question of how feeling distrusted by other members of society (such as immigrants or ethnic minorities), affects compliance with rules and regulations, exploring both the positive and negative effects thereof. They suggested that felt distrust affects compliance through adjustments in opportunism (Williamson, 1996). Contrary to previous findings on the negative

effects of felt distrust, the researchers found that distrust can induce both positive and conditionally negative effects and suggest that interventions signaling distrust may not always be inefficient. The authors regarded the negative effect as attributable to the theoretical notions of retaliation (Kruglanski, 1970) and negative reciprocity (Frey, 1993), and the positive effect as attributable to either a normative mechanism (i.e., the internalization of values through ethical deliberation and awareness of responsibility) or a strategic mechanism (i.e., strategic cooperation).

Consistent with the idea of retaliation, Rudolf et al. (2018) explained the effect of felt distrust through recourse to Reactance Theory (Brehm, 1966; Miron & Brehm, 2006). In contrast, Cha & Wong-On-Wing (2020) attributed the negative effect of felt distrust on employees' dysfunctional behaviors to the crowding-out effect of employees' reciprocity (Frey, 1993). Overall, while these studies inform us of the individual consequences of institutional or organizational distrust, they do not provide us with an understanding of how employees who feel distrusted by an important other at work, such as their direct supervisor, may react emotionally and behaviorally to that distrust.

2.2.3 Research on Interpersonal Felt Distrust in Organizational Setting

Only two studies to date have empirically examined the effect of interpersonal felt distrust on individual attitudes and behaviors in an organizational setting (Lanaj et al., 2018; Liu et al., 2024). Lanaj et al. (2018) found that employees feeling distrusted by their coworkers led to emotional exhaustion and, consequently, to withdrawal behaviors at work and increased conflict at home. The authors view felt distrust as a stressful event and explained its outcomes through the Conservation of Resources Theory (Hobfoll, 2001). They measured felt distrust through traits that capture perceptions of trustworthiness, such as integrity, ability and benevolence (Mayer et al., 1995). Liu et al. (2024) set to examine behavioral outcomes of employees feeling ability-distrusted

by supervisors. Drawing on the Self-Regulation Theory (Bandura, 1986), they found that shame mediated the relationship between feeling ability-distrusted by a supervisor and employees' knowledge hiding and deviant innovation behaviors. While these studies are helpful in understanding some of the emotional and behavioral outcomes to felt distrust, they have been limited in their focus to particular "trust-relevant factors" (Lanaj et al., 2018) as determinants of distrustworthiness. For example, Lanaj et al. (2018) focused on competence- and integrity-based mistrust when they examined the effects of felt distrust by coworkers on employee attitudes and behaviors. And Liu et al. (2024) focused on feeling ability-distrusted by supervisor in their study of the impact of felt distrust on employee behaviors in the workplace. Essentially, these studies examined the outcomes to perceived distrustworthiness, rather than to felt distrust. However, current theoretical thought in organizational science still lacks agreement on the dimensions of distrustworthiness (Six & Latusek, 2023), while empirical research provides mixed empirical support for any particular dimension (J. Cho, 2006; Dimoka, 2010; Sitkin & Roth, 1993; Sitkin & Stickel, 1996). Moreover, using opposite-valence dimensions of trustworthiness, i.e., ability, integrity and benevolence, in measuring felt distrust is not consistent with viewing distrust as separate and distinct from trust.

2.2.4 Research on the Outcomes of Felt Trust

Because some researchers do not distinguish between low felt trust and distrust (Mayer et al., 1995; Rotter, 1980; Schoorman et al., 2007) or measure felt distrust with opposite-valence factors of trustworthiness (Lanaj et al., 2018; Liu et al., 2024), I also cover the literature on the outcomes of felt trust in this chapter. Existing conceptualizations portray felt trust as a cognition or psychological perception of how much one is trusted by another; that is, a belief about how much vulnerability another person or entity is willing to assume in their expectation of another's

positive conduct. For example, *felt trust* has been defined as “the trusted other’s own perception of whether he or she is trusted by others” (Lau et al., 2007; Lau & Lam, 2008) or “the perception and realization of others’ positive expectations and exposed willingness to be vulnerable” (Lau et al., 2014). Overwhelmingly, research has suggested that felt trust is a positive individual experience that is beneficial to both the individual who experiences it and to the organization as a whole. As such, it has been associated with the positive emotion of pride (Baer et al., 2015), positive work attitudes in general (Gill et al., 2019; Skiba & Wildman, 2019), positive workplace behaviors such as increased organization citizenship behaviors (Lau et al., 2014; Lau & Lam, 2008; Zheng et al., 2019), loyalty to one’s supervisor (J. Deng & Wang, 2009), knowledge sharing (Nerstad et al., 2018), cooperation (Campagna et al., 2019; Chiu & Chiang, 2019), and enhanced work (Lau et al., 2014) and organizational performance (Salamon & Robinson, 2008). The theoretical mechanisms that explain the consequences of felt trust include responsibility norms (Salamon & Robinson, 2008), organization-based self-esteem (Lau et al., 2014; Zheng et al., 2019), felt obligation and moral licensing (X. Chen et al., 2019), psychological empowerment (Gill et al., 2019), and overall fairness perceptions (Baer et al., 2021).

Existing research on felt trust has been predominantly concentrated on understanding the positive effects of felt trust. However, less is known about how individuals react to not feeling trusted or feeling trusted to a lesser degree than expected, the lower range of felt trust. Past research has found that when individuals are not as trusted as they wish to be, their performance decreases (Baer et al., 2021) and work-life conflict is exacerbated (Gong & Sims, 2021). A study on the effects of low felt trust following the loss of felt trust found that it stimulated cooperation (Campagna et al., 2019). In the same vein, a study on the effect of low expectations (Nurmohamed, 2020) found that low expectations motivated better individual performance as those individuals

were motivated to prove others wrong. The existing research and theory on felt trust suggest that negative reactions to low felt trust are mostly explained by principles of social exchange and perceived unfairness. Positive reactions to low felt trust motivate the trustees to enhance the trustor's positive perceptions of their trustworthiness and promote the trustor's willingness to be vulnerable.

2.2.5 Summary of Research on the Outcomes of Felt Distrust

Research on felt distrust remains at a nascent stage. This research is dispersed across diverse literatures, such as economics, management, politics, and public administration. At this time, the organizational science domain, and specifically the distrust literature, lacks a dedicated branch of research on the study of interpersonal felt distrust.

Existing research on the outcomes of felt distrust can be found in the study of different issues, such as control and monitoring or punishment, whereby felt distrust is merely assumed to be signaled rather than empirically examined to confirm this assumption. Most of the previous research emphasizes impersonal distrust signaled by the organization, the public, or the authorities. Only a few of the existing studies have delved into the effects of interpersonal felt distrust, wherein distrust is signaled by one individual to another individual (Lanaj et al., 2018; Liu et al., 2024; Mooijman & Graham, 2018; Rudorf et al., 2018). However, most of this research to date is not conducted in the context of supervisor-employee relations (except for Liu et al. (2024) whose focus was limited to feeling ability-distrusted by supervisor). Considering that those in power are often inclined to distrust others, especially when their position or status is unstable (Inesi et al., 2012; Mooijman et al., 2015, 2017, 2019; Mooijman & Graham, 2018; Schilke et al., 2015), research is needed to understand how employees may react to feeling distrusted by a supervisor at work.

Theoretical explanations of the outcomes of felt distrust have to date predominantly focused on the notions of retaliation, negative reciprocity, and reactance. Importantly, these mechanisms have been suggested but not empirically tested. The only theoretical mechanisms of interpersonal felt distrust that have been tested to date include the Conservation of Resources Theory by Lanaj et al. (2018) and Self-Regulation Theory by Liu et al. (2024). However, these studies helped us better understand individual responses to perceived distrustworthiness rather than to felt distrust. Importantly, felt distrust was operationalized in line with factors of trustworthiness, which is inconsistent with contemporary view of distrust as a separate and unique construct from trust.

The neighboring literature on the outcomes of felt trust has mostly focused on positive outcomes for individuals, for example through improved job satisfaction (Lester & Brower, 2003), pride (Baer et al., 2015), loyalty to one's supervisor (J. Deng & Wang, 2009), organizational citizenship behaviors (Lau et al., 2014; Lau & Lam, 2008; Lester & Brower, 2003), and job performance (Brower et al., 2009; Nurmohamed, 2020). A handful of studies on low felt trust that examined individual negative (Baer et al., 2021) and positive (Campagna et al., 2019; Nurmohamed, 2020) reactions apply similar theoretical explanations to those in the existing research on the outcomes of felt distrust, such as injustice and retaliation.

In conclusion, this current review of research on the outcomes of felt distrust indicates a significant gap in our understanding of individual responses to feeling distrusted by another individual at work. Focusing on the essence of interpersonal felt distrust, which I define as one's social cognition that another individual thinks one has ill intentions towards them and fears that one intends to harm them, I suggest that being on the receiving end of distrust has an important impact on one's identity. Drawing on the distrust literature that views distrust as a separate

construct from trust, I offer an alternative theoretical perspective on the outcomes of interpersonal felt distrust that centers around its impact on the individual's identity at work and their innate desire to be seen as a good and moral human being. I review the literature on work-related identity, moral identity centrality, and identity threat in the following sections.

2.3 Work-related identity

Work-related identity is a domain-specific form of an individual's identity that contributes to the development of one's work-related self-concept (Miscenko & Day, 2016). The following paragraphs will address the origins of work-related identity to provide a better understanding of this complex and unique manner of understanding oneself.

Identity and the self-concept. The literature on identity spans sociology and social psychology, each offering distinct theoretical perspectives. Structural symbolic interactionists (e.g., McCall & Simmons, 1978; Stryker & Burke, 2000) emphasize role-based identities — internalized meanings tied to social roles that are embedded in institutional structures and organized hierarchically by salience. In contrast, social psychological theories such as Social Identity Theory (Tajfel, 1982; Tajfel & Turner, 1979; Turner & Oakes, 1986) and Self-Categorization Theory (Turner et al., 1987) center on group-based identities, focusing on how individuals define themselves through group memberships and categorize self and others in intergroup contexts. Complementing both is a personal identity perspective, which highlights the individual's internal self-definitions, values, and moral commitments that persist across roles and contexts (e.g., Blasi, 1984; Erikson, 1968; Gecas, 1982). Stryker's (1980) and Stryker & Serpe's (1994) work on identity salience also reflects the view that individuals prioritize certain identities based on their self-concept and the meanings those identities hold for them. These frameworks

collectively illustrate identity as socially situated, psychologically constructed, and dynamically shaped by roles, groups, and personal self-concepts.

In this dissertation, I draw on both structural symbolic interactionist (e.g., McCall & Simmons, 1978; Stryker, 1980; Stryker & Burke, 2000) and psychological (e.g., Blasi, 1984; Erikson, 1968; Gecas, 1982) perspectives to conceptualize personal identity. Broadly, identities are meanings that individuals attach to themselves; they form the foundation of one's self-concept (Gecas, 1982; Mead, 1934). *Personal identities* encompass meanings that develop from a combination of one's individual characteristics and life experiences (Blasi, 1993; Erikson, 1968), such as what it is like to be a caring person, and role-identities (McCall & Simmons, 1978; Stryker & Bruke, 2000), such as being an employee or a manager. Identity theory (Stryker & Burke, 2000) posits that these social roles provide structure and meaning to human behavior because each role and the identity associated with that role is associated with certain social expectations. The identity literature highlights the critical role of relationships with others in developing and shaping one's identity. For example, individuals use "reflected appraisals" (Cooley, 1902) to verify or confirm their identities at least partially through perceptions of how others treat and interact with them (Burke & Stets, 2009). And since one's self-identity is associated with their sense of self-worth (Gecas, 1982; Hogg & Terry, 2000), people generally attach positive value to their identities (Daniels & Robinson, 2019; Dutton et al., 2010) and seek to present their self-identity to others in any given situation (Aquino & Douglas, 2003; Erez & Early, 1993).

Identity in the Workplace. Research in sociology and social psychology has long established that an individual's identity is a multi-dimensional construct that consists of many different and interconnected subidentities (James, 1890; Stryker & Burke, 2000). One such subidentity is one's work-related identity. Different concepts have been used to refer to one's

identity at work: work identity (Miscenko & Day, 2016; Selenko et al., 2018), work-based identity (Van Dijk & van Dick, 2009), and work-related identity (Ashforth & Johnson, 2001; Dutton et al., 2010; Selenko et al., 2022). Since they essentially refer to the same phenomenon in the work domain, being a collection of meanings attached to the self (Gecas, 1982), I draw on research on identity at work from all of these authors in my dissertation, but I adopt the concept of “work-related identity” for consistency of definition.

A person’s *work-related identity* refers to a constellation of personal and social identities that are activated and made salient in the workplace (Ashforth & Johnson, 2001). This domain-specific identity is “tied to participation in the activities of work (i.e., a job) or membership in work-related groups, organizations, occupations, or professions” (Dutton et al., 2010, p. 266). Similarly to the more general notion of a person’s identity described in previous subsection (Petriglieri, 2011), a work-related identity includes aspects of both personal identity (such as one’s character, morals, values, associated roles) and social identity (such as one’s group membership in the workplace) (Miscenko & Day, 2016). The workplace is known to be a salient environment for the development and maintenance of our identities (Daniels & Robinson, 2019). People are motivated to maintain positive work-related identities because through those identities, they develop a sense of their work-related self (Dutton et al., 2010).

2.4 Moral identity centrality

The concept of *moral identity centrality* stands at the intersection of the rich literatures on identity and morality. It generally refers to the importance or salience of morality to one’s self-identity (Hardy & Carlo, 2011).

Morality and Integrity in the Workplace. Morality can be understood as a set of standards, rules, norms, and values that are expected of employees in the workplace with the intent to promote the general good of the organizational members (T. R. Cohen et al., 2014; Stoner et al., 2013). It is determined by adherence to a set of moral principles with integrity, which in turn refers to a steadfast consistency within one's commitments or principles and coherence between one's principles and their actions (McFall, 1987). To companies and employers worldwide, acting ethically and according to the law translates into saving billions of dollars annually in potential thefts, lawsuits, and settlements (J. W. Weiss, 2021). A high moral calibre of employees is beneficial for organizations because it is associated with ethical and responsible work practices and relationships, with enhanced profitability and competitive advantage, and with organizational resilience in the face of challenges and changes (Adeyeye et al., 2015).

Moral Identity. Early theorists have defined moral identity from the trait or character perspective. For example, Blasi (1984) conceptualized moral identity as an individual difference reflecting the degree to which acting morally is central or characteristic of a person's sense of self. Along these lines, in the organizational context, Dutton et al. (2010) have referred to moral identity as a type of a work-related identity imbued with virtuous qualities, such as being caring, compassionate, honest, kind, hardworking, and generous. In summary, the dispositional perspective on moral identity suggests that individuals with a strong moral identity are people who feel that moral values (such as being honest, fair, compassionate, and generous) are essential to their sense of personal identity (Hertz & Krettenauer, 2016).

More recently, building on and extending the early theorists' ideas on moral identity, other scholars (Aquino et al., 2009; Aquino & Reed, 2002; Lapsley & Narvaez, 2004; Mulder & Aquino, 2013) have offered a social cognitive approach to the conceptualization of moral identity

(Schwartz et al., 2011). Social cognitive theory takes a systems perspective on identity (Cervone & Tripathi, 2009). Specifically, scholars who follow social cognitive theory explain personality and identity as coherent and dynamic systems of social-cognitive and affective processes in a given individual that interact with situational influences (Cervone & Shoda, 1999). In this way, identity can be understood as a conception of interconnected beliefs, goals, and personal standards (Cervone, 2004). A social cognitive perspective on a person's moral identity states that it can be either more or less accessible within their overall self-concept and the degree of its accessibility can vary as an individual difference. In this dissertation, I adopt a social cognitive perspective on moral identity, meaning a self-conception organized around a set of moral traits (Aquino & Reed, 2002) or a "cognitive schema a person holds about his or her moral character" (Aquino et al., 2009, p. 124).

Moral Identity Centrality. According to the social cognitive approach, Aquino et al. (2009) suggest that situational cues that heighten moral knowledge content can temporarily enhance one's level of moral identity accessibility. In this way, certain individuals are able to maintain a more rather than less stable accessibility of moral identity. Aquino et al. (2009) describe individuals who experience chronic accessibility of moral identity as being high in *moral identity centrality*. According to Aquino & Reed (2002), the construct of moral identity consists of two dimensions, namely internalization and symbolization. *Internalization* refers to the degree to which moral traits (such as being caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind) are central to one's self-concept (Aquino & Reed, 2002). *Symbolization*, on the other hand, taps into the degree to which these moral traits are reflected in one's actions in everyday life, such as wearing clothes or engaging in hobbies that clearly identify one as having these characteristics and the degree to which one strives to communicate to others that one has these characteristics

(Aquino & Reed, 2002). The internalization dimension of moral identity is a valid measure of moral identity centrality (Aquino et al., 2009) as a strong indicator of the persistence of one's experience of embracing a moral identity as a part of their self-concept (Aquino et al., 2009; Aquino & Reed, 2002; Boegershausen et al., 2015).

2.5 Work-related identity threat

Identity Threat. Organizational behavior literature lacks agreement on what constitutes identity threat and offers multiple definitions of this concept. Early perspectives feature social psychological process theories on identity threat, viewing as a threat any obstacle to the enactment of the core identity processes. For example, one of the leading identity scholars in social psychology, Breakwell (1986), claimed that “a threat to identity occurs when the processes of identity, assimilation-accommodation and evaluation are, for some reason, unable to comply with the principles of continuity, distinctiveness and self-esteem, which habitually guide their operation” and that the very reason for this obstruction of the core identity processes constitutes that threat (pp. 46-47). Later, Aquino & Douglas (2003) defined identity threat in terms of particular behavior, as “any overt action by another party that challenges, calls into question, or diminishes a person's sense of competence, dignity, or self-worth” (p. 196).

Petriglieri (2011) defined individual-level identity threat as “experiences appraised as indicating potential harm to the value, meanings, or enactment of an identity” (Petriglieri, 2011, p. 644). This definition of identity threat is grounded in two established theories on identity in the organizational behavior literature: the identity theory (Stryker, 1987; Stryker & Burke, 2000) and the social identity theory (Tajfel & Turner, 1986). Petriglieri (2011) highlighted the critical role of a person's subjective appraisal of various life experiences as possible identity threats. She stressed that identity threats emerge from present cues of future harm to one's integrity and to the enactment

of one's particular identity, suggesting three potential sources of identity threat perception: the individual themselves, other people, and the material world. Most recently, Kroeper et al. (2022) defined identity threat as "a set of psychological concerns pertaining to the value of one's identity in a setting, including worries about belonging and authenticity as well as concerns about being devalued, disrespected, stereotyped, and marginalized in a setting" (p. 400). In my dissertation, I adopt their definition of identity threat to capture a set of employee concerns regarding how their identity is viewed and valued by their supervisors at work.

In the next chapter, I develop a model of the outcomes of interpersonal felt distrust in the workplace building significantly on the literature covered in this chapter.

Chapter 3: Theory and Hypotheses

In this chapter, I develop a theoretical model of the consequences of interpersonal felt distrust to fill the void in our understanding of how individuals react to feeling distrusted at work. This model considers the essence of interpersonal felt distrust (i.e., being perceived by another as having ill intentions to harm) and offers a novel identity perspective to explain individual reactions to distrust. Specifically, given the high importance of holding and presenting a positive work-related identity (Aquino & Douglas, 2003; Daniels & Robinson, 2019; Dutton et al., 2010; Erez & Early, 1993) and the influence others' opinions wield on our senses of self (Burke & Stets, 1999; Cooley, 1902; Mead, 1934), I suggest that feeling distrusted by a supervisor connotes being viewed as an ill-intentioned and harmful employee, which clashes with one's core identity as a good and moral being at work.

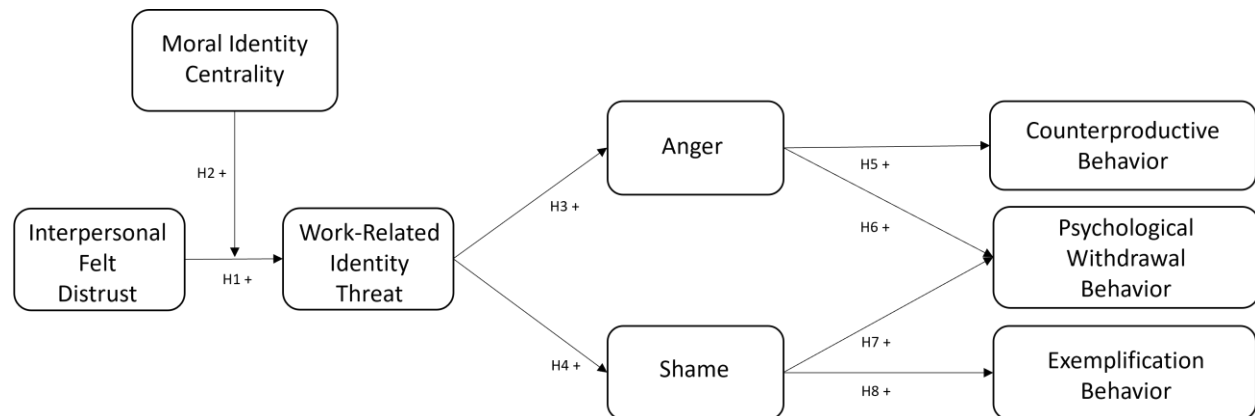
The key theoretical premise of this model is that interpersonal felt distrust triggers a work-related identity threat, and that this effect is more pronounced for individuals who hold moral identity central to their personal identity. As a theoretical mechanism, work-related identity threat is grounded in identity theory (Stryker, 1987; Stryker & Burke, 2000) and social psychological perspectives on identity (Blasi, 1993; Breakwell, 1986; Erez & Early, 1993; McCall & Simmons, 1978; Rosenberg, 1979), moral identity (Aquino et al., 2009; Aquino & Reed, 2002; Blasi, 1984; Hart, 2005; Hart et al., 1998; McFall, 1987), work-related identity (Dutton et al., 2010; Mischenko & Day, 2016; Selenko et al., 2018), and identity threat (Aquino & Douglas, 2003; Breakwell, 1986; Petriglieri, 2011).

Drawing on this theory and research, I develop hypotheses on the emotional and behavioral consequences of interpersonal felt distrust at work. Building on empirical research that finds that identity threat may trigger strong negative moral emotions of anger (Bies, 1999; Gilligan, 1996;

Steele, 1988) and shame (Daniels & Robinson, 2019; de Hooge et al., 2010; Tracy et al., 2007; Tracy & Robins, 2004), I suggest that work-related identity threat caused by feeling distrusted by a supervisor at work will trigger negative moral emotions of anger and/or shame in the individuals who experience felt distrust. As a result, employees are likely to engage in either counterproductive, psychological withdrawal, or exemplification behaviors at work. The choice of these behavioral outcomes is informed by the literature on the outcomes of anger and shame in the workplace. These behaviors are seen as important individual-level variables that are critical for organizational performance. This conceptual model and hypothesized relationships are presented in Figure 1.

Figure 1

Conceptual Model of the Outcomes of Interpersonal Felt Distrust in the Workplace



3.1 Felt distrust as a trigger of work-related identity threat

Since the quality of one's interpersonal relationships serves as an important source of identity validation (Darwin, 1872; Goffman, 1967; Lind & Tyler, 1988; Sennett & Cobb, 1973; Steele, 1988), I suggest that interpersonal felt distrust threatens one's work-related identity as a

moral being and a valued employee. Specifically, when an individual perceives that their supervisor distrusts them (i.e., ascribes ill intentions to harm the other party to them) (Lewicki et al., 1998), they come to realize that they are being seen as an evil or otherwise harmful person, which contradicts their general aspiration to be seen as a good human being and consequently damages their self-image. Individuals are strongly motivated to enact and maintain identities that are important to them because doing so provides them with a source of high self-regard and a sense of stability and continuity over time (Shamir, 1991). Interpersonal felt distrust by one's supervisor is therefore expected to evoke a particularly strong threat to one's work-related identity because of the higher weight placed by individuals on evaluations of their character made by organizational members with high (and higher) status and power roles.

Research on the social origins of self-knowledge indicates that individuals' self-concepts are greatly influenced by how they perceive themselves as being viewed by others, an assertion framed by scientists as the "looking glass self" (Shrauger & Schoeneman, 1979). Felt distrust by an important other in the workplace, such as one's supervisor, signals that the distruster sees one as ill-intended and a potential source of harm. Much of pro-social and leadership behavior generally is driven by the motivation to be seen and to see oneself as a positive and influential being at work (Dutton et al., 2010). Moreover, important established social norms are violated by distrust, which in turn damages one's positive work-related identity insofar as individuals generally see themselves as trustworthy and trusting of others (J. D. Brown, 2012; Sedikides et al., 2014) and expect others to trust them in return (Ellemers, 2012; Tyler & Lind, 1992). Because interpersonal felt distrust by a supervisor connotes a negative expectation of harm from an authority figure at work, it results in employees questioning their positive self-view in the

workplace and serves as a threat to employees' work-related identity. I therefore suggest the following:

H1: Interpersonal felt distrust leads to work-related identity threat.

Since people are more likely to pay attention to identity-related information when it relates to an identity they consider important (Stryker & Serpe, 1994), they are more likely to deem a particular experience as identity threatening when it relates to an important identity (Petriglieri, 2011). Research suggests that individuals high (but not low) on moral identity centrality (i.e., the internalization dimension of moral identity) are particularly attuned to threats to moral identity posed by situational factors (Effron, 2014). Individuals who place a high value on their moral identity experience even stronger work-related identity threat following an experience of felt distrust. I therefore argue that moral identity centrality (Aquino & Reed, 2002) in a focal employee will strengthen their experience of work-related identity threat following felt distrust and suggest the following second hypothesis:

H2: The positive relationship between interpersonal felt distrust and work-related identity threat will be stronger for individuals with higher (vs. lower) moral identity centrality.

3.2 Emotional reactions to a threatened work-related identity

While we know that distrust is a heavily emotionally loaded attitude (Lewicki et al., 1998), we do not know if and what emotions may emerge on the receiving end of distrust and how these emotions may influence employee behaviors in the aftermath of felt distrust. So far, no research previous has investigated the effects of felt distrust on individuals' emotions. In their conceptual paper, Mooijman & Graham (2018) mentioned that felt distrust can increase negative emotion, such as anger; however, this assertion was not empirically tested.

A rich body of research reveals that daily experiences at work have a profound impact on employees' emotions, attitudes, and behaviors overall (Ashkanasy et al., 2000; Averill, 1982; Brief & Weiss, 2002; Frijda, 1986; Izard, 1991; Lazarus, 1999; Schwarz & Clore, 1996). Affective events theory (H. M. Weiss & Cropanzano, 1996) is often used to explain the impact of employee perceptions and experiences of salient events at work on their reactions in the work setting. Specifically, an affect-initiated processing route (Douglas et al., 2008) has been suggested as one of the theoretical pathways through which to understand organizational members' emotional reactions to negative events at work and associated behaviors thereto.

I suggest that work-related identity threat is a salient workplace experience that triggers strong emotions in individuals. Indeed, the existing research documents that identity threat is often associated with intense negative moral emotions that signal a specific problem to be solved and stimulate certain behaviors to solve the problem (de Hooe et al., 2010; Douglas et al., 2008). In particular, empirical studies provide strong evidence that threats to identity may trigger strong negative moral emotions of anger (Bies, 1999; Gilligan, 1996; Steele, 1988) and/or shame (Daniels & Robinson, 2019; de Hooe et al., 2010; Tracy et al., 2007; Tracy & Robins, 2004).

Past research provides robust evidence that anger is one of the primary emotions generated by identity threat at various levels of its conceptualization, whether those be interpersonal, relational, or collective (Bies, 1999; Gilligan, 1996; Steele, 1988). For example, when individuals perceive interpersonal rejection (events in which they feel devalued, unaccepted, or rejected), they experience anger (Leary et al., 2006). In a similar vein, research shows that employees express anger in response to actions they perceive as interpersonal mistreatment (Aquino et al., 2004). Building on this research, I suggest that work-related identity threat triggered by feeling distrusted

by a supervisor at work will evoke a moral negative emotion of anger in the target employee. As such, I propose the following:

H3: Work-related identity threat (a) leads to anger; and (b) mediates the positive relationship between interpersonal felt distrust and anger.

Another emotion that may be experienced as a result of threatened work-related identity following felt distrust by a supervisor is shame. Just like anger, shame is a negative moral emotion (Barclay et al., 2005; Tangney & Dearing, 2002). As a moral emotion, shame functions as “an emotional moral barometer, providing immediate and salient feedback on our social and moral acceptability” (Tangney et al., 2009, p. 192). In fact, in their empirical study, Woodyatt & Wenzel (2014) found shame to be a marker of threat to moral identity. Unlike anger, shame is an internally focused and painful social emotion (Barclay et al., 2005; Tangney & Dearing, 2002). As a social emotion, shame arises from one’s realization of one’s own shortcomings when compared to accepted social standards and is commonly experienced in the workplace (Daniels & Robinson, 2019).

Like guilt, pride, and embarrassment, shame belongs to the family of self-conscious emotions, which means that it is evoked through self-reflection and self-evaluation (Tangney et al., 2009). Due to its self-conscious nature, shame typically arises when one evaluates a threat to the self in relation to having failed to live up to an important standard tied to one’s identity (M. Lewis, 1992; Tangney & Dearing, 2002; Tracy et al., 2007; Tracy & Robins, 2004). It is a painful social emotion (Daniels & Robinson, 2019a) whereby its social facet reflects one’s concerns about their social standing (P. Gilbert, 1997) and appearance in the eyes of others (H. B. Lewis, 1971). Daniels & Robinson (2019) emphasized that “the role of other people likely has a strong effect on the self-evaluation that lies at the heart of shame because people often look to feedback from others

when evaluating the self” (p. 2454). They highlighted the important role of other people’s condemnations in triggering a process of self-devaluation that leads to shame (Daniels & Robinson, 2019). Shame is most likely to be felt when a person’s identity attracts the disapproval of others and is particularly likely to be strong when disapproval comes from an authority figure or from a significant other (Braithwaite et al., 2008; Daniels & Robinson, 2019). Such characteristics of shame make it a likely emotion to develop following a threatened sense of work-related identity after feeling distrusted by a supervisor at work.

Moreover, in their review of theory and research of self-conscious motives and emotions, Leary (2007) stated that self-conscious emotions, including shame, “arise in response to events that have real or imagined implications for others’ judgments of the individual” (p. 317). Leary (2007) further stated that shame results from one’s sense of failure to live up to certain accepted standards and worry about how one appears to others for having violated these standards. In their empirical study, Smith et al. (2002) confirmed earlier research findings that shame is strongly associated with individuals’ negative self-evaluation and public exposure of both moral (transgressions) and nonmoral (incompetence and inferiority) experiences. In their dissertation, Welten (2011) found that shame increased both as a function of threats to one’s personal self (i.e., falling short of one’s own personal standards) and to one’s social self (i.e., falling short of the standards of others). Therefore, the existing research provides strong evidence for an association between one’s threatened identity, including moral identity, and the emotional response of shame.

Consistent with the research on shame-evoking processes in organizational settings (Daniels & Robinson, 2019), I suggest that the realization that one is perceived as a source of harm to a partner in an interpersonal relationship and, therefore, that one falls short of virtuous moral standards may evoke a moral emotion of shame. This emotional response is made even stronger

when felt distrust is experienced as coming from a significant other at work, such as one's supervisor. Therefore, I suggest the following:

H4: Work-related identity threat (a) leads to shame; and (b) mediates the positive relationship between interpersonal felt distrust and shame.

3.3 Behaviors following emotional reactions to a threatened work-related identity

Individuals value their identities as means to sustaining their senses of self-worth (Bies, 1999; Brockner, 1988; Gecas, 1982; Tajfel & Turner, 1986); therefore, individuals are motivated to defend themselves against acts that threaten their identities (Baumeister et al., 1996; Felson & Steadman, 1983; Steele, 1988). Emotions experienced as a result of identity threat signal a specific problem to be solved (Frijda, 1986; Zeelenberg & Pieters, 2006) and activate a distinct motivation that triggers certain behaviors to solve the problem (Plutchik, 1962; Roseman, 1984; Zeelenberg et al., 2008), according to the functional approach to emotions (de Hooij et al., 2010). Consistent with past research that has found that emotions influence individual behaviors in the workplace (Douglas et al., 2008; Forgas & George, 2001; Forgas & Smith, 2003; Shaver et al., 1987), I anticipate that anger and shame, as moral self-conscious emotions that emerge as a result of work-related identity threat, will motivate individuals to engage in retaliatory and/or impression management behaviors. Specifically, in my model, I suggest that anger, following work-related identity threat triggered by felt distrust, may lead to counterproductive and psychological withdrawal behaviors, and shame may lead to psychological withdrawal and exemplification behaviors in the workplace. I will discuss this suggestion further in the following paragraphs.

Anger to Counterproductive Work Behavior. Anger is a moral (Haidt, 2003), outward-focused, and negative (Barclay et al., 2005; Tangney & Dearing, 2002) other-condemning (Rozin

et al., 1999) emotion. Studies have found that anger can energize individuals to “right wrongs” and deter future injustices (Bies & Tripp, 2002). Anger can also motivate individuals to engage in retaliatory behaviors against the transgressor (Allred, 1999, 2000; Allred et al., 1997; Averill, 1982). Existing research provides strong evidence that anger leads to interpersonal aggression (Aquino & Douglas, 2003; Averill, 1982; Felson, 1992; Tedeschi & Felson, 1994), retaliation (Allred, 1999, 2000; Barclay et al., 2005; Skarlicki et al., 1999), and abuse against others (Fox et al., 2001; Khan et al., 2013). Aggression allows one to restore their sense of self and release their anger (Baumeister et al., 1996; Crocker & Park, 2004). Past research has also found support for the mediating role of anger in the relationship between perceived injustice and retaliation (Barclay et al., 2005), as well as perceived injustice and counterproductive work behaviors (Khan et al., 2013; Krehbiel & Cropanzano, 2000).

Counterproductive work behaviors (CWBs) refer to a group of aggressive (e.g., production deviance and abuse against others) or passive (e.g., employee failing to follow instructions) workplace behaviors aimed at harming the organization and/or its members (Fox et al., 2001). They represent an intentional, approach-oriented form of workplace deviance (Mangione & Quinn, 1975; Robinson & Bennett, 1995; Robinson & Greenberg, 1998) in which workplace deviance is understood as “voluntary behavior that violates significant organizational norms and in so doing threatens the well-being of an organization, its members, or both” (Robinson & Bennett, 1995, p. 556). CWBs have detrimental effects on individuals and organizations because of their harmful effects on the organization and/or its employees (Aubé et al., 2009; Jensen et al., 2010; D. A. Jones, 2009; Penney & Spector, 2005). Since distrust is felt from the supervisor, on whom the employee depends for performance evaluation and career outcomes, I propose that employees are less likely to display “personal aggression” or abuse their supervisor. Studies find that individuals are less

likely to take revenge when their targets are highly likely to retaliate (Petersen, 2010), more powerful than they are (Aquino et al., 2006; Jermier et al., 1994), or when third parties are present (S. H. Kim et al., 1998). Therefore, the anger-provoked counterproductive behavior will be shifted towards the organization in general, as a way of retaliating against the distruster. Therefore, I suggest the following:

H5: Anger: (a) leads to counterproductive work behavior; and (b) mediates the positive relationship between work-related identity threat and counterproductive behavior.

Anger to Psychological Withdrawal Behavior. In certain situations, employees may find that overt retaliation when experiencing anger may be too costly (Heese & Pérez-Cavazos, 2021; Homans, 1961) or socially inappropriate (Zinner et al., 2008). In such cases, an employee who feels angry in response to a threat to their work-related identity may consider more avoidance-based forms of behavior as a form of retaliation. Specifically, the distrusted individual may engage in psychological withdrawal behaviors, i.e. cognitive disengagement from their work situation, such as withholding effort and thinking of leaving the job (Lehman & Simpson, 1992). Such behavior represents a more covert form of retaliation (Homans, 1961; Skarlicki & Folger, 1997). While less dramatic than overt forms of retaliation, psychological withdrawal behaviors nevertheless pose adverse consequences for an organization's effective functioning (Skarlicki & Folger, 1997).

Justice research provides evidence for the association between negative emotions and psychological withdrawal behaviors (Barclay & Kiefer, 2014). In their study, Barclay & Kiefer (2014) found that negative emotions fully mediated the relationship between overall justice perceptions and psychological withdrawal. While much of the literature finds that anger affects

counterproductive behavior (e.g., Douglas & Martinko, 2001; Ferris et al., 2016; Fox & Spector, 1999; Fox, Spector, & Miles, 2001; Goldman, 2003; Lee & Allen, 2002), some scholars have suggested that psychological withdrawal could also be a way of coping with the anger (Deffenbacher, 2011; Palfai & Hart, 1997). Although this reasoning has not received much empirical attention, I draw on its logic to argue that employees who experience anger following work-related identity threat after having felt distrusted by their supervisor will engage in psychological withdrawal as a form of covert retaliation. Therefore, I suggest the following:

H6: Anger: (a) leads to psychological withdrawal behavior; and (b) mediates the positive relationship between work-related identity threat and psychological withdrawal behavior.

Shame to Psychological Withdrawal Behavior. The existing body of research provides ample evidence that shame may lead to psychological withdrawal or prosocial impression management behavior in organizational setting (Daniels & Robinson, 2019; de Hooge et al., 2008, 2010, 2011, 2018). Shame has been described as a “global, painful, and devastating experience in which the self, not just behavior, is painfully scrutinized and negatively evaluated” and “is often accompanied by a sense of shrinking and being small, and by a sense of worthlessness and powerlessness” (Tangney, 1991, p. 599). It has long been associated with social withdrawal, hiding, disappearing, inhibited social interactions, and isolation of oneself from others at an interpersonal level (Breugelmans & Poortinga, 2006; Fontaine et al., 2006; Haidt, 2003; Keltner & Buswell, 1996; Keltner & Harker, 1998; H. B. Lewis, 1971; M. Lewis, 1992; Lindsay-Hartz, 1984; Tangney, 1989, 1991; Tangney et al., 1996; Wicker et al., 1983).

Organizational scholars note that shame leads to withdrawal behaviors in an attempt to protect oneself from further threats to self and social image, and to distance oneself from a shame-inducing social environment (Daniels & Robinson, 2019). People who experience shame following

an interpersonal event tend to avoid contact with the other party for fear of inducing further shame (Tangney et al., 2009). Based on research on behavioral responses to shame following a threat to self and social image, individuals are likely to avoid the source of the threat to further protect their moral identity (Daniels & Robinson, 2019), especially when engaging in approach behaviors is not possible, is too difficult, or is too risky (e.g., de Hooge et al., 2010, 2011). Building on this knowledge, I argue that shame will trigger psychological withdrawal in the workplace as a way to protect the threatened identity from further harm. Therefore, I suggest the following:

H7: Shame: (a) leads to psychological withdrawal behavior; and (b) mediates the positive relationship between work-related identity threat and psychological withdrawal behavior.

Shame to Exemplification Behavior. While shame often leads to negative behaviors, it can also lead to positive ones. Existing research provides plenty of evidence of prosocial interpersonal behavioral responses following shame that aim to repair the threatened self (Daniels & Robinson, 2019; de Hooge et al., 2008, 2010, 2011, 2018) and present more positive aspects of the self (Gausel & Leach, 2011; Gruenewald et al., 2007; Schmader & Lickel, 2006; Tangney et al., 1996). Such prosocial behaviors include exemplification (Daniels & Robinson, 2019), which refers to a form of strategic self-presentation behavior used to communicate some information about oneself to others (Baumeister, 1982). Self-presentation behaviors are guided by two core motives: first, to please others, and second, to construct (create, maintain, and adapt) one's public self as consistent with one's ideal self (Baumeister, 1982). Prosocial impression management behaviors help minimize relational losses (Goffman, 1955; Leary et al., 1996; Leary & Miller, 2000) and improve the way that one is viewed by others (Baumeister, 1982; B. R. Brown, 1970; Leary & Kowalski, 1990; Leary & Miller, 2000), which are both desired goals following one's experience of shame.

Exemplification is a specific type of prosocial impression management behavior used to enhance one's self-image in relation to those around them (E. E. Jones & Pittman, 1982; Turnley & Bolino, 2001). Exemplification emphasizes one's desirability, dedication, and value as a relational partner (Bolino, 1999; Bolino & Turnley, 2003) and demonstrates to others that one is moral, helpful, and cooperative (Bonner et al., 2017). Exemplification includes self-sacrificial behaviors that exceed others' expectations (Bolino et al., 2008; Rosenfeld et al., 1995), for example taking on heavy workloads, arriving to work early and staying late, and working on the weekends (Bolino & Turnley, 1999). Prior research suggests that individuals pursue exemplification behaviors to signal their social and moral worthiness to those around them (D. T. Gilbert & Jones, 1986; E. E. Jones & Pittman, 1982; Stone et al., 1997). Recent empirical research suggests that shame motivates employees to pursue exemplification behaviors to protect one's threatened and devalued self as well as to demonstrate to others that one is a dedicated and hardworking employee (Bonner et al., 2017). I argue that individuals who feel shame following a work-related identity threat may display exemplification behavior to improve how they are seen by important others at work and to signal to people around them that they are worthy of association. Therefore, I suggest the following:

H8: Shame: (a) leads to exemplification behavior; and (b) mediates the positive relationship between work-related identity threat and exemplification behavior.

3.4 Conclusion

My model suggests that interpersonal felt distrust at work will trigger work-related identity threat, which will in turn evoke emotions of anger and shame in individuals who feel distrusted by their supervisor. My model suggests that upon the experience of anger, the distrusted individual will engage in counterproductive interpersonal work behavior or psychological withdrawal

behavior. Upon the experience of shame, in contrast, the distrusted individual will engage in psychological withdrawal or exemplification behavior. In the following two chapters (chapters 4 and 5), I describe two empirical studies I undertake to test the stated hypotheses in this model and present my findings.

Chapter 4: Study 1 Methods and Results

4.1 Overarching Approach to Research Design

To test the hypotheses presented in this dissertation, I undertook a multimethod research design (Wellman et al., 2023) by utilizing complementary deductive approaches across two studies. Study 1 comprised an experiment and Study 2 comprised a critical incident technique-based survey. Study 1 was designed to test the causal effect of interpersonal felt distrust by a supervisor on an individual's work-related identity threat and subsequent emotional reactions thereto (Hypotheses 1, 3 and 4) as well as the moderating role of moral identity centrality (Hypothesis 2). Study 1 was set up as an experiment to ensure a high level of internal validity in the study results (Shadish et al., 2002). The sample consisted of individuals who had experienced being supervised at work. The experimental manipulation of felt distrust allowed examining the causal effect of felt distrust by a supervisor on one's work-related identity threat. Further, the mediating effect of work-related identity threat on emotional outcomes was examined in the context of a serial mediation, hence limiting the study's ability to establish a firm causal inference from felt distrust to work-related identity threat and onto emotions.

Study 2 aimed to replicate Study 1 and test Hypotheses 1–4 through recourse to a different research design to enhance the external validity of the research findings. In addition, Study 2 tested the individuals' behavioral reactions to interpersonal felt distrust by a supervisor (Hypotheses 5--8). It was designed by integrating the critical incident technique (CIT) (Flanagan, 1954) in a survey. Integrating CIT into survey design has been recommended as an effective approach to capturing individual reactions to real-life workplace events, such as various forms of mistreatment (Hershcovis, 2011). This technique has been widely used in organizational science research as a valid tool to examine individuals' responses to personal offense (Aquino et al., 2001), conflict and

trust (Korsgaard et al., 2002; Lind et al., 1997; Lissak & Sheppard, 1983), unethical behavior (Mitchell et al., 2015), moral objection (Wellman et al., 2016), gender threats (Leavitt et al., 2022), and organizational error management culture (Van Dyck et al., 2005). As part of the critical incident technique, Study 2 participants were asked to recall and reflect on their experience of feeling distrusted by a supervisor in the workplace, followed by a series of survey questions assessing their emotional and behavioral reactions to the event. The sample consisted of individuals who had experienced feeling distrusted by a supervisor at work. Study participants were recruited from Prolific for both studies. Analysis of the serial mediation within this complex cognition–emotion–behavior pathway provided insight into the associations among the variables of interest. However, the study was not designed to support causal inferences, as the absence of temporal precedence between variables and the potential for alternative explanations limit the ability to determine directionality or rule out confounding influences.

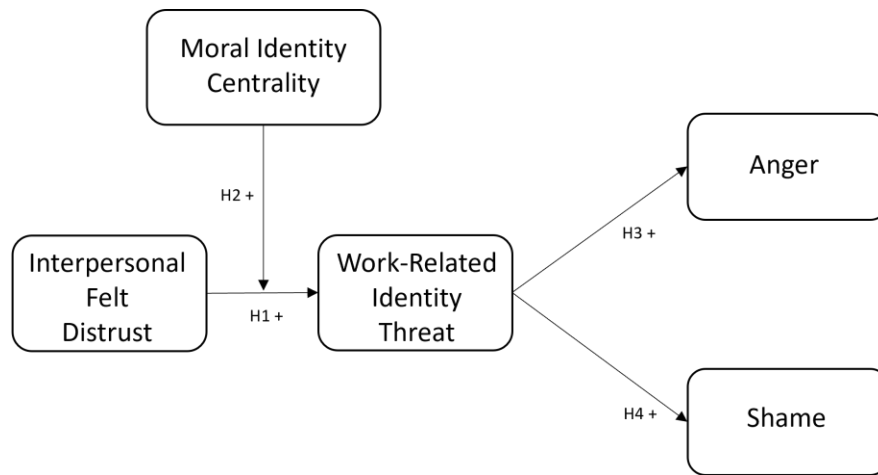
Combining these different quantitative methodologies and diversifying the target population across the two studies provided for more rigorous testing of the core mediating mechanism, namely work-related identity threat, and the associated emotional reactions to interpersonal felt distrust (i.e., anger and shame). This triangulation served to enhance the internal and external validity of the research to the extent possible by the decisions on the studies' designs.

4.2 Study 1 Description

In this study, I empirically tested hypotheses 1—4. The hypothesized relationships and the hypotheses under study are depicted in Figure 2 below.

Figure 2

The Core of the Conceptual Model Analyzed in Study 1



4.2.1 Sample and Procedure

4.2.1.1 Sample size determination. Sample size was determined using G*Power 3.1.9.7 (Erdfelder et al., 2009; Faul et al., 2007). An a priori power analysis (effect size = .50, α = .05, power = .80, df = 1378) indicated a required sample size of 545. The output parameters were: noncentrality parameter λ = 136.25, critical χ^2 = 1465.47, and actual power = .80. This recommended sample size for SEM analysis is consistent with the recommendations of Kline (2016) (i.e., the n:q rule, where n is the sample size and q is the number of parameters in the model, with a recommended ratio of at least 10:1 for models with latent variables). With 52 free parameters in the Study 1 model, the 10:1 rule produced a minimum 520 sample size to run the SEM model and provide for decent effect size, power, and significance levels. Eventually, I decided to oversample and recruited 610 participants. This sample size is expected to yield good estimates of effect size detected and a high statistical power of the analysis itself.

4.2.1.2 Sample characteristics. The sample participants were recruited through Prolific, an online platform that specializes in maintaining a high quality pool of respondents that complete

surveys for fair pay. Among researchers, Prolific data has been of higher quality in terms of participant attention to survey questions, diligence, and responsibility, especially when compared to other popular platforms, such as Amazon Mechanical Turk (Palan & Schitter, 2018; Peer et al., 2017). Prolific has been a source of high-quality respondent data for a number of experimental studies in the fields of management, organizational behavior, and human resources management to date (i.e., Cho & Jiang, 2022; Harari et al., 2022; Kim et al., 2021; Lyubikh et al., 2022; Yan et al., 2022).

The sample was representative of the population of the United States according to sex, age, and race. The following eligibility criteria were applied to Prolific participants and further validated through select survey questions: age (20-85 years old), employment status (full-time, part-time, and due to start a new job within the next month), whether the person has a direct supervisor at work, and years lived in the country of residence (filtered to a minimum of three years residing in the United States).

4.2.1.3 Survey design. Study 1 included an experiment based on a two-condition (Distrust vs. Neutral) between-person factorial vignette design with independent measures, thus resulting in two experimental scenario frameworks. The respondents were randomly assigned to read one of the two hypothetical workplace scenarios (i.e., the written “paper people” vignettes) (Aguinis & Bradley, 2014). Both scenarios reflect an organizational setting in which employees are required to request a password from their supervisor in order to access a sensitive dataset. Such an organizational context is realistic in that it resembles that of organizations that prioritize the security and confidentiality of their resources and sensitive data. This setting was held constant in both scenarios to provide for a similar baseline. The Neutral scenario presented this organizational setting as a policy that applies to all employees. The Distrust scenario, in contrast, included a

continuation of the storyline involving a case where a supervisor exhibits a high degree of interpersonal distrust on the focal employee in the scenario. This scenario was designed by drawing on existing theoretical and empirical research on distrust in the workplace (Bijlsma-Frankema et al., 2015; Guo et al., 2017; Lewicki et al., 1998; Sitkin & Bijlsma-Frankema, 2018; Six & Latusek, 2023) to depict behaviors that are typical when one person distrusts another person. Both scenarios are included in Appendix A. The study was set up on Qualtrics and study participants were randomly assigned to read one of the two scenarios as they accessed the survey link.

4.2.1.4 Survey procedure. Eligible Prolific panel members were invited to follow the link to the online survey. Consenting participants were asked to respond to a series of questions that consisted of three validation questions (i.e., self-reported responses to the questions enumerated above regarding employment status, experience being supervised at work, and number of years lived in the country of residence). Those who passed the validation questions were shown one of the two randomly assigned scenarios. This was followed by a manipulation check, eight questions measuring the variables of interest (being identity threat, anger-experience, shame-experience, moral identity centrality, anger-trait, shame-trait, narcissism, and neuroticism), and three demographic questions (sex, age, and race). Participants were required to respond to each survey question without skipping any of the questions. The study upheld ethical integrity requirements by allowing participants to withdraw at any time while still receiving compensation for their participation. Throughout the survey, the participants were offered clear instructions on what to expect on the next page to ease the flow of the survey and provide a better survey experience for the participants. After completing the survey, the respondents were thanked and debriefed in the end-of-study message.

4.2.1.5 Testing the survey. Before launching the core study, I conducted meticulous testing of the survey for clarity, ease of flow, aesthetic appeal, and technical issues. The survey was tested first on a sample of ten Prolific participants, and then on ten Ph.D. students and recent graduates. I used their feedback to improve survey design and its technical characteristics. I met with the ten Ph.D. students and recent graduates for brief conversations to collect their feedback and reflections on the survey. As a result, important revisions were made to the survey. Hyperlinks to online Dictionary entries were added for complex words to help with definitions and synonyms, a progress bar was added at the bottom of the survey to increase transparency and provide for realistic expectations on survey progression, the end of survey message edited for consistency with the survey questions and the content in the informed consent form, the font and size adjusted for better readability, and other small adjustments.

4.2.1.6 Launching the survey. Data was collected on Prolific between June and July 2024. Altogether, a total of 687 members initiated the survey, but only 610 completed it and were considered valid by Prolific. The 77 invalid attempts at the survey were either due to (1) the survey not being completed (11 members started and never finished, hence timed out); (2) returned by Prolific (64 returned submissions were due to participants not passing validation check questions or screening-out questions); or (3) rejected by members (informed consent not given by 2 members). The valid 610 responses all included consent to participate in the study and subsequently passed validation, attention, and seriousness checks. Since all of the survey questions were enforced, there was no missing data in the survey responses. The 610 respondents were approved by me as the researcher and were paid for their participation in the survey.

4.2.2 Ensuring data quality

The Study 1 Qualtrics survey design incorporated elements of security control and data quality checks in line with best practices for the design of surveys for online panel participants (Aguinis & Bradley, 2014; Atzmüller & Steiner, 2010; Palan & Schitter, 2018; Tristram et al., 2012a, 2012c, 2012b). Participants' Prolific IDs were captured through embedded data in the Qualtrics survey to guarantee there was no unauthorized access to the survey beyond registered panel members. QCaptcha scores (i.e., Prolific's in-built system for bot detection) were collected for each participant response based on their pattern of interaction with the survey (i.e., the length of time it took to complete, response pattern, etc.). The participants' IP addresses were captured in the metadata to ensure that each survey participant completed the survey only once.

Special care was taken to address insufficient efforts in responding (IER) or careless responding (CR) answering behavior. Existing research suggests that a minimum of 10-12% of participants in an online survey (Meade & Craig, 2012) and as much as 46% of participants of a crowdsourced sample (Brühlmann et al., 2020) may exhibit these behaviors. Failure to identify CR undermines survey data quality and poses a threat to the validity of the survey (Meade & Craig, 2012). Following best practices (Eerola et al., 2021; Landers & Behrend, 2015; Risser et al., 2024), the Study 1 online survey included two attention checks and one seriousness check, which were required to be passed for the survey responses to be considered quality data. Attention and seriousness screeners are recognized as effective means to identify the least attentive respondents in applied survey research (Berinsky et al., 2021; DeSimone & Harms, 2018; Porter et al., 2019). The questions used in the two attention checks were: "If you are reading this statement carefully, please select "Agree"" and "If you're reading this question now, please respond "Disagree"". The participants had to choose the requested option from a list of five possible options ranging from

(1) Strongly Disagree to (5) Strongly Agree. The wording of the question is consistent with other researchers' approach to attention questions in online surveys (i.e. Kundro & Nurmohamed, 2021; Lyubykh et al., 2022; Tost et al., 2022) The participants were explicitly told that the question checks their attention and were instructed to respond to each attention check carefully. They were also reminded that they have to pass both attention checks, and the seriousness check to be eligible to contribute to this survey.

Past research suggests that a considerable portion of online survey participants (i.e. 3–28%) misrepresent information about themselves, their experiences, and their skills, which contributes to the problem of insufficient and careless responses (MacInnis et al., 2020). Respondents' potential dishonesty in representing their identity or experiences may introduce systematic bias into research findings, which can influence the accuracy of estimates and distort the strength of the effect (i.e., by making the relationships appear weaker (Siegel et al., 2015; Siegel & Navarro, 2019) or stronger (Wessling et al., 2017) than they would be in a true population of respondents). To address this problem, scholars have recommended the inclusion of some form of a seriousness check in an organizational survey whereby the survey participants respond to a straightforward and transparent question about whether they took the survey seriously and responded honestly (Aust et al., 2013; DeSimone & Harms, 2018; Eerola et al., 2021; Meade & Craig, 2012).

To improve the data accuracy and validity in the Study 1 survey, I used a revised version of a seriousness check recommended by Aust et al. (2013, p. 530). The participants were told that this question checks the seriousness of their efforts on the survey, that an appropriate response is required to consider their input eligible for data analysis, and that they should answer honestly. The revision consisted of adding "please use my data for analysis" to the positive answer option. The seriousness check question was included at the very end of the survey and reads as follows:

“It would be very helpful if you could tell us at this point whether you have taken part seriously, so that we can use your answers for our scientific analysis, or whether you were just clicking through to take a look at the survey?” The respondents were asked to select one of the two options: “1. I have taken part seriously, please use my data for analysis. 2. I have just clicked through, please throw my data away.” Past research has found that survey participants who did not respond to survey questions with due care were likely to honestly report this, which improved the quality of data collected and the validity of research findings (Aust et al., 2013; Meade & Craig, 2012).

4.2.3 Measures

4.2.3.1 Dependent Variables

Anger-Experience. This variable was measured with a 6-item Hostility subscale developed by Watson & Clark (1994). This scale consists of words and phrases that describe different feelings and emotions, with sample items including: (1) “Angry”, (2) “Hostile”, and (3) “Irritable”. The participants were asked to reflect on how imagining themselves as the focal employee in the scenario they have just read makes them feel. They indicated their responses on a 4-point Likert type scale ranging from (1) very slightly or not at all to (5) extremely in relation to the items on the Hostility subscale. The Cronbach’s alpha in this study was .95. The Omega 3 (Flora, 2020) composite reliability was also .95.

Shame-Experience. This variable was measured using a 5-item Shame subscale of the State Shame and Guilt Scale (SSGS; Marschall et al. (1994)). The participants were asked to reflect on how imagining themselves as the focal employee in the scenario they have just read makes them feel. They indicated their responses on a 5-point Likert type scale ranging from (1) not feeling this way at all to (5) feeling this way very strongly. The original version consisted of five items, with a sample item being: “I would want to sink into the floor and disappear.” The only revision I made

to the scale is adding “would” to each response item to make it more appropriate to assess emotions in a hypothetical situation. The Cronbach’s alpha in this study was .95. The Omega 3 (Flora, 2020) composite reliability was .96.

4.2.3.2 Mediating Variable

Work-Related Identity Threat. This variable was measured using a revised scale of identity threat from Kroeper et al. (2022). The original version consisted of six items, with a sample item being: “How much might you worry that you won’t belong at the company?” Participants were asked to indicate the amount of worry they might have regarding core aspects of their personal and social identity at work following a threat to their sense of self at work. They indicated their responses on a 5-point Likert type scale ranging from (1) not at all to (5) an extreme amount. My revision of their scale consisted of (1) taking out one question about feeling stereotyped as being irrelevant to the purpose of my study; and (2) adding three more questions that capture threat to moral identity in the broader context of work-related identity threat. These additional items were informed by the literatures on identity, work-related identity, morality, and integrity (Blasi, 1993; T. R. Cohen et al., 2014; Kurtines & Gewirtz, 1984; Selenko et al., 2018). The additional questions were discussed with the pilot group participants (all Ph.D. students in HRM/OB), none of whom had any concerns about these specific questions as a part of the work-related identity threat measure. The additional items are: “How much might you worry that the supervisor challenges your sense of moral integrity?”, “How much might you worry that the supervisor calls into question your morality?”, and “How much might you worry that the supervisor challenges your sense of self-worth?”. The Cronbach’s alpha in this study was .97. The Omega 3 (Flora, 2020) composite reliability was .96.

4.2.3.3 Moderating Variable

Moral Identity Centrality (MIC). This variable was measured using the internalization subscale of Aquino & Reed's (2002) self-importance of moral identity measure. Participants were given a list of characteristics that describe a person (i.e., caring, compassionate, fair, friendly, generous, helpful, hardworking, honest, and kind) and asked to visualize in their mind the kind of person who has these characteristics, imagining how that person would think, feel and act. The participants were then given five statements and asked to indicate their agreement using a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). A sample item includes: "It would make me feel good to be a person who has these characteristics." Items 3 and 4 were reverse-coded. After re-coding items 3 and 4 so that higher values reflect greater centrality of moral identity, the scores for the five items were averaged to calculate the moral identity centrality score for each participant. The Cronbach's alpha in this study was .76. The Omega 3 (Flora, 2020) composite reliability was .75.

4.2.3.4 Independent Variable

Interpersonal Felt Distrust (Experimental condition). Interpersonal felt distrust was experimentally manipulated by randomly assigning participants to read one of the two scenarios: Distrust or Neutral (see Appendix A). The text of the scenarios was guided by the literature on distruster behaviors (Bijlsma-Frankema et al., 2015; Guo et al., 2017; Lewicki et al., 1998; Sitkin & Bijlsma-Frankema, 2018; Sitkin & Stickel, 1996).

Interpersonal Felt Distrust (Manipulation Check). Following exposure to the scenario, survey participants were asked how distrusted they felt by the supervisor as a manipulation check for the experimental condition. Interpersonal felt distrust as a manipulation check was measured using a revised version of feeling of distrust scale from Mooijman et al. (2017). The original

version consisted of three items, with a sample item being: “I feel distrusted by the supervisor.” Participants were asked to indicate their agreement with each item on a 7-point Likert type scale ranging from (1) disagree completely to (7) agree completely.

I have revised the scale in three ways: (1) I changed the wording of the verb in the scale to read in simple past tense as opposed to present tense so as to aid the participants in relating the question to the situation they have read in the scenario; (2) I revised the reference to “the supervisor”; and (3) I have added one more item to increase the number of items in the scale, namely: “I thought the supervisor considers me distrustworthy.” Expanding the scale from three to four items strengthens its internal consistency and enhances construct validity. Research suggests that scales with a higher number of items generally produce more reliable measurements by reducing error and capturing a broader range of the intended construct (Onwuegbuzie & Daniel, 2002). A three-item scale may be susceptible to instability in reliability estimates, whereas four items allow for improved coherence among responses. The Cronbach’s alpha for the scale in this study was .97. The Omega 3 composite reliability (Flora, 2020) was also .97.

4.2.3.5 Control Variables

Research suggests that some personality factors, e.g., neuroticism, narcissism, anger and shame, influence an individual’s appraisal and reactivity to stressful events and experience of emotions (e.g., Bolger & Schilling, 1991; Bolger & Zuckerman, 1995; Coleman et al., 2019; Douglas et al., 2008; Garofalo et al., 2025; Kuppens & Tuerlinckx, 2007; Penley & Tomaka, 2002; Orth & Luciano, 2015; Schneider, 2004; Tong et al., 2006). Similarly, empirical evidence suggests that self-conscious emotional experiences differ across gender (Else-Quest et al., 2012), ethnicity (Durik et al., 2006), and age (Henry et al., 2018). Guided by this research, I anticipate that shame, anger, narcissism, neuroticism, sex, age, and race may exert influence on how individuals react to

events that trigger identity threat, and associated experiences of anger and shame. Therefore, in the survey, I statistically control these variables in the SEM and multivariate regression analyses to make sure that the regression coefficients more accurately capture the effects of interpersonal felt distrust and work-related identity threat on the variables of interest in the conceptual model.

Shame-Trait. This variable was measured with a 4-item Shame-Negative Self-Evaluation (NSE) subscale of the Guilt and Shame Proneness Scale (GASP) (T. R. Cohen et al., 2011). Participants were exposed to four situations and asked to indicate the likelihood that they would react in the way described on a 7-point Likert scale ranging from (1) very unlikely to (7) very likely. The sample items include: “You give a bad presentation at work. Afterwards your boss tells your coworkers it was your fault that your company lost the contract. What is the likelihood that you would feel incompetent?” The Cronbach’s alpha in this study was .73. The Omega 3 (Flora, 2020) composite reliability was also .73.

Anger-Trait. This variable was measured using a 10-item STAXI Trait Anger (T-Anger) Scale (Forgays et al., 1997) to measure individual differences in personal disposition to experience anger. The original 10 STAXI T-Anger scale consists of two dimensions: Angry-Temperament (T-Anger/T) and Angry-Reaction (T-Anger/R) (Spielberger, 1988; Spielberger et al., 1983). The T-Anger/T subscale assesses individual differences in one’s disposition to experience angry feelings without provocation, while the T-Anger/R subscale assesses how often a person feels angry when unfairly criticized or when they are treated unjustly or badly (Spielberger et al., 1988). The participants are instructed to rate themselves according to “how they generally feel” on a 4-point Likert scale, ranging from (1) almost never to (4) almost always. Two sample items from the scale are: “I have a fiery temperament” and “I feel infuriated when I do a good job and get poor

evaluation”. The Cronbach’s alpha in this study was .86. The Omega 3 (Flora, 2020) composite reliability was .73.

Neuroticism. This variable was measured using a 12-item neuroticism subscale of the Eysenck Personality Questionnaire–Brief Version (EPQ-BV) (Sato, 2005), which is a shortened and validated version of the Eysenck Personality Questionnaire–Revised (EPQ-R). The participants were instructed to answer the following questions on a 5-point Likert-type scale, indicating the frequency from (1) not at all to (5) extremely. Two sample items include: “Does your mood often go up and down?” and “Do you ever feel miserable for no reason?”. The Cronbach’s alpha in this study was .94. The Omega 3 (Flora, 2020) composite reliability was .93.

Narcissism. This variable was measured using the 16-item Narcissistic Personality Inventory-16 (NPI-16) (Ames et al., 2006) that was developed to measure subclinical narcissism. The NPI-16 is a shortened form of the NPI-40 (Raskin & Hall, 1979). The NPI-16 items are drawn from across the dimensions of the NPI-40. Scoring is performed by computing the proportion of responses consistent with narcissism, as suggested by the scale key. The participants are asked to read sixteen pairs of statements and then, for each pair, to select one statement that comes closest to describing their feelings and beliefs about themselves. A sample pair of statements is: “a. I really like to be the center of attention; b. It makes me uncomfortable to be the center of attention”. The Cronbach’s alpha in this study was .78. The Omega 3 (Flora, 2020) composite reliability was .86. Since the Narcissism scale consists of categorical indicators, Zumbo et al.’s (2007) ordinal alpha, which treats ordinal variables as numeric, was also calculated and was .88.

Sex. Participants’ sex has been operationalized as sex at birth (male or female). Past research provides evidence that women score significantly higher on shame-NSE than men (T. R. Cohen et al., 2011). Empirical research also suggests that a desire to physically express anger is a

more salient decision for women than for men (Forgays et al., 1997; Fuqua et al., 1991; van der Ploeg et al., 1982).

Age. This variable has been operationalized in groups according to the number entered by the respondent: “Please indicate your age (in number of full years, e.g. 26 or 35):18-25, 26-35, 36-45, 46-55, 56-65, over 65”. Age of participants may correlate with some of the variables measured in the model. For example, past research has found that older participants score significantly higher on shame-NSE than younger participants (T. R. Cohen et al., 2011).

Race. This variable has been measured as a categorical variable with a set of standard categories on Prolific for Ethnicity: White, Black, Asian, Mixed-race, Other and Prefer not to say. It is possible that representatives of racial/ethnic minorities may feel more vulnerable in the workplace and hence react to felt distrust by supervisor differently than their White counterparts.

4.3 Study 1 Data Analysis Strategy

After reviewing the data for cleanness, normality, and multivariate outliers, the data was analyzed in R in several steps. The initial step involved exploratory descriptive, correlational, and group-comparison analysis of all core variables in the model. The latent structure and reliability of measures for Identity Threat, Moral Identity Centrality, Anger-Experience, and Shame-Experience was explored with the confirmatory factor analysis (CFA). The reliability of each scale was assessed with Cronbach’s alpha and Omega 3 composite reliability (Flora, 2020).

Common method bias (Podsakoff et al., 2003) was tested following an assessment of the latent factor measurement model fit by comparing model fit indices across one- through four-factor confirmatory factor analysis (CFA) models. Fit was evaluated using AIC, BIC, CFI, TLI, RMSEA, and SRMR. The results are reported in section 4.4.8.2 of this dissertation. The results indicated

that models with four latent variables exhibited comparable fit, with lower AIC and BIC values when compared to simpler models.

The model hypotheses were tested through structural equation modeling (SEM) using *lavaan* in R. SEM offers a comprehensive solution to the analysis of experimental designs (Breitsohl, 2019). Its strength lies in correction for the measurement error in the model and robust estimations of individual path coefficients and overall model fits for models utilizing observed and latent variables (Kline, 2005; Meyers et al., 2017). For the purposes of statistical analysis, the core variables in the model (namely identity threat, moral identity centrality, anger-experience, shame-experience, and the interaction term between interpersonal felt distrust and moral identity centrality) were treated as latent. The control variables (anger-trait, shame-trait, narcissism, neuroticism, sex, age, and race) were treated as observed. A complementary analysis of hypotheses testing was performed through a series of linear regression models in R using formal mediation analysis with the *mediate* package. This allowed me to verify the results of SEM and provide a more nuanced understanding and visualization of moderation and mediation effects.

4.4 Study 1 Results

4.4.1 Assessing Data Quality

Survey participants' Q-captcha scores were above .8 overall and equal to 1 for 590 respondents (97% of the sample), which indicates that the survey was completed by humans and not by bots. Qualtrics' built-in Expert Review function showed that all respondents passed their system's quality check. Finally, all 610 participants passed the validation questions, both attention checks, and the seriousness check. Since all of the survey questions were enforced, meaning that the respondents could not move forward without providing an answer to a question, the sample dataset does not have missing data.

4.4.2 Assessing Data for Normality

4.4.2.1 Testing for Univariate Normality

The core variables were explored for normality and distribution patterns using visualization techniques such as histograms, q-q plots, scatterplots, and parallel coordinate plots. The visual inspection of each individual indicator of the latent variables suggested that the data was distributed consistently with the expected effects of the experimental conditions, meaning that Interpersonal Felt Distrust, Identity Threat, Anger-Experience, and Shame-Experience exhibited skewed distributions across the two conditions. This pattern aligns with the manipulation, as participants were exposed to two radically different scenarios (Neutral vs. Distrust). Consequently, the responses as these concern the outcome variables were expected to be (1) significantly different based on condition exposure; and (2) predominantly low in the Neutral condition, while predominantly high in the Distrust condition.

To formally assess univariate normality, the Shapiro-Wilk test, skewness, and kurtosis were computed for all indicator variables. The Shapiro-Wilk test yielded highly significant results ($p < .001$) for nearly all variables, indicating departures from normality. Skewness values ranged from -4.38 to 2.52, showing varying degrees of left- and right-skewed distributions. Kurtosis values ranged from 1.45 to 22.84, with several variables, such as Moral Identity Centrality and Anger-Trait, displaying heavy tails. Given these findings, the assumption of univariate normality was violated, supporting the need for robust estimation methods and non-parametric approaches in subsequent analyses.

4.4.2.2 Testing for Multivariate Normality

To assess multivariate normality, Mardia's test was conducted, examining multivariate skewness and kurtosis across all indicator variables. Results revealed significant deviations from

normality, with Mardia's skewness (Statistic = 86,282.90, $p < .001$) and Mardia's kurtosis (Statistic = 88.37, $p < .001$) indicating non-normal multivariate distributions. This further confirmed the choice of robust estimation methods and bootstrapping approaches in subsequent statistical analyses.

4.4.3 Multivariate Outlier Detection and Management

Following best practice recommendations on exploring the dataset for potential outliers in the organizational science literature (Aguinis et al., 2013), I used two approaches to identify multivariate outliers in my dataset. First, I calculated the Mahalanobis distance on the core numeric variables (i.e., Felt Distrust (continuous variable used as a manipulation check), Identity Threat, Moral Identity Centrality, Anger-Expression, and Shame-Expression), using the covariance matrix of these columns. A chi-square test ($p < 0.001$) was applied to determine the threshold for outliers. The chi-square critical values for the lower and upper tails were 0.210 and 20.515, respectively. Observations with Mahalanobis distances exceeding this threshold were identified as potential outliers. This process resulted in the identification of 16 observations as potential outliers in the Distrust condition (constituting 5% of the data) and 8 observations as potential outliers in the Neutral condition (constituting less than 3% of the data) for a total of 24 observations.

As a second approach, I fit a multivariate regression model for the Neutral and Distrust condition datasets and calculated the Cook's distance to further quantify the influence of each observation. Cook's distance provides a reliable measure of how much an individual data point impacts the overall regression model by taking into account the changes in the fitted values when that observation is excluded. Following the recommendations of Aguinis et al. (2013), I used the F distribution with $df = (k + 1, n - k - 1)$ and $\alpha = .5$ (J. Cohen et al., 2003) to derive a threshold from the F distribution and compare the Cook's distance values to the critical threshold. As a result,

no significant outliers were identified beyond the Cook's D threshold. Such a combination of several methods to assess multivariate outliers provides a solid understanding of the data. Since both methods identified less than 5% of the dataset as potential multivariate outliers, it is considered safe to assume multivariate normality and apply the statistical tests of regression and SEM to test the model of Study 1.

4.4.4 Descriptives

Table 1 presents the means, standard deviations, and additional descriptive information for the categorical and continuous variables of Study 1. To compare the means and medians of these variables across the two experimental conditions, the descriptives are also presented for the same variables grouped by experimental conditions, being Neutral (Table 2) and Distrust (Table 3)². As visible in the tables, across the full dataset ($n = 610$), Interpersonal Felt Distrust and Work-Related Identity Threat show moderate dispersion, while Moral Identity Centrality and Shame-Trait reflect more stable individual differences with a narrower spread.

The proportional distribution of sex, age, and race is representative of the population of the United States. The demographics distribution of the study sample included 350 female (57.38%) and 260 male (42.62%) respondents, a median age of 36 years (with the youngest respondent being 20 years of age and the oldest being 74 years of age), and a race distribution of 413 White (67.70%), 90 Black (14.75%), 61 Asian (10%), 31 Mixed-race (5%), 14 Other (2.3%), and 1 Prefer not to say (.0016%) participants. Five hundred and eight respondents (83.28%) were employed full-time, 99 (16.23%) were working part-time, and 1 (.0016%) were due to start a new job within the next month. Two participants (.0033%) did not respond to the question about employment

² No participants were excluded based on manipulation check scores for interpersonal felt distrust. This decision reflects the study's aim to preserve natural variability in subjective responses to the vignette and acknowledges that interpersonal experience of felt distrust may not be uniformly triggered in all individuals. While this approach may introduce noise, it enhances ecological validity and avoids artificially restricting the range of the dependent variable.

status; however, they passed the validation check for employment status based on their indication of employment status on the Prolific member portal.

Table 1

Study 1. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables

#	Variable	Mean	Standard Deviation (sd)	Median	Minimum	Maximum	Range	Standard Error (se)
	Interpersonal Felt							
1	Distrust	4.06	2.32	4.75	1.00	7.00	6.00	0.09
	Work-Related							
2	Identity Threat	2.73	1.34	2.88	1.00	5.00	4.00	0.05
3	Anger-Experience	2.25	1.20	2.00	1.00	5.00	4.00	0.05
4	Shame-Experience	2.34	1.33	2.00	1.00	5.00	4.00	0.05
	Moral Identity							
5	Centrality (MIC)	4.65	0.53	4.80	1.20	5.00	3.80	0.02
6	Shame-Trait	4.37	0.65	4.50	1.00	5.00	4.00	0.03
7	Anger-Trait	1.77	0.52	1.70	1.00	4.00	3.00	0.02
8	Neuroticism	2.43	0.93	2.33	1.00	5.00	4.00	0.04
9	Narcissism	0.20	0.19	0.19	0.00	0.88	0.88	0.01
10	Age	37.54	11.22	36.00	20.00	74.00	54.00	0.45

Note: n = 610

As evident in Tables 2 and 3, Work-Related Identity Threat is much more pronounced in responses to the Distrust Scenario compared to the Neutral Scenario, reinforcing the anticipated harmful effect of interpersonal felt distrust on one's work-related identity. Likewise, participants in the Distrust condition report significantly higher levels of Anger-Experience and Shame-Experience when compared to those in the Neutral condition, please see tables 2 and 3.

Table 2

Study 1. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables for the Neutral Condition

#	Variable	Mean	Standard Deviation (sd)	Median	Minimum	Maximum	Range	Standard Error (se)
1	Interpersonal Felt Distrust	2.10	1.40	1.75	1.00	7.00	6.00	0.08
2	Work-Related Identity Threat	1.65	0.81	1.38	1.00	5.00	4.00	0.05
3	Anger-Experience	1.41	0.74	1.00	1.00	5.00	4.00	0.04
4	Shame-Experience	1.42	0.75	1.00	1.00	5.00	4.00	0.04
5	Moral Identity Centrality (MIC)	4.65	0.50	4.80	2.20	5.00	2.80	0.03
6	Shame-Trait	4.40	0.65	4.50	1.50	5.00	3.50	0.04
7	Anger-Trait	1.74	0.52	1.70	1.00	4.00	3.00	0.03
8	Neuroticism	2.38	0.91	2.25	1.00	5.00	4.00	0.05
9	Narcissism	0.20	0.19	0.12	0.00	0.88	0.88	0.01
10	Age	37.09	10.85	36.00	20.00	69.00	49.00	0.62

Note: n = 303

Table 3

Study 1. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables for the Distrust Condition

#	Variable	Mean	Standard Deviation (sd)	Median	Minimum	Maximum	Range	Standard Error (se)
	Interpersonal							
1	Felt Distrust	6.00	1.06	6.25	1.00	7.00	6.00	0.06
	Work-Related							
2	Identity Threat	3.80	0.79	3.88	1.00	5.00	4.00	0.05
	Anger-							
3	Experience	3.08	0.96	3.17	1.00	5.00	4.00	0.05
	Shame-							
4	Experience	3.25	1.13	3.40	1.00	5.00	4.00	0.06
	Moral Identity							
5	Centrality (MIC)	4.65	0.56	4.80	1.20	5.00	3.80	0.03
6	Shame-Trait	4.34	0.65	4.50	1.00	5.00	4.00	0.04
7	Anger-Trait	1.80	0.52	1.70	1.00	4.00	3.00	0.03
8	Neuroticism	2.47	0.94	2.42	1.00	4.75	3.75	0.05
9	Narcissism	0.21	0.19	0.19	0.00	0.88	0.88	0.01
10	Age	37.99	11.57	36.00	20.00	74.00	54.00	0.66

Note: n = 307

4.4.5 Validity and Reliability of Measures

To ensure the construct validity of the measurement scales used in the survey and verify their reliability, I performed confirmatory factor analysis (CFA) for each of the core constructs in the model with *lavaan* package in R (Maximum Likelihood Robust (MLR) estimator). The MLR estimator provides robust standard errors and a scaled test statistic that are adjusted for non-normality in the data. This method is especially useful when dealing with data that may not meet the assumption of multivariate normality, as it helps to produce more reliable and accurate parameter estimates and fit indices. For transparency, I reported standard, scaled, and robust estimates of these model fit indices. The results of the CFA are presented in Appendix B, Table B1.

For the *Interpersonal Felt Distrust* (manipulation check) variable, the CFA model demonstrated acceptable fit, $\chi^2(2) = 12.30$, $p = .002$, using the Yuan-Bentler scaled correction. Global fit indices supported this conclusion: CFI values were high across estimators (Standard = .995; Scaled = .988; Robust = .995), as were TLI values (Standard = .984; Scaled = .964; Robust = .985). The SRMR was low (.005 for both Standard and Scaled), indicating minimal residuals. Although the standard RMSEA (.124) and robust RMSEA (.120) exceeded conventional thresholds, the scaled RMSEA (.092) approached acceptable fit ($< .08$), suggesting modest misfit. The standardized loadings for the Interpersonal Felt Distrust (manipulation check) variable are all strong and significant, suggesting that each indicator contributes substantially to the overall construct. All of the estimated reliability indices were above their respected thresholds (i.e., Cronbach's alpha ($\alpha = 0.97$), Omega 3 Composite Reliability ($\omega_3 = 0.97$), Composite Reliability (CR = 0.97) and Average Variance Extracted (AVE = 0.90)), which indicates excellent construct reliability and validity.

For the *Identity Threat* variable, the CFA output indicates that the model fits well overall. The chi-square test is non-significant for the scaled estimate, $\chi^2(12) = 18.291$, $p = 0.107$. The global fit indices also suggest excellent fit: CFI (Standard and Scaled: 0.998; Robust: 0.999), TLI (Standard: 0.994; Scaled: 0.995; Robust: 0.997), RMSEA (Standard: 0.046; Scaled: 0.029; Robust: 0.036), and SRMR (0.009 for both Standard and Scaled). The standardized loadings for the Identity Threat variable indicate strong relationships with its observed variables with standardized loadings ranging above 0.79, all being significant. These high values suggest that the indicators are reliable measures of the latent variable Identity Threat. All estimated reliability indices were above their respected thresholds (i.e., Cronbach's alpha ($\alpha = 0.97$), Omega 3 Composite Reliability ($\omega_3 = 0.96$),

Composite Reliability (CR = 0.97), and Average Variance Extracted (AVE = 0.80)), which indicates excellent construct reliability and validity.

For the *Moral Identity Centrality* variable, the CFA output indicates generally good fit indices and reliability, with some areas for potential model improvement. Although the chi-square test is significant, $\chi^2(5) = 22.635, p < 0.001$, the model fit indices suggest good fit: CFI (Standard: 0.954; Scaled: 0.931; Robust: 0.960), TLI (Standard: 0.909; Scaled: 0.862; Robust: 0.921), RMSEA (Standard: 0.111; Scaled: 0.076; Robust: 0.103), and SRMR (0.038 both Standard and Scaled). The standardized loadings for the MIC latent variable are all significant. They suggest that MIC_2 and MIC_5 have the strongest relationships with the latent variable, while MIC_3 has the weakest. Most of the estimated reliability indices were above their respected thresholds (i.e., Cronbach's alpha ($\alpha = 0.76$), Omega 3 Composite Reliability ($\omega_3 = 0.75$), Composite Reliability (CR = 0.78), and Average Variance Extracted (AVE = 0.40)), which indicates reasonable construct reliability and validity. Overall, the loadings indicate that the observed variables are reliable measures of the latent construct.

For the *Anger-Experience* variable, the CFA output indicates that the model fits well overall, even though the chi-square test is significant, $\chi^2(9) = 85.063, p = < 0.001$, the model fit indices suggest overall good fit: CFI (Standard: 0.960; Scaled: 0.959; Robust: 0.962), TLI (Standard: 0.934; Scaled: 0.931; Robust: 0.937), RMSEA (Standard: 0.164; Scaled: 0.118; Robust: 0.160), and SRMR (0.022 for both Standard and Scaled). The standardized loadings for the Anger-Experience variable indicate strong relationships with its observed variables, with standardized loadings ranging above 0.840, all being significant. These high values suggest that the indicators are reliable measures of the latent variable Identity Threat. All estimated reliability indices were above their respected thresholds (i.e., Cronbach's alpha ($\alpha = 0.95$), Omega 3 Composite Reliability

($\omega_3 = 0.95$), Composite Reliability (CR = 0.95), and Average Variance Extracted (AVE = 0.77)), which indicates excellent construct reliability and validity.

For the *Shame-Experience* variable, the CFA output indicates that the model fits very well overall. While the chi-square test is significant, $\chi^2(5) = 15.197$, $p = 0.010$, the model fit indices suggest excellent fit: CFI (Standard and Scaled: 0.993; Robust: 0.994), TLI (Standard: 0.986; Scaled: 0.987; Robust: 0.989), RMSEA (Standard: 0.085; Scaled: 0.058; Robust: 0.077), and SRMR (0.008 for both Standard and Scaled). The standardized loadings for the Shame-Experience variable indicate strong relationships with its observed variables, with standardized loadings ranging above 0.840, all being significant. These high values suggest that the indicators are reliable measures of the latent variable Identity Threat. All estimated reliability indices were above their respected thresholds (i.e., Cronbach's alpha ($\alpha = 0.95$), Omega 3 Composite Reliability ($\omega_3 = 0.96$), Composite Reliability (CR = 0.96), and Average Variance Extracted (AVE = 0.81)), which indicates excellent construct reliability and validity.

The CFAs for control variables (i.e. anger-trait, shame-trait, narcissism, and neuroticism) also showed good model fit and reasonable estimates of reliability. The results for the standardized factor loadings and measures of reliability are available in the Appendix B, Table B2. Unlike other controls, narcissism (as a latent variable) is composed of categorical items. Therefore, an additional version of coefficient alpha has been calculated for narcissism and is included in the notes section, i.e., ordinal alpha (Zumbo et al., 2007; Zumbo & Kroc, 2019). Likewise, the average variance extracted (AVE) for narcissism has been calculated from polychoric (polyserial) correlation.

4.4.6 Correlational analysis

Table 4 provides a summary of correlations between the continuous variables in the model. The entries on the diagonal line are the Cronbach alpha values. As expected, the data shows that

there exists a high and statistically significant correlation between the core variables in the model. A correlation between the continuous measure of interpersonal felt distrust (manipulation check questions) and work-related identity threat indicates a strong positive relationship ($r = .89, p < .001$), which. The work-related identity threat has a strong and positive correlation with anger-experience ($r = .85, p < .001$) and shame-experience ($r = .82, p < .001$). Notably, there is also a strong and positive correlation between interpersonal felt distrust (manipulation check variable) and anger-experience ($r = .79, p < .001$) and shame-experience ($r = .75, p < .001$). There is a strong positive correlation between the two outcome variables of anger-experience and shame-experience ($r = .80, p < .001$). The correlation of moral identity centrality with the core variables of the model, being interpersonal felt distrust, work-related identity threat, anger-experience and shame-experience, is very weak (r ranging from .01 to .03) and statistically non-significant.

Table 4

Study 1. Intercorrelations of Continuous Variables

Variable	1	2	3	4	5	6	7	8	9
1. Interpersonal Felt Distrust	(.97)								
2. Work-Related Identity Threat	.89***	(.97)							
3. Anger-Experience	.79***	.85***	(.95)						
4. Shame-Experience	.75***	.82***	.80***	(.95)					
5. Moral Identity Centrality	.03	.03	-.01	-.01	(.76)				
6. Shame-Trait	-.02	.04	.04	.16***	.37***	(.73)			
7. Anger-Trait	.06	.15***	.26***	.20***	.20***	-.03	(.86)		
8. Neuroticism	.09*	.16***	.21***	.30***	-.11**	.14***	.57***	(.94)	
9. Narcissism	0	-.02	.05	-.03	-.16**	.27***	.23***	-.08	(.78)

Note. Study 1. $n = 610$. The measure of the Manipulation Check variable serves as a proxy for Interpersonal Felt Distrust. Entries on the main diagonal are Cronbach's alpha. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The p -value for the correlation between Neuroticism and Narcissism was .051

From the four dispositional traits examined as control variables (i.e., anger-trait, shame-trait, narcissism, and neuroticism), neuroticism is the only control variable that has weak to moderately statistically significant correlations with all of the variables in the model. Correlations of all control variables with moral identity centrality are very weak ($r < .1$) and statistically non-significant, except for a statistically significant, weak, and positive correlation between moral identity centrality and neuroticism ($r = .09, p < .05$).

Generally, all of the correlations were in expected directions and magnitude of association. It was interesting, however, that shame-trait did not have a meaningful correlation with work-related identity threat and that narcissism did not have meaningful correlations with work-related identity threat, anger-experience, or shame-experience. Moreover, it was interesting to find out that the moderator (i.e., moral identity centrality (MIC)) had very weak and non-significant correlations with the core variables in the model (i.e., interpersonal felt distrust (manipulation check), identity threat, anger-experience, and shame-experience). At the same time, MIC had weak/moderate/moderate and statistically significant correlations with all of the control variables (i.e., shame-trait being negatively correlated with MIC; and anger-trait, narcissism, and neuroticism being positively correlated with MIC).

4.4.7 Group Comparison Analyses

After having explored the descriptive statistics and correlations of the variables, I ran a series of Mann-Whitney U tests (Wilcoxon rank-sum tests) as a part of an exploratory analysis to assess whether there is a significant difference between the means of the two groups for each of the core variables in the model. The choice of this non-parametric test is appropriate given observed violations of univariate normality assumptions. Additionally, I used Rank-Biserial Correlation to estimate the effect size, as it is a suitable measure for non-parametric comparisons,

capturing the strength of the difference between groups. As in previous analyses, I used the Neutral condition (control group) as the reference to examine the impact of the Distrust condition (treatment group). The Mann-Whitney U test served as an exploratory step to identify significant differences between the groups and understand the direction of effects.

To assess the effectiveness of experimental manipulation, I compared Interpersonal Felt Distrust (manipulation check) across the two conditions using the Mann-Whitney U test. Results indicated a statistically significant difference between the groups ($W = 3292.5$, $p < .001$), demonstrating that participants in the Distrust condition reported substantially higher levels of Interpersonal Felt Distrust compared to those in the Neutral condition. The Rank-Biserial Correlation ($r = -0.93$, 95% CI $[-0.94, -0.92]$) confirms a very large effect size, suggesting a clear and substantial impact of the experimental manipulation on participants. These findings support the expected pattern wherein the Neutral condition produced lower distrust levels while the Distrust condition resulted in markedly increased distrust. An earlier test of the Mahalanobis distance identified 24 observations as potential outliers, so I ran the Mann-Whitney U t-test on a cleaned dataset ($n = 586$). The test statistic and effect size were identical. The consistency between the full dataset and cleaned dataset results suggests the manipulation effect is robust and not driven by extreme values.

Next, I conducted the Mann-Whitney U t-test for Identity Threat, Anger Experience, Shame Experience, and Moral Identity Centrality to detect significant group differences for each of the core variables in the model. Testing the difference of Identity Threat by experimental conditions (mean in group Distrust = 3.80, mean in group Neutral = 1.65) suggests that the effect is positive, statistically significant, and large (difference = 2.14, 95% CI $[2.02, 2.27]$, $W = 88,513$, $p < .001$; Rank-Biserial Correlation = 0.90, 95% CI $[0.88, 0.92]$). This means that survey participants who

read the Distrust scenario experienced a much higher level of Identity Threat than survey participants who were exposed to the Neutral scenario. Moreover, this difference is statistically significant. The results are robust across the original and cleaned datasets, with a similarly positive, statistically significant, and large effect in the cleaned dataset (difference = 2.11, $W = 81,440$, $p < .001$; Rank-Biserial Correlation $r = 0.90$, 95% CI [0.88, 0.91]).

The Mann-Whitney U test, which examined the difference across experimental conditions in the two outcome variables, namely Anger-Experience and Shame-Experience, indicates that survey participants in the treatment group (the Distrust condition) experienced more anger and shame compared to survey participants in the control group (the Neutral condition). Specifically, testing of Anger-Experience by experimental conditions (mean in group Distrust = 3.08, mean in group Neutral = 1.41) indicates that the effect is positive, statistically significant, and large (difference = 1.68, 95% CI [1.54, 1.81], $W = 84,989$, $p < .001$; Rank-Biserial Correlation $r = 0.83$, 95% CI [0.80, 0.85]). The effect was very similar in the cleaned dataset (difference = 1.63, 95% CI [1.49, 1.77], $W = 77,974$, $p < .001$; Rank-Biserial Correlation $r = 0.82$, 95% CI [0.78, 0.85]). Likewise, the difference in Shame-Experience across experimental conditions (mean in group Distrust = 3.25, mean in group Neutral = 1.42) suggests that the effect is positive, statistically significant, and large (difference = 1.84, 95% CI [1.68, 1.99], $W = 83,942$, $p < .001$; Rank-Biserial Correlation $r = 0.80$, 95% CI [0.77, 0.83]), with a very similar effect in the cleaned dataset (difference = 1.79, 95% CI [1.64, 1.95], $W = 76,993$, $p < .001$; Rank-Biserial Correlation $r = 0.79$, 95% CI [0.76, 0.83]). These results are expected due to the differential effect of the experimental conditions for the outcome variables.

Importantly, the moderating variable of Moral Identity Centrality (MIC), which was measured in both groups toward the end of the survey, did not yield a statistically significant

difference between the two experimental groups. This is in line with expectations because Moral Identity Centrality is essentially a personal characteristic, which should not have been influenced by the experimental group to which the survey participants were assigned. The difference of MIC across experimental conditions (mean in group Distrust = 4.648, mean in group Neutral = 4.652) suggests that the effect is negative, statistically not significant, and very small (difference = $-3.29\text{e-}03$, 95% CI $[-0.09, 0.08]$, $W = 47,584$, $p = 0.5997$; Rank-Biserial Correlation $r = 0.02$, 95% CI $[-0.07, 0.11]$). The effect was very similar in the cleaned dataset: negative, statistically not significant, and very small (difference = $-9.60\text{e-}03$, 95% CI $[-0.09, 0.08]$, $W = 43,395$, $p = 0.8066$; Rank-Biserial Correlation $r = 0.01$, 95% CI $[-0.08, 0.10]$). This confirms that on average, survey participants were no different in their MIC across the two experimental groups.

4.4.8 Structural Equation Modeling (SEM) Analysis

SEM is a powerful statistical technique that allows researchers to analyze complex relationships between variables, including moderation and mediation, by modeling their effects simultaneously (Kline, 2016; J. Lin, n.d.; Meyers et al., 2017). It allows for an understanding of the underlying structure of the data, to estimate the relationships between latent and observed variables in a single model, and to handle measurement errors. It has been recognized as an appropriate analysis method for experimental designs in organizational research (Breitsohl, 2019). I conducted SEM using the *lavaan* package in R using the Maximum Likelihood (ML) estimator with bias-corrected bootstrapping with 5000 resamples to obtain robust parameter estimates and confidence intervals, following recommended practices for handling non-normality in SEM models (Efron & Tibshirani, 1993). Through the SEM structural model, I tested the following hypotheses:

- 1) Hypothesis 1 – The effect of Interpersonal Felt Distrust (binary: Distrust Scenario and Neutral Scenario) on Identity Threat (latent);
- 2) Hypothesis 2 – The effect of the double-mean centered interaction between Moral Identity Centrality (MIC, latent) and Interpersonal Felt Distrust (binary) on Identity Threat (latent), while controlling for Sex (binary: male, female), Age (continuous), Race (binary: White, Non-White);
- 3) Hypothesis 3 – The mediating effect of Identity Threat (latent) on Anger-Experience (latent), while controlling for Anger Trait (observed), Shame Trait (observed), Neuroticism (observed), and Narcissism (observed); and
- 4) Hypothesis 4 – The mediating effect of Identity Threat (latent) on Shame-Experience (latent), while controlling for Anger Trait (observed), Shame Trait (observed), Neuroticism (observed), and Narcissism (observed).

4.4.8.1 Assessing the Measurement Model

As the first step in the Structural Equation Modeling (SEM) analysis, a measurement model was assessed using Confirmatory Factor Analysis (CFA) in *lavaan*, focusing on four latent variables: Work-Related Identity Threat (IDthreat.lv), Moral Identity Centrality (MIC.lv), Anger-Experience (Angerexp.lv), and Shame- Experience (Shameexp.lv). The latent interaction term was excluded, as interactions can bias model fit estimates (Regorz, 2022). The model was estimated using robust maximum likelihood (MLR), which accounts for non-normality in data. Fit indices were evaluated to determine model adequacy, including CFI, TLI, RMSEA, and SRMR, ensuring that the factor structure appropriately represents the data. Results indicate a significant chi-square test ($\chi^2(246) = 946.53, p < 0.001$); however, its sensitivity to sample size ($n = 610$) limits interpretability. Model fit indices indicate strong model fit, with CFI = 0.940, TLI = 0.932, RMSEA

= 0.080, 90% CI =[0.075, 0.086], SRMR = 0.039. These results suggest that the specified factor structure adequately represents the data, supporting the validity of the measurement model.

4.4.8.2 Testing for Common Method Variance

Based on the Confirmatory Factor Analysis (CFA) results, the model was assessed for the presence of common method bias (Podsakoff et al., 2003). Measurement models ranging from a single-factor solution to a four-factor model, which represents the conceptual structure, were compared. Models were estimated using MLR (Robust) estimator in R with *lavaan*, and fit was assessed using CFI, TLI, RMSEA, SRMR, AIC, and BIC indices.

Results indicate that the four-factor model provided the best fit, aligning with theoretical expectations (CFI = 0.934, TLI = 0.926, RMSEA = 0.084, 90% CI [0.080, 0.089], SRMR = 0.039, AIC = 32468.06, BIC = 32706.39). The three-factor model demonstrated reasonably strong fit (CFI = 0.882, TLI = 0.869, RMSEA = 0.112, SRMR = 0.046, AIC = 33310.39, BIC = 33535.47), though lower than the four-factor model. In contrast, the two-factor model exhibited even weaker fit (CFI = 0.833, TLI = 0.816, RMSEA = 0.133, SRMR = 0.088, AIC = 34105.17, BIC = 34321.43). The one-factor model performed the worst (CFI = 0.787, TLI = 0.767, RMSEA = 0.150, SRMR = 0.091, AIC = 34845.80, BIC = 35057.65), indicating that a simplified structure fails to adequately represent the data. The poor fit of the single-factor model supports the absence of common method variance, reinforcing the need to treat the four latent constructs as distinct factors. The four-factor model was identified as the optimal solution.

4.4.8.3 Double-Mean Centering the Latent Interaction (Moderation)

As the next step, to account for interaction effects within the Structural Equation Modeling (SEM) framework, the moderator Moral Identity Centrality (MIC) was mean-centered to reduce

multicollinearity (G. C. Lin et al., 2010). This involved centering each observed Moral Identity Centrality (MIC) indicator before creating interaction terms by subtracting the mean of each MIC indicator variable from its corresponding observations. This process ensures that each MIC indicator had a mean of zero, reducing multicollinearity when interacting with the categorical independent variable Interpersonal Felt Distrust. Interpersonal Felt Distrust, operationalized through experimental conditions (Neutral and Distrust) in the Scenario variable, then interacted with the centered MIC indicators to generate the latent interaction terms. Interaction terms were computed using *model.matrix()* in R, ensuring that the correct reference group was maintained. Following established methodological recommendations (G. C. Lin et al., 2010; Regorz, 2022), a second mean-centering procedure was conducted on the resulting interaction terms to ensure proper variance structure and numerical stability in SEM estimation. This double mean-centering approach minimizes multicollinearity, which can inflate standard errors and distort parameter estimates. By ensuring that both the predictor and interaction term are appropriately processed before their inclusion in the structural model, this method enhances the interpretability and the reliability of statistical inference.

4.4.8.4 Assessing Multicollinearity

Since bivariate correlations among core variables were high (0.75 to 0.89), it was important to address potential multicollinearity before running the full SEM model. To mitigate collinearity effects, double-mean centering was applied to the interaction term, ensuring model stability and improving parameter estimation. After centering, Variance Inflation Factor (VIF) scores were computed for each regression model to assess collinearity among predictors. For Identity Threat, VIF values ranged from 1.004 to 2.325, indicating moderate but acceptable collinearity. Similarly, for Anger Expression, VIF scores ranged from 1.052 to 2.958; while for Shame Expression, they

ranged from 1.062 to 2.962—none exceeding the conventional threshold of 5 (Hair et al., 2010). Given these results, the predictors were deemed suitable for inclusion in the SEM model, with minimal risk of inflated standard errors or unreliable estimates.

4.4.8.5 Hypotheses Testing

Structural equation modeling (SEM) was conducted using *lavaan* in R. Since multicollinearity was not detected and double-mean centering was applied to the interaction term prior to estimation, it was deemed appropriate to estimate the model using Maximum Likelihood (ML), a standard estimation method that efficiently handles latent variables and provides unbiased parameter estimates when multivariate normality assumptions are met (Hair et al., 2010). Additionally, bias-corrected bootstrapping with 5000 resamples was applied to obtain robust parameter estimates and confidence intervals, following recommended practices for handling non-normality in SEM models (Efron & Tibshirani, 1993). This approach improves statistical inference by generating dataset-based standard errors, reducing dependence on theoretical distributional assumptions, and enhancing the reliability of confidence interval estimation. After estimating the model, the residual covariance matrix was examined for normality. The results confirmed that residuals followed an approximately normal distribution, supporting the multivariate normality assumption required for ML estimation to provide valid results. The full structural model included five latent constructs—Work-Related Identity Threat; Moral Identity Centrality (MIC); the interaction between Interpersonal Felt Distrust and MIC, Anger-Experience, and Shame-Experience;; the observed control variables (Anger-Trait, Shame-Trait, Narcissism, Neuroticism, Sex, Age, and Race); and the covariation of individual indicators of the double-mean centered latent variable.

The chi-square test ($\chi^2(575) = 3865.31, p < 0.001$) was significant, though its sensitivity to sample size ($n = 610$) limits interpretability. The full model fit indices (CFI = 0.841, TLI = 0.823, RMSEA = 0.097, 90% CI [0.094, 0.100], SRMR = 0.066) showed a decline compared to the measurement model. However, these values should not be interpreted as indicators of overall model fit, as the inclusion of a latent interaction term introduces bias to model evaluation (Regorz, 2022). The measurement model previously demonstrated excellent fit (CFI = 0.940, TLI = 0.932, RMSEA = 0.080, 90% CI [0.075, 0.086], SRMR = 0.039), confirming that the latent factor structure was robust. The lower fit observed here reflects added complexity from the added interaction term and error covariances, which unduly increased model constraints.

Next, I present the unstandardized estimates (beta b) and their bootstrapped confidence intervals, the corresponding p -values of the effects, and the standardized estimates (beta β) in the SEM model. Examining the regression paths in the SEM, there is a significant positive effect of Interpersonal Felt Distrust on Work-Related Identity Threat ($b = 2.210$, 90% CI [2.074, 2.341], $p < 0.001$, $\beta = 0.819$), indicating that interpersonal felt distrust by supervisor leads to increased perceived identity threat (**Hypothesis 1 was supported**). The effect of Moral Identity Centrality (MIC) on Work-Related Identity Threat is not significant ($b = -0.080$, 90% CI [-0.283, 0.170], $p = 0.487$, $\beta = -0.024$), suggesting that levels of MIC do not influence work-related identity threat directly. However, the interaction of Interpersonal Felt Distrust and MIC significantly affects the level of Work-Related Identity Threat ($b = 0.282$, 90% CI [0.134, 0.400], $p < 0.001$, $\beta = 0.084$), indicating that MIC moderates the relationship between interpersonal felt distrust by supervisors and work-related identity threat (**Hypothesis 2 was supported**). Work-Related Identity Threat is strongly associated with Anger-Experience ($b = 0.835$, 95% CI [0.75, 0.92], $p < 0.001$, $\beta = 0.86$) and Shame-Experience ($b = 0.71$, 90% CI [0.616, 0.795], $p < 0.001$, $\beta = 0.765$), suggesting that

experiencing identity threat at work is associated with intense emotional reactions of anger (**Hypothesis 3a was supported**) and shame (**Hypothesis 4a was supported**). The direct effect of Interpersonal Felt Distrust on Anger-Experience is not significant ($b = 0.048$, 90% CI [-0.174, 0.270], $p = 0.675$, $\beta = 0.018$), which suggests that Identity Threat fully mediates (Meyers et al., 2017) the relationship between Interpersonal Felt Distrust and Anger-Experience (**Hypothesis 3b was supported**). The direct effect of Interpersonal Felt Distrust on Shame-Experience ($b = 0.243$, 90% CI [0.021, 0.461], $p = 0.031$, $\beta = 0.098$) is moderate and statistically significant, which suggests partial mediation (Meyers et al., 2017) (**Hypothesis 4b was supported**).

Among personality-based control variables, Anger-Trait is significantly associated with Anger Experience ($b = 0.309$, 90% CI [0.175, 0.443], $p < 0.001$, $\beta = 0.123$), and Shame-Trait is significantly associated with Shame Experience ($b = 0.203$, 90% CI [0.099, 0.308], $p < 0.001$, $\beta = 0.106$). Neuroticism is significantly associated with Shame Experience ($b = 0.214$, 90% CI [0.145, 0.284], $p < 0.001$, $\beta = 0.159$), while Narcissism did not show any significant effects. The effects of Sex, Age, and Race on the predictors were not significant, which indicates that these demographic factors do not contribute to differences in work-related identity threat, anger, or shame following interpersonal felt distrust by a supervisor at work.

Latent variable variances reported in the SEM output confirm meaningful differentiation of constructs, with Identity Threat (0.588), Anger-Experience (0.329), and Shame-Experience (0.347) showing substantial variance. Even though the latent interaction term's variance (0.161) was much smaller, it was still significant, thereby confirming its moderating effect. Standard errors were properly estimated, ensuring statistical reliability. Table 5 presents the unstandardized and standardized estimates, standard errors, the bias-corrected confidence intervals, and the p-values

of the SEM for Study 1. Figure 3 presents the conceptual model with SEM results (coefficients and significance).

Table 5

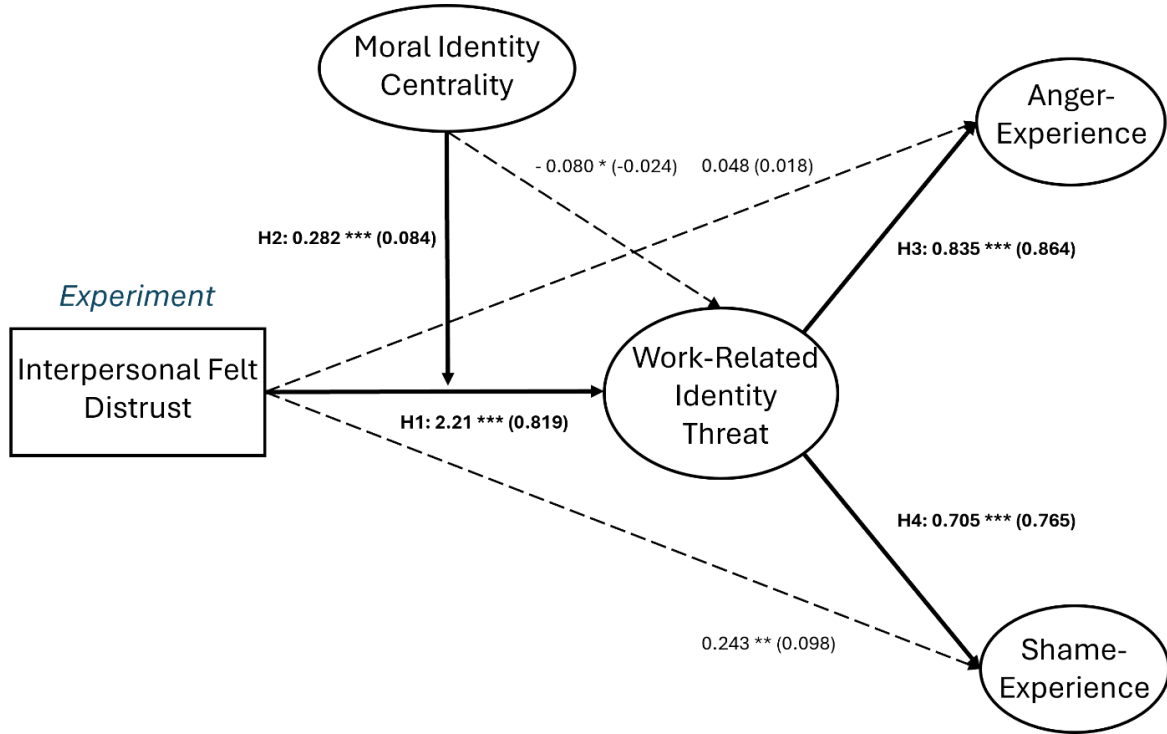
Study 1. SEM: Regression Path Estimates (Unstandardized and Standardized), Confidence Intervals for Unstandardized Coefficients, Standard Errors and P-Values

Path	Estimate (Unstandardized)	CI (LL)	CI (UL)	Std. Error	Estimate (Standardized)	p-value
<i>Predicting Identity Threat</i>						
IFD → Identity Threat	2.210	2.074	2.341	0.069	0.819	< 0.001
MIC → Identity Threat	-0.080	-0.283	0.170	0.115	-0.024	0.487
IFD x MIC → Identity Threat	0.282	0.134	0.400	0.069	0.084	< 0.001
Sex → Identity Threat	-0.068	-0.200	0.057	0.066	-0.025	0.303
Age → Identity Threat	-0.004	-0.010	0.003	0.003	-0.030	0.268
Race → Identity Threat	-0.020	-0.170	0.130	0.077	-0.007	0.799
<i>Predicting Anger Experience</i>						
IFD → Anger-Experience	0.048	-0.174	0.270	0.113	0.018	0.675
Identity Threat → Anger-Experience	0.835	0.751	0.918	0.042	0.864	< 0.001
Anger Trait → Anger-Experience	0.309	0.175	0.443	0.068	0.123	< 0.001
Neuroticism → Anger-Experience	0.024	-0.053	0.104	0.040	0.017	0.544
Narcissism → Anger-Experience	0.198	-0.070	0.468	0.137	0.029	0.150
Sex → Anger-Experience	0.069	-0.042	0.180	0.057	0.026	0.223
Age → Anger-Experience	-0.003	-0.008	0.002	0.003	-0.022	0.310
Race → Anger-Experience	0.060	-0.061	0.180	0.061	0.021	0.326
<i>Predicting Shame Experience</i>						
IFD → Shame-Experience	0.243	0.021	0.461	0.113	0.098	0.031
Identity Threat → Shame-Experience	0.705	0.616	0.795	0.046	0.765	< 0.001
Shame Trait → Shame-Experience	0.203	0.099	0.308	0.053	0.106	< 0.001
Neuroticism → Shame-Experience	0.214	0.145	0.284	0.035	0.159	< 0.001
Narcissism → Shame-Experience	0.141	-0.155	0.436	0.151	0.022	0.353
Sex → Shame-Experience	-0.091	-0.207	0.019	0.057	-0.036	0.114
Age → Shame-Experience	-0.002	-0.007	0.003	0.003	-0.018	0.435
Race → Shame-Experience	0.030	-0.093	0.149	0.062	0.011	0.626

Note: n = 610 (307 in Distrust condition; 303 in Neutral condition). Statistically significant coefficients in bold. Confidence intervals (CI) [LL, UL] were computed from 5,000 bootstrapped resamples. IFD - Interpersonal Felt Distrust; MIC - Moral Identity Centrality; Sex: Female, Male; Race: White, Non-White

Figure 3

Study 1. SEM Hypothesis Testing Results



Note: $n=610$; all variables latent. Bold paths are hypothesized relationships. Dashed paths are direct effects (not hypothesized but examined to test mediation). Values in each path: Hypothesis, unstandardized estimate, significance, standardized estimate in brackets. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$

4.4.8.6 Simple Slopes Analysis

A simple slopes analysis was conducted using the *semTools* package in R (Regorz, 2022). This analysis examined how the interaction between Scenario condition (Neutral vs. Distrust) and Moral Identity Centrality (latent, mean-centered; MIC.lv) influenced Work-Related Identity Threat. The results presented in Figure 4 show that the effect of being in the Distrust condition (vs. neutral) was significant at low ($b = 1.91$, $SE = 0.57$, $z = 3.37$, $p = .001$), average ($b = 2.21$, $SE = 0.07$, $z = 31.98$, $p < .001$), and high levels of MIC ($b = 2.51$, $SE = 0.56$, $z = 4.46$, $p < .001$), indicating that the effect of distrust on identity threat became stronger as MIC increased. Figure 4

illustrates the conditional effect of interpersonal felt distrust (vs. neutral condition) on work-related identity threat across varying levels of moral identity centrality. The plotted simple slopes reveal a clear upward trend, indicating that higher values of the moderator amplify the impact of distrust on identity threat. Confidence intervals at each moderator level highlight the variability at each level, thus reinforcing the reliability of these effects. These findings suggest that moral identity centrality plays a significant role in threatening one's work-related identity when individuals perceive interpersonal distrust from a supervisor.

Figure 4

Study 1. Simple Slopes Plot Depicting Conditional Effects of Scenario (Distrust vs. Neutral) by Moral Identity Centrality

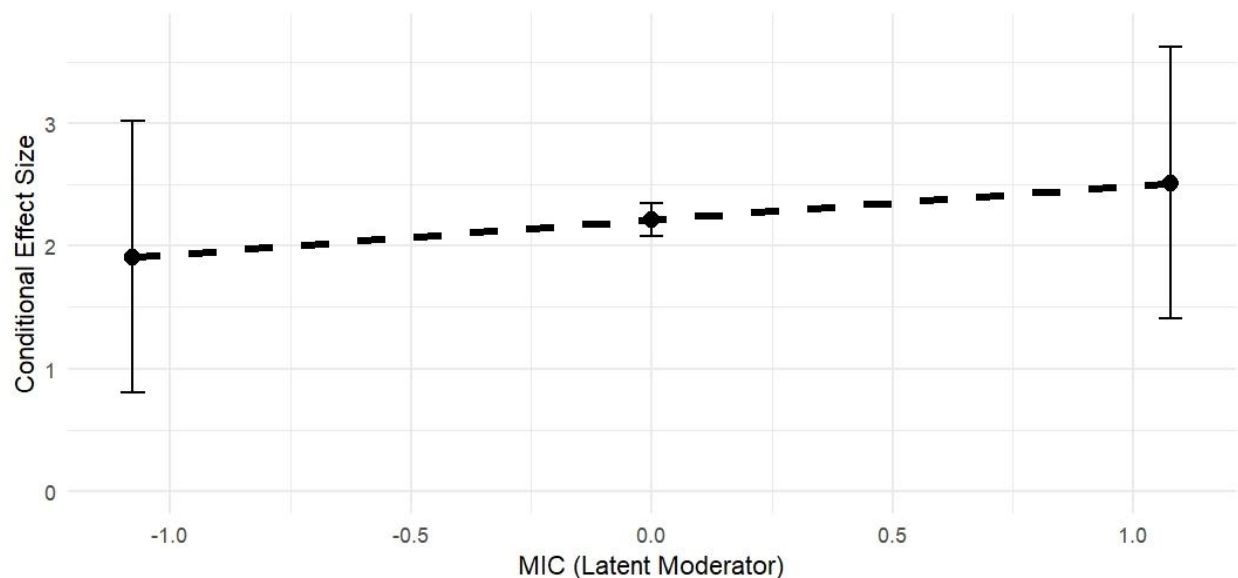
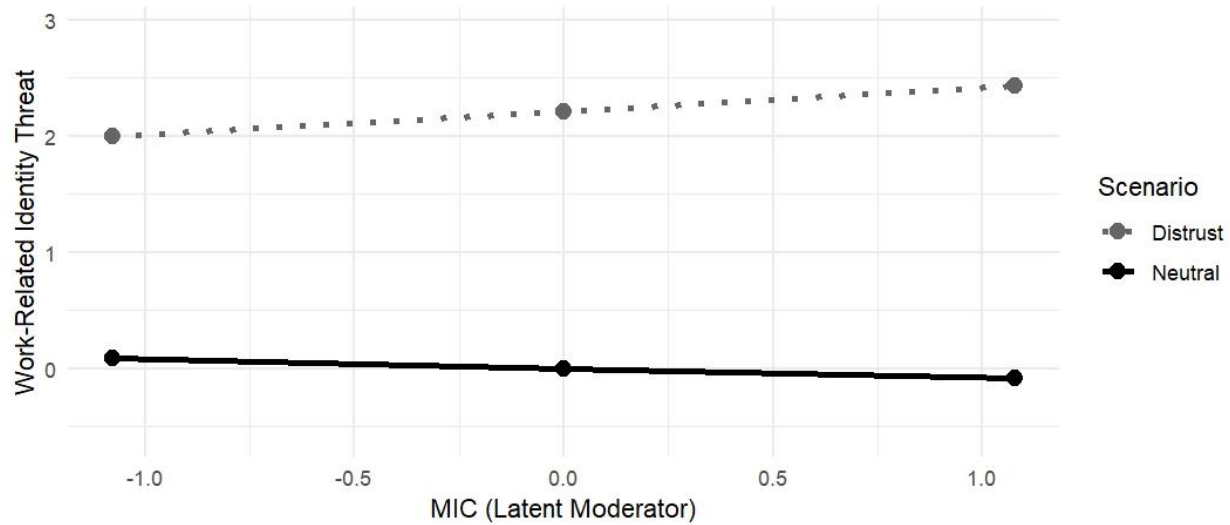


Figure 5 further displays predicted values of work-related identity threat across levels of MIC for each scenario condition. While the latent interaction term was not statistically significant in the formal SEM analysis, the pattern of conditional effects (Figure 4) and the divergence in

predicted values of work-related identity threat (Figure 5) suggest that moral identity centrality may amplify the impact of interpersonal distrust on work-related identity threat.

Figure 5

Study 1. Predicted Work-Related Identity Threat Across MIC Levels by Scenario Condition



4.4.8.7 Mediation Analysis. To provide a more detailed understanding of the mediation effect, I estimated the direct, indirect, and total effects for the model, along with the proportional value of the indirect effects. I performed the calculations on the unstandardized estimates in the regression output of the SEM model. Based on the output, the estimate from Interpersonal Felt Distrust to Identity Threat (path *a*) is 2.210 ($p < 0.001$), the estimate from Identity Threat to Anger-Experience (path *b*) is 0.835 ($p < 0.001$), and the estimate of the direct effect of Interpersonal Felt Distrust on Anger-Experience (path *c*) is 0.048 (n.s.). The indirect effect of Interpersonal Felt Distrust on Anger-Experience through Identity Threat ($a \times b$) is 1.845 ($p < 0.001$). The total effect of Interpersonal Felt Distrust on Anger Experience ($c + a \times b$) is 1.893. The proportional indirect effect ($a \times b / c + a \times b$) is 0.975. This means that the indirect effect of Interpersonal Felt Distrust through Identity Threat accounts for 97.5% of the variance in Anger-

Experience. Given the non-significant direct effect, this suggests that work-related identity threat fully mediates the relationship between interpersonal felt distrust and anger-experience.

With respect to the second mediating path in the model, the effect of Interpersonal Felt Distrust on Identity Threat (path *a*) is 2.210 ($p < 0.001$), the estimate from Identity Threat to Shame-Experience (path *b*) is 0.705 ($p < 0.001$), and the estimate of the direct effect of Interpersonal Felt Distrust on Shame-Experience (path *c*) is 0.243 ($p < 0.01$). The indirect effect of Interpersonal Felt Distrust on Shame-Experience through Identity Threat ($a \times b$) is 1.557 ($p < 0.001$). The total effect of Interpersonal Felt Distrust on Shame-Experience ($c + a \times b$) is 0.243. The proportional indirect effect ($a \times b / c + a \times b$) is 0.865, which means that the indirect effect of Interpersonal Felt Distrust through Identity Threat accounts for 86.5% of the variance in Shame-Experience. However, the direct effect is still significant, which suggests partial mediation. The summary of these effects is included in Table 6.

Table 6

Study 1. SEM Results: Direct, Indirect, Total and Proportional Indirect Effects in the Mediation

Path	Direct Effect	Indirect Effect	Total Effect	Proportional Indirect Effect
Interpersonal Felt Distrust → Identity Threat → Anger-Experience	0.048	1.845	1.893	0.975
Interpersonal Felt Distrust → Identity Threat → Shame-Experience	0.243	1.557	1.800	0.865

Note: $n = 610$

4.4.8.8 Conclusion. The results of this SEM analysis indicate that Interpersonal Felt Distrust significantly and positively affects Identity Threat (**Hypothesis 1 was supported**). The significant interaction effect between the scenario and MIC suggests that the relationship between interpersonal felt distrust and identity threat is stronger at higher levels of MIC (**Hypothesis 2 was supported**). Identity Threat is positively associated with Anger-Experience and Shame-

Experience (**Hypotheses 3a and 4a were supported**). Moreover, Identity Threat significantly mediates the relationship between Interpersonal Felt Distrust and both Anger-Experience and Shame-Experience (**Hypotheses 3b and 4b were supported**). Interpersonal Felt Distrust influences Anger-Experience and Shame-Experience through the mediating effect of Work-Related Identity Threat. The direct effect of Interpersonal Felt Distrust on Anger-Experience is minimal, positive, and statistically non-significant. In contrast, the direct effect of Interpersonal Felt Distrust on Shame-Experience is moderate, positive, and statistically significant, suggesting that Work-Related Identity Threat only partially explains the relationship. These findings support a mediating role of Work-Related Identity Threat, with stronger evidence for indirect effects in the case of Anger-Experience.

4.4.9 Complementary Analysis: Multiple Linear Regression

Having analyzed the model with SEM, I decided to also test my hypotheses through a series of multiple linear regressions. In SEM, the variables of interest were treated as latent constructs, meaning that they were represented by underlying factors measured through multiple indicators. In contrast, in the multiple linear regression analyses, the same variables are treated as observed by calculating the composite average of the individual items representing each variable. Running these complementary simple linear regressions provides for a more complete understanding of the mediation and moderation relationships in the model and validates the results obtained from SEM to ensure the robustness of the findings.

4.4.9.1 Hypothesis 1. Hypothesis 1 states that felt distrust leads to work-related identity threat. The first linear regression analysis examines Identity Threat as the outcome, with Interpersonal Felt Distrust as the main predictor, while controlling for Sex, Age, and Race (see Table 7). The model is statistically significant ($F(8, 601) = 137.4, p < 2.2e-16$) and accounts for

approximately 64.65% of the variance in Identity Threat ($R^2 = 0.65$, adjusted $R^2 = 0.64$). The effect of Interpersonal Felt Distrust on Identity Threat is positive and significant ($b = 2.15$, 95% CI [2.02, 2.27], $t(601) = 32.97$, $p < .001$; Std. $\beta = 1.61$, 95% CI [1.51, 1.70]). As such, **Hypothesis 1 is supported.**

Table 7

Study 1. Linear Multiple Regression Predicting Identity Threat from Interpersonal Felt Distrust (Controlling for Sex, Race and Age) (Hypothesis 1)

Identity Threat	
Variables	Coefficient (Standard Error)
Intercept	1.682*** (.06)
Scenario Distrust	2.146*** (.07)
Controls:	
Sex Male	- 0.060 (.07)
Age_centered	-0.003 (.003)
Race Asian	0.011 (.11)
Race Black	0.045 (.09)
Race Mixed	-0.148 (.15)
Race Other	-0.090 (.22)
Race Prefer Not to Say	-0.624 (.81)
R squared	0.647
Adjusted R-squared	0.642
F (8, 601)	137.4***

Note: $n = 610$. Two-tailed test, unstandardized beta coefficients reported. Sex (coded as 1 = female; 2 = male; female as reference group). Age (continuous variable) is mean-centered. Race/Ethnicity (coded as 1 = White; 2 = Black; 3 = Asian; 4 = Mixed; 5 = Other; 6 = Prefer not to say; White as reference group). *** $p < 0.001$

Regression diagnostics confirmed the validity of the model and its assumptions. The residuals had a median of -0.0607, with a standard error of 0.8 (601 degrees of freedom). The Durbin-Watson test (1.8942, $p = 0.09637$) indicated no significant autocorrelation. Nine outliers, with standardized residuals beyond ± 3 , were identified but not removed, as Cook's distance was within the acceptable range (< 0.50). Removing them slightly increased the effect size ($b = 2.22$, t

= 37.198, $p < 2e-16$; $R^2 = 0.70$) and the effect of Sex became marginally significant ($b = -0.10$, $p = 0.0944$). Homoscedasticity was supported by visual observation of residual plots, q-q plots, and a histogram showing normality and constant variance.

4.4.9.2 Hypothesis 2. Hypothesis 2 states that moral identity centrality strengthens the relationship between felt distrust and work-related identity threat. I tested the interaction effect in a linear regression model by adding the interaction term for Moral Identity Centrality (MIC, scale-continuous) and Interpersonal Felt Distrust (binary) variables. Following the recommendations of Meyers et al. (2017) on assessing interaction effects in regression models, I mean-centered the moderating variable MIC. This allowed me to view the estimates of the parameters at a more typical level of the moderator rather than at the default level of zero, which is non-existent in my dataset (the minimum value of MIC in the dataset is 1.2). The regression model of Identity Threat, predicted by the interaction of Interpersonal Felt Distrust (binary) and mean-centered MIC, is statistically significant ($F(3, 606) = 376.8$, $p < 2.2e-16$) and demonstrates a good fit insofar as it accounts for 65.10% of the variance in Identity Threat ($R^2 = 0.65$, adjusted $R^2 = 0.65$). The results are presented in Table 8. The effect of Interpersonal Felt Distrust (binary) $\times$ MIC (mean-centered) is positive and statistically significant ($b = 0.396$, 95% CI [0.16, 0.64], $t(606) = 3.24$, $p = 0.001$; Std. $\beta = 0.16$, 95% CI [0.06, 0.25]). This indicates that the interaction effect of Interpersonal Felt Distrust and MIC increases Identity Threat by 0.396 units of measurement for each unit increase in MIC. This means that **Hypothesis 2 is supported**.

Table 8

Study 1. Linear Multiple Regression Predicting Identity Threat from Interaction of Interpersonal Felt Distrust and Moral Identity Centrality mean-centered (Hypothesis 2)

Identity Threat	
Variables	Coefficient (Standard Error)
Intercept	1.655*** (.05)
Interpersonal Felt Distrust	2.14*** (.06)
MIC_centered	-0.150 . (0.09)
Interpersonal Felt Distrust x MIC	0.396** (.12)
R squared	0.651
Adjusted R-squared	0.649
F (3, 606)	376.8***

Note: n = 610. Two-tailed test, unstandardized beta coefficients reported. Intercept corresponds to Scenario = Neutral and MIC = Mean. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$

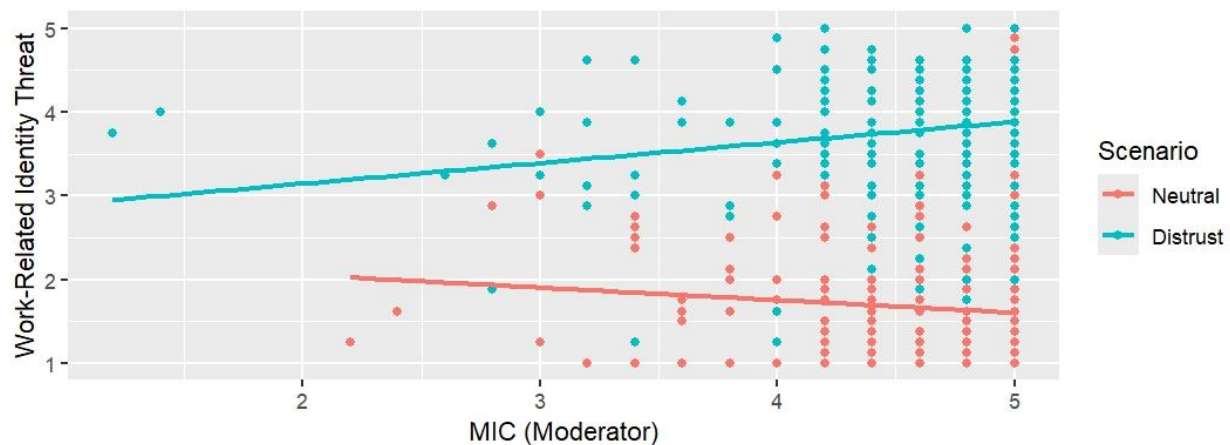
In summary, the interaction term reveals that the relationship between Interpersonal Felt Distrust and Identity Threat depends on the level of MIC. Figure 6 visualizes this effect across the two experimental conditions (Distrust and Neutral). The interaction plot shows that survey participants in the Distrust condition experience a greater level of Identity Threat with each unit of increase in MIC. Interestingly, individuals in the Neutral condition show the opposite trend, experiencing less Identity Threat with each unit increase in MIC. Although this finding lies beyond the scope of this dissertation, it underscores the powerful impact of interpersonal felt distrust, which dominates and reshapes the overall relationship.

Regression diagnostics confirmed that key assumptions were met. A visual inspection of residual and q-q plots indicated homoscedasticity and normality. The residuals had a median of -

0.0703, with a standard error of 0.7916 (606 degrees of freedom). Without the six outliers that exceeded 3 standard deviations, the interaction term effect improved ($b = 0.45$, $t = 3.97$, $p < 8.19\text{e-}05$) along with overall model fit ($R^2 = 0.70$, adjusted $R^2 = 0.69$). The MIC effect at the Neutral Scenario level was also significant ($b = -0.21$, $t = -2.48$, $p = 0.0135$). However, the Cook's distance test (values < 0.50) indicated that none of the outliers were influential, so they were ultimately retained. The Durbin-Watson test ($DW = 1.9084$, $p = 0.242$) confirmed no significant autocorrelation.

Figure 6

Study 1. Interaction of Scenario and Moral Identity Centrality (MIC) on Work-Related Identity Threat

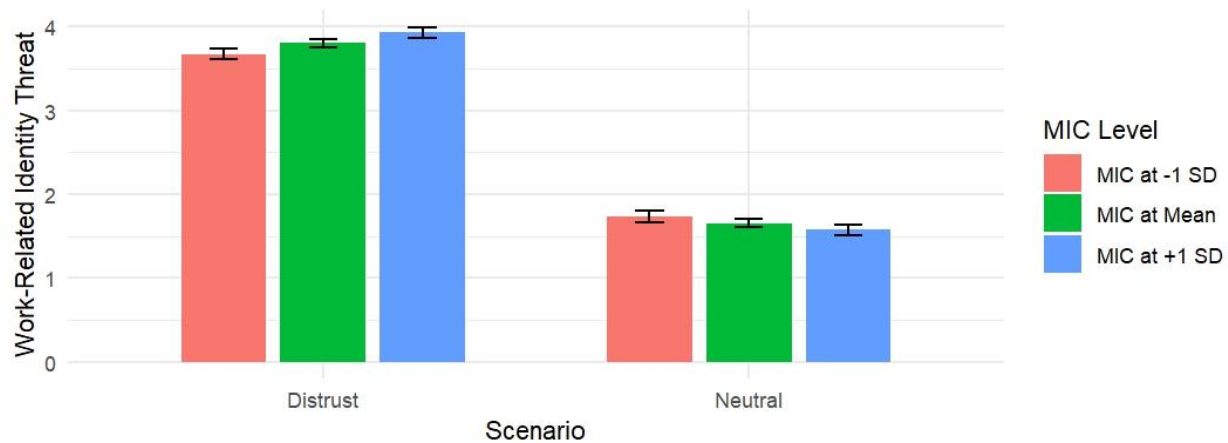


The simple slopes analysis (Preacher et al., 2006), presented in figure 7, further examines the effects of Interpersonal Felt Distrust (binary: Distrust vs. Neutral) on Identity Threat at different levels of MIC (low, medium, and high, calculated as mean $\pm$ 1 standard deviation). The analysis shows that the effect is positive and significant at all critical levels of the mean-centered MIC. This includes when MIC is less central (i.e., minus one standard deviation from the mean ($b = 1.93$, $SE = 0.09$, $t = 21.26$, $p < .001$)), at the mean value of MIC ($b = 2.14$, $SE = 0.06$, $t = 33.45$, $p < .001$),

and when it is highly central (i.e., plus one standard deviation from the mean ($b = 2.35$, $SE = 0.09$, $t = 25.87$, $p < .001$)). These coefficients correspond to the vertical differences between the bars in the Distrust and Neutral conditions within each MIC level. The results show that the effect of Interpersonal Felt Distrust on Identity Threat grows stronger (from $b = 1.93$ to $b = 2.35$) at higher levels of the moderator, in line with hypothesis 2.

Figure 7

Study 1. Effect of the Interaction Between Interpersonal Felt Distrust (Neutral vs. Distrust) and MIC (Mean-Centered) on Work-Related Identity Threat

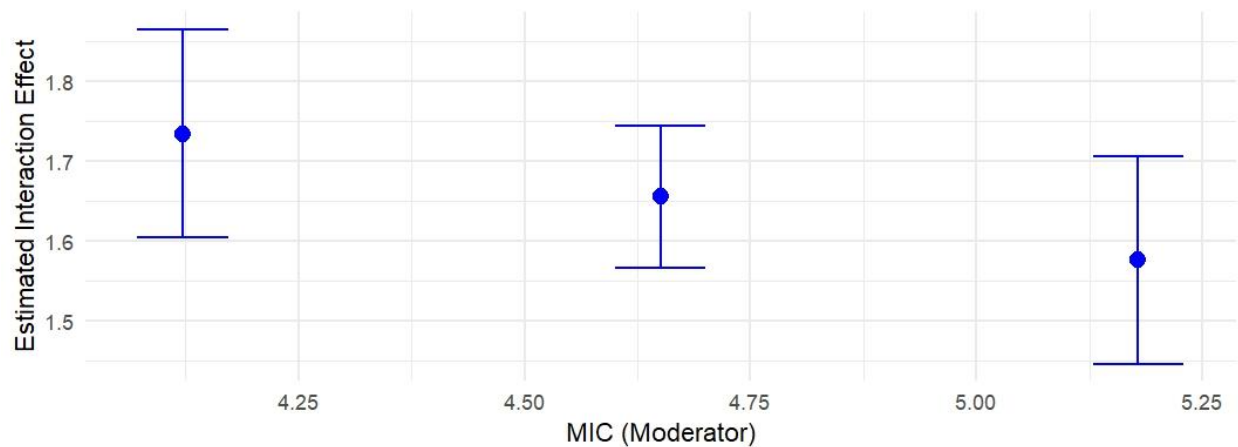


The simple slopes analysis and interaction plot highlight that as MIC increases, the effect of Interpersonal Felt Distrust on Identity Threat strengthens. This suggests that individuals with higher MIC are more sensitive to variations in Interpersonal Felt Distrust, as clearly depicted in the interaction plot. Building on these findings, I examined the changes in the overall size of the interaction effect (Interpersonal Felt Distrust $\times$ MIC) at different levels of MIC. Unlike the slopes analysis, this examination focuses on how the magnitude of the interaction effect changes at different levels of MIC. The results reveal a slight decrease in the magnitude of the interaction effect as MIC increases, though it remains significant at all levels: at minus one standard deviation:

$b = 1.73$, $SE = 0.66$, 95% CI [1.60, 1.86], $t = 26.13$, $p < .001$; at the mean of MIC: $b = 1.65$, $SE = 0.05$, 95% CI [1.57, 1.74], $t = 36.39$, $p < .001$; and at plus one standard deviation: $b = 1.58$, $SE = 0.66$, 95% CI [1.45, 1.71], $t = 23.82$, $p < .001$. This change in magnitude of the interaction effect of Interpersonal Felt Distrust and MIC on the Identity Threat at three critical levels of MIC is shown in Figure 8.

Figure 8

Study 1. Confidence Intervals of the Interaction Effect (Scenario Distrust x MIC) on Identity Threat at Different Levels of MIC



4.4.9.3 Hypotheses 3a and 3b. Hypothesis 3a states that interpersonal felt distrust leads to anger, which Hypothesis 3b posits is mediated by work-related identity threat. Before conducting the regression analysis, I explored the relationship between Interpersonal Felt Distrust (binary: Distrust vs. Neutral) and Anger Experience. A visual inspection of the box plot, violin plot, and bar plot with error bars revealed significant differences, with the Distrust scenario consistently showing higher Anger Experience values across median, distribution, and mean measures when compared to the Neutral scenario. Next, I performed a simple mediation analysis using a standard ordinary least squares regression following the steps outlined in Meyers et al. (2017). I ran three

separate regression analyses and compared the regression coefficients to determine the presence, significance, and type of mediation present in the structural model:

Regression analysis 1: I used Interpersonal Felt Distrust (binary: Distrust vs. Neutral) to predict Anger-Experience (unmediated model, path c , i.e., the direct effect), while controlling for Age, Sex, and Race.

Regression analysis 2: I used Interpersonal Felt Distrust (binary: Distrust vs. Neutral) to predict Identity Threat (mediated model, path a , i.e., the indirect effect), while controlling for Age, Sex, and Race.

Regression analysis 3: I used Interpersonal Felt Distrust (binary: Distrust vs. Neutral) and Identity Threat together to predict Anger-Experience (mediated model, path b , i.e., the indirect effect, and path c' , i.e., the direct effect), while controlling for Age, Sex, Race, Anger Trait, Narcissism, and Neuroticism.

For Analysis 1 (direct effect of X on Y, path c), I fitted a linear model (estimated using OLS) to predict Anger-Experience with Interpersonal Felt Distrust, Sex, Age, and Race. The model explains a statistically significant and substantial proportion of variance ($R^2 = 0.49$, $F(4, 605) = 148.07$, $p < .001$, adj. $R^2 = 0.49$). The model's intercept, corresponding to Scenario = Neutral, Sex = Female, Age = 0, and Race = White is at 1.65 (95% CI [1.38, 1.93], $t(605) = 11.99$, $p < .001$). Within this model, the effect of Interpersonal Felt Distrust is statistically significant and positive ($b = 1.68$, 95% CI [1.55, 1.82], $t(605) = 24.25$, $p < .001$; Std. $\beta = 1.40$, 95% CI [1.29, 1.52]). Among the control variables, the effect of Age is statistically significant and negative ($b = -7.51\text{e-}03$, 95% CI [-0.01, -1.35e-03], $t(605) = -2.39$, $p = 0.017$; Std. $\beta = -0.07$, 95% CI [-0.13, -0.01]) while the effects of Sex (Male) ($b = 0.02$) and Race (Non-White) ($b = 0.07$) are positive and

statistically non-significant. A visual inspection of diagnostic plots indicated no unusual patterns, confirming the normality and homoscedasticity of residuals. While four outliers exceeded three standard deviations, their removal slightly improved the model fit and effect of Interpersonal Felt Distrust on Anger-Experience; however, Cook's distance tests showed that no outliers were influential, so none were removed.

Analysis 2 (effect of X on M, path *a*) (i.e., the effect of Interpersonal Felt Distrust on Identity Threat while controlling for Sex, Age, and Race) was reported upon earlier in the context of testing Hypothesis 1. Assumptions for normality, homoscedasticity, and lack of multivariate outliers have been confirmed at that time. The only change in the model this time was a change to a more condensed, binary variable of Race (White and Non-White). This model explains a statistically significant and substantial proportion of variance ($R^2 = 0.65$, $F(4, 605) = 275.09$, $p < .001$, adj. $R^2 = 0.64$). The model's intercept, corresponding to Scenario = Neutral, Sex = Female, Age = 0, and Race = White, is at 1.80 (95% CI [1.55, 2.05], $t(605) = 13.99$, $p < .001$). Within the model, the effect of Interpersonal Felt Distrust is statistically significant and positive ($b = 2.14$, 95% CI [2.02, 2.27], $t(605) = 33.07$, $p < .001$; Std. $\beta = 1.60$, 95% CI [1.51, 1.70]). The effects of the control variables are statistically non-significant and negative (e.g., Sex (Male) ($b = -0.06$), Age ($b = -3.16e-03$), Race (Non-White) ($b = -8.86e-03$)).

In Analysis 3, I fitted a linear model to predict Anger Experience with Identity Threat (path *b*) and Interpersonal Felt Distrust (path *c'*), while controlling for Sex, Age, Race, Anger Trait, Narcissism, and Neuroticism. The model explains a statistically significant and substantial proportion of variance ($R^2 = 0.75$, $F(8, 601) = 223.06$, $p < .001$, adj. $R^2 = 0.74$). The model's intercept, corresponding to Scenario = Neutral, Identity Threat = 0, Sex = Female, Age = 0, Race = White, Anger Trait = 0, Narcissism = 0, and Neuroticism = 0, is at -0.25 (95% CI [-0.54, 0.05],

$t(601) = -1.66, p = 0.098$). Within the model, the effect of Identity Threat (indirect effect, path b) is statistically significant and positive ($b = 0.70$, 95% CI [0.64, 0.76], $t(601) = 22.13, p < .001$; Std. $\beta = 0.78$, 95% CI [0.71, 0.85]), while the effect of Interpersonal Felt Distrust (direct effect, path c') is statistically non-significant and positive ($b = 0.16$, 95% CI [-1.62e-03, 0.33], $t(601) = 1.94, p = 0.052$; Std. $\beta = 0.14$, 95% CI [-1.35e-03, 0.27]). Among the control variables, only Anger Trait shows a statistically significant and positive effect ($b = 0.27$, 95% CI [0.15, 0.39], $t(601) = 4.43, p < .001$; Std. $\beta = 0.12$, 95% CI [0.07, 0.17]). Visual inspections of diagnostic plots confirmed normality, homoscedasticity, and lack of multivariate outliers, meeting regression assumptions. The Durbin-Watson test ($DW = 2.06, p = 0.464$) showed no autocorrelation. Eight potential outliers (residuals > 3 SD) were identified, but testing without them resulted in minor changes, including slightly larger and significant effects for Interpersonal Felt Distrust ($b = 0.16$ to $0.18, p = 0.052$ to 0.019) and Sex ($b = 0.06$ to $0.09, p = 0.222$ to 0.069). As these differences were minimal, the full-sample regression output was retained for consistency. The output of the three regression analyses is presented in Table 9.

An examination of these regression analyses reveals that the effects of Interpersonal Felt Distrust on Identity Threat (indirect effect, path a , $b = 2.14, p < 0.001$) and Identity Threat on Anger-Experience (indirect effect, path b , $b = 0.70, p < 0.001$) are statistically significant and move in the expected direction (positive). Thus, **Hypothesis 3a is supported**.

The initial observation that the unstandardized regression coefficient beta (b) of Interpersonal Felt Distrust predicting Anger-Experience becomes much lower in magnitude and almost non-significant (at $p = 0.052$) with the mediator (Identity Threat) in the model suggests perfect mediation (Preacher & Hayes, 2004). To ensure robust findings for Hypothesis 3b, I conducted a formal mediation analysis using the *mediate* function from R's mediation package.

Table 9*Study 1. Mediation Testing. Results of the Three Regression Analyses (Hypothesis 3)*

Analysis	Path from Figure M	b	SE of b	β	r	t	p	sig.	Adj. R ²	F (DF)	p	sig.
1	Interpersonal Felt Distrust to Anger											
	c Experience	1.68	0.070	1.40	0.700	24.25	< .001	***	0.491	148.07 (4, 605)	< .001	***
	Controls:											
	Sex (Male)	0.18	0.070	0.01	-0.020	0.249	0.803					
	Age	-0.01	0.003	-0.07	-0.050	-2.392	0.017	*				
	Race (Non-White)	0.07	0.075	0.06	0.036	0.919	0.359					
2	Interpersonal Felt Distrust to Identity Threat											
	a Threat	2.14	0.065	1.60	0.803	33.08	< .001	***	0.643	275.09 (4, 605)	< .001	***
	Controls:											
	Sex (Male)	-0.06	0.066	-0.04	-0.061	-0.90	0.370					
	Age	-0.00	0.003	-0.03	0.008	-1.08	0.281					
	Race (Non-White)	-0.00	0.070	-0.00	-0.000	-0.13	0.899					
3	Identity Threat to Anger											
	b Experience	0.70	0.032	0.78	0.850	22.13	< .001	***	0.745	223.1 (8, 601)	< .001	***
	Interpersonal Felt Distrust to Anger											
	c' Experience	0.16	0.083	0.14	0.700	1.94	0.052	.				
	Controls:											
	Sex (Male)	0.06	0.052	0.05	-0.020	1.22	0.222					
	Age	-0.00	0.000	-0.03	-0.050	-1.27	0.206					
	Race (Non-White)	0.06	0.054	0.05	0.036	1.16	0.245					
	Anger Trait	0.27	0.062	0.12	0.260	4.43	< .001	***				
	Narcissism	0.14	0.139	0.02	0.050	1.04	0.298					
	Neuroticism	0.02	0.035	0.02	0.210	0.60	0.547					

Note: n=610. *b* is the unstandardized coefficient; SE of *b* is the standard error of beta; β is the standardized coefficient; *r* is the Pearson correlation of the predictor and the outcome variable. Point-biserial correlation *r* is calculated between continuous variables Identity Threat and Anger-Expression and binary variables Scenario (0 = Neutral, 1 = Distrust), Sex (0 = Female, 1 = Male) and Race (0 = White, 1 = Non-White). The *t* statistic and the *F* statistic (with degrees of freedom) are provided with their *p*-value and indicator of significance. ****p* < 0.001, ***p* < 0.01, **p* < 0.05, .*p* < 0.1. The three regressions test the following paths in the mediation analysis: c = total path between independent variable (X) and dependent variable (Y); a = first indirect path between X and mediator (M); b = second indirect path between M and Y; c' = direct path between X and Y after controlling for the effect of M. Relevant control variables are included at each step of the regression series.

This Quasi-Bayesian approach, with 1000 simulations, estimated the indirect (ACME: 1.1792, 95% CI [1.0332, 1.33]), direct (ADE: 0.1617, 95% CI [0.0016, 0.32]), and total (1.3409, 95% CI [1.2077, 1.48]) effects, along with the proportion mediated (88.07%, 95% CI [0.7731, 1.00]). The analysis confirmed significant and positive effects, highlighting that Identity Threat mediates most of the relationship between Interpersonal Felt Distrust and Anger Experience, as detailed in Table 10. The results from the *mediate* package differ slightly from earlier analytical

calculations because this package uses a simulation-based Quasi-Bayesian approach to account for uncertainty across all model components. This method provides more robust estimates, which may lead to small variations in effect sizes and confidence intervals when compared to standard estimates. Based on this output, I conclude that the mediation effect is significant, with a substantial proportion of the total effect being mediated. The presence of both significant direct and indirect effects indicates partial mediation, as the treatment (Interpersonal Felt Distrust) affects the outcome (Anger-Experience) both directly and indirectly through the mediator (Identity Threat). The use of the Quasi-Bayesian approach and 1000 simulations ensure robust and reliable estimates of the mediation effects. Thus, **Hypothesis 3b is supported.**

Table 10

Study 1. Causal Mediation Analysis, Interpersonal Felt Distrust Predicting Anger-Experience Through Identity Threat (Hypothesis 3b)

Effect	Estimate	95% CI	95% CI	p-value
		Lower	Upper	
ACME	1.1792	1.0332	1.33	< 2e-16 ***
ADE	0.1617	0.0016	0.32	0.048 *
Total Effect	1.3409	1.2077	1.48	< 2e-16 ***
Prop. Mediated	0.8807	0.7731	1	< 2e-16 ***

Note: n = 610. Simulations: 1000 using Quasi-Bayesian Confidence Intervals

4.4.9.4 Hypotheses 4a and 4b. Hypothesis 4a states that interpersonal felt distrust affects shame experience through workplace-related identity threat and through a second mediation in Hypothesis 4b. As was the case in testing the previous hypothesis, I completed another set of three linear regressions to test the direct (c and c') and indirect effects (a , b , $a \times b$, and proportional indirect effect) in the mediation model. I controlled for Sex, Age, Race, Shame Trait, Narcissism, and Neuroticism. Before the regression, visual inspections of box, violin, and bar plots showed a

clear difference in Shame Expression between Neutral and Distrust, with no signs of non-normality.

Fitting a linear model (OLS) to predict Shame-Experience with Interpersonal Felt Distrust, Sex, Age, and Race (path *c*), resulted in a model that explains a statistically significant and substantial proportion of variance ($R^2 = 0.49$, $F(4, 605) = 147.42$, $p < .001$, adj. $R^2 = 0.49$). The model's intercept, corresponding to Scenario = Neutral, Sex = Female, Age = 0, and Race = White, is significant at 1.92 (95% CI [1.62, 2.22], $t(605) = 12.56$, $p < .001$). Within this model, the effect of Interpersonal Felt Distrust is statistically significant and positive ($b = 1.83$, 95% CI [1.68, 1.98], $t(605) = 23.75$, $p < .001$; Std. $\beta = 1.38$, 95% CI [1.26, 1.49]). The effects of all control variables are statistically significant and negative: Sex (Male) ($b = -0.28$, 95% CI [-0.44, -0.13], $t(605) = -3.64$, $p < .001$; Std. $\beta = -0.21$, 95% CI [-0.33, -0.10]); Age ($b = -9.77\text{e-}03$, 95% CI [-0.02, -2.93e-03], $t(605) = -2.81$, $p = 0.005$; Std. $\beta = -0.08$, 95% CI [-0.14, -0.02]); and Race (Non-White) ($b = -0.05$, 95% CI [-0.21, 0.12], $t(605) = -0.57$, $p = 0.567$; Std. $\beta = -0.04$, 95% CI [-0.16, 0.09]). This suggests that the effect of interpersonal felt distrust results in less experience of shame for men (compared to women), older people (compared to younger people), and visible minorities (compared to White people). Visual inspection of residuals (q-q plot, histogram, and residuals vs. fitted values) showed no unusual patterns, supporting normality and homoscedasticity. The Durbin-Watson test ($DW = 1.85$, $p = 0.06$) indicated no autocorrelation. Five potential outliers (residuals > 3 SD) were identified, but testing the model without them yielded similar coefficients, so they were not deemed influential and were retained.

Path *a* – the effect of Interpersonal Felt Distrust on Identity Threat (i.e., the first indirect path in the mediation model) has already been tested in the previous mediation analysis (see above). The model meets our assumptions for the normal distribution and independence of

residuals, lacks multivariate outliers and homoscedasticity, and explains a statistically significant and substantial proportion of variance ($R^2 = 0.65$, $F(4, 605) = 275.09$, $p < .001$, adj. $R^2 = 0.64$). The intercept, corresponding to Scenario = Neutral, Sex = Female, Age = 0, and Race = White, is at 1.80 (95% CI [1.55, 2.05], $t(605) = 13.99$, $p < .001$). Within the model, the effect of Interpersonal Felt Distrust is statistically significant and positive ($b = 2.14$, 95% CI [2.02, 2.27], $t(605) = 33.07$, $p < .001$; Std. $\beta = 1.60$, 95% CI [1.51, 1.70]) and the effects of control variables are statistically negative and non-significant, e.g. Sex ($b = -0.06$), Age ($b = -3.16e-03$), Race ($b = -8.86e-03$).

The final regression analysis included an estimation of paths b (the second indirect effect of M on Y in the mediation model) and c' (the direct effect of X on Y while controlling for the presence of M). Both paths have been estimated by fitting a linear regression model to predict Shame-Experience with Identity Threat, Interpersonal Felt Distrust, Sex, Age, Race, Shame Trait, Narcissism, and Neuroticism. The model explains a statistically significant and substantial proportion of variance ($R^2 = 0.73$, $F(8, 601) = 200.69$, $p < .001$, adj. $R^2 = 0.72$). The intercept, corresponding to Scenario = Neutral, Identity Threat = 0, Sex = Female, Age = 0, Race = White, Shame Trait = 0, Narcissism = 0, and Neuroticism = 0, is statistically significant at -1.13 (95% CI [-1.65, -0.61], $t(601) = -4.29$, $p < .001$). Within the model, the effect of Interpersonal Felt Distrust on Shame-Experience (path c') is statistically significant and positive ($b = 0.37$, 95% CI [0.18, 0.56], $t(601) = 3.87$, $p < .001$; Std. $\beta = 0.28$, 95% CI [0.14, 0.42]). Therefore, **Hypothesis 4a is supported**. Likewise, the effect of Identity Threat on Shame-Experience (path b) is statistically significant and positive ($b = 0.68$, 95% CI [0.61, 0.75], $t(601) = 18.57$, $p < .001$; Std. $\beta = 0.68$, 95% CI [0.61, 0.75]). Among control variables, only Shame Trait ($b = 0.23$, 95% CI [0.14, 0.33], $t(601) = 5.00$, $p < .001$; Std. $\beta = 0.11$, 95% CI [0.07, 0.16]) and Neuroticism ($b = 0.22$, 95% CI

[0.16, 0.29], $t(601) = 6.75$, $p < .001$; Std. $\beta = 0.15$, 95% CI [0.11, 0.20]) have a statistically significant and positive effect on Shame-Experience.

A visual inspection of residuals confirmed normality and homoscedasticity, while the Durbin-Watson statistic (1.81) indicated no autocorrelation, though the p-value (0.018) suggested autocorrelation at the 5% level. Newey-West standard errors were computed to address this, but results were nearly identical to the original Ordinary Least Squares (OLS) model, so OLS estimates were retained for consistency. Nine potential outliers (residuals > 3 SD) were identified, but their removal caused minimal changes, except for a slight increase in the significance of the Sex (Male) coefficient ($b = -0.12$, $p = 0.033$). As these changes were minor, the outliers were deemed non-influential and retained. The results of the three regression analyses are presented in Table 11.

Next, I performed a formal mediation analysis with the *mediate* function from the *mediation* package in R (Quasi-Bayesian approach with the use of 1000 simulation-based confidence intervals and p-values for the mediation effects). The summary of the mediation analysis is included in Table 12. The Average Causal Mediation Effect (ACME) shows a significant positive indirect effect (1.453, 95% CI [1.288, 1.64]) of Interpersonal Felt Distrust on Shame-Experience through Identity Threat. The Average Direct Effect (ADE) is also significant and positive (0.372, 95% CI [0.172, 0.56]), resulting in a Total Effect of 1.825 (95% CI [1.679, 1.96]). Approximately 79.6% of the total effect is mediated by Identity Threat (95% CI [0.706, 0.90]). Therefore, **Hypothesis 4b is supported.**

Based on the mediation analysis output, I conclude that the mediation effect is significant, with a substantial proportion of the total effect being mediated. The presence of both significant direct and indirect effects indicates partial mediation, as the treatment (Interpersonal Felt Distrust) affects the outcome (Shame-Experience) both directly and indirectly through the mediator

(Identity Threat). The use of the Quasi-Bayesian approach and 1000 simulations ensures robust and reliable estimates of the mediation effects. Overall, the complementary regression analysis results are consistent with the SEM analysis results.

Table 11

Study 1. Mediation Testing. Results of the Three Regression Analyses. (Hypothesis 4)

Analysis	Path from Figure M	b	SE of b	β	r	t	p	sig.	Adj. R ²	F (DF)	p	sig.
1	Interpersonal Felt Distrust to Shame											
	c Experience	1.83	0.077	1.38	0.691	23.75	< .001	***	0.490	147.4 (4, 605)	< .001	***
	Controls:											
	Sex (Male)	-0.28	0.078	-0.21	-0.134	-3.64	< .001	***				
	Age	-0.01	0.003	-0.08	-0.043	-2.81	0.005	**				
	Race (Non-White)	-0.05	0.083	-0.04	-0.006	-0.57	0.567					
2	Interpersonal Felt Distrust to Identity											
	a Threat	2.14	0.065	1.60	0.803	33.08	< .001	***	0.643	275.09 (4, 605)	< .001	***
	Controls:											
	Sex (Male)	-0.06	0.066	-0.04	-0.061	-0.90	0.370					
	Age	-0.00	0.003	-0.03	0.008	-1.08	0.281					
	Race (Non-White)	-0.00	0.070	-0.00	-0.000	-0.13	0.899					
3	Identity Threat to Shame Experience	0.68	0.036	0.68	0.850	18.57	< .001	***				
	Interpersonal Felt Distrust to Shame											
	c' Experience	0.37	0.097	0.28	0.691	3.87	< .001	***	0.724	200.7 (8, 601)	< .001	***
	Controls:											
	Sex (Male)	-0.10	0.060	-0.08	-0.134	-1.74	0.082	.				
	Age	-0.00	0.003	-0.03	-0.043	-1.19	0.234					
	Race (Non-White)	0.03	0.062	0.02	-0.006	0.46	0.646					
	Shame Trait	0.23	0.047	0.11	0.160	5.00	< .001	***				
	Narcissism	0.15	0.156	0.02	-0.030	0.96	0.336					
	Neuroticism	0.22	0.033	0.15	0.300	6.75	< .001	***				

Note: n=610. *b* is the unstandardized coefficient; SE of *b* is the standard error of beta; β is the standardized coefficient; *r* is the Pearson correlation of the predictor and the outcome variable. Point-biserial correlation is calculated between continuous variables Identity Threat and Shame-Expression and binary variables Scenario (0 = Neutral, 1 = Distrust), Sex (0 = Female, 1 = Male) and Race (0 = White, 1 = Non-White). The *t* statistic and the *F* statistic (with degrees of freedom) are provided with their *p*-value and indicator of significance. ****p* < 0.001, ***p* < 0.01, **p* < 0.05, .*p* < 0.1. The three regressions test the following paths in the mediation analysis: *c* = total path between independent variable (*X*) and dependent variable (*Y*); *a* = first indirect path between *X* and mediator (*M*); *b* = second indirect path between *M* and *Y*; *c'* = direct path between *X* and *Y* after controlling for the effect of *M*. Relevant control variables are included at each step of the regression series.

Table 12

Study 1. Causal Mediation Analysis, Interpersonal Felt Distrust Predicting Shame-Experience Through Identity Threat (Hypothesis 4b)

Effect	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	1.453	1.288	1.64	< 2e-16 ***
ADE	0.372	0.172	0.56	< 2e-16 ***
Total Effect	1.825	1.679	1.96	< 2e-16 ***
Prop. Mediated	0.796	0.706	0.9	< 2e-16 ***

Note: n = 610. Simulations: 1000 using Quasi-Bayesian Confidence Intervals

4.5 Study 1 Discussion and Conclusion

In Study 1, I examined whether employees who felt interpersonally distrusted would experience greater workplace-related identity threat and stronger emotions of anger and/or shame when compared to those exposed to a neutral workplace scenario. Additionally, I explored whether moral identity centrality moderates the level of identity threat experienced. Testing these hypotheses using an experimental design helped isolate the impact of interpersonal felt distrust on identity threat, ruling out to some extent other potential explanations through controlled conditions (Neutral and Distrust), careful design of the two experimental scenarios (based on the distrust literature), and controlling for other possible explanatory variables that could account for the hypothesized effect. Examination of the effect of interpersonal felt distrust on anger and shame through a serial mediation involving work-related identity threat allowed to discover significant correlations between the variables of interest and the mediating role of work-related identity, but the study design did not intend to provide causal inferences on the relationship between felt distrust and emotions.

The four hypotheses tested in Study 1 were tested using two different statistical approaches: SEM and multiple linear regression. In both analyses, all of the hypotheses were supported. The results are consistent, supporting the direct effects of felt distrust on identity threat (H1), the moderating effect of moral identity centrality (H2), the mediation of identity threat between felt distrust and anger experience (H3), and between felt distrust and shame experience (H4) (see Table 13).

The comparison table highlights consistent direction, significance, and magnitude of coefficients for all effects (paths a, b, and proportional indirect effects) across the SEM and regression analyses. Interpersonal Felt Distrust increases Identity Threat by an average of 2.15 units, with MIC (moderator) being significant in both methods, though the SEM coefficient is larger (1.264 vs. 0.396). This difference reflects SEM's ability to account for measurement error, item-level precision, and complex covariance structures; in contrast, regression relies on simpler estimation methods using observed variables. Both methods confirm that Identity Threat significantly impacts Anger-Experience and Shame-Experience (path b), with effects at or above 0.68 per unit change and consistently show that Work-Related Identity Threat mediates the relationship between Interpersonal Felt Distrust and experienced Anger and Shame. About 88% of the total effect is mediated in the Distrust–Anger pathway and 80% in the Distrust–Shame pathway, with results aligned across methods.

Table 13*Study 1. Comparison of Effects Between SEM and Multiple Regression Analyses*

Path	SEM Analysis						Multiple Regression Analysis							
	Moderat. Effect	Indirect Effect (path <i>a</i>)	Indirect Effect (path <i>b</i>)	Direct Effect (path <i>c'</i>)	Total Indirect Effect	Total Effect	Proportional Indirect Effect	Moderat. Effect	Indirect Effect (path <i>a</i>)	Indirect Effect (path <i>b</i>)	Direct Effect (path <i>c'</i>)	Total Indirect Effect	Total Effect	Proportional Indirect Effect
Hypothesis 1: IFD → Identity Threat	0.282***	2.210***						0.396**	2.146***					
Hypothesis 2: IFD x MIC → Identity Threat														
Hypothesis 3a: Identity Threat → Anger-E.		0.835***							0.700***					
Hypothesis 3b: IFD → Identity Threat → Anger-E.				0.048	1.845***	1.893	97.50%				0.162 .	1.179***	1.341	88.07%
Hypothesis 4a: Identity Threat → Shame-E.		0.705***							0.680***					
Hypothesis 4b: IFD → Identity Threat → Shame-E.				0.243*	1.557*	1.800	86.50%				0.372***	1.453***	1.825	79.60%

Note: n = 610. All coefficients reported are unstandardized coefficients from SEM and Regression models. Moderat. Effect - Moderating Effect; IFD - Interpersonal Felt Distrust; MIC - Moral Identity Centrality; Anger-E - Anger-Experience; Shame-E - Shame-Experience. ****p* < 0.001, ***p* < 0.01, **p* < 0.05, .*p* < 0.1 .

Slight differences in results between SEM and regression analyses stem from the complexities each method captures and their underlying assumptions, such as sample size and normality. SEM provides a holistic view, accounting for moderation, latent variables, and measurement error, with the measurement model showing good fit indices (CFI = 0.940, TLI = 0.932, RMSEA = 0.080 [CI 0.075, 0.086], SRMR = 0.039). Regression analysis, while more focused on individual relationships and not accounting for measurement error, explains substantial variance (R^2 and adjusted R^2). Despite differing effect estimates, both methods consistently demonstrate significant and substantive effects supporting the theoretical model (see Appendix C Table C1).

Chapter 5: Study 2 Methods and Results

5.1 Study 2 Description

Study 2 was designed to assess individuals' actual emotional and behavioral reactions to interpersonal felt distrust by their supervisors in the workplace, testing the full conceptual model provided across Hypotheses 1 – 8. The objectives were the following: (1) to strengthen the external validity of Study 1's findings on the impact of interpersonal felt distrust by one's supervisor on an individual's sense of work-related identity and emotions (i.e., anger and shame); and (2) to extend the scope of these findings by assessing individuals' actual behavioral reactions to interpersonal felt distrust, specifically by measuring how often they engaged in counterproductive, psychological withdrawal and exemplification behaviors after feeling distrusted by a supervisor.

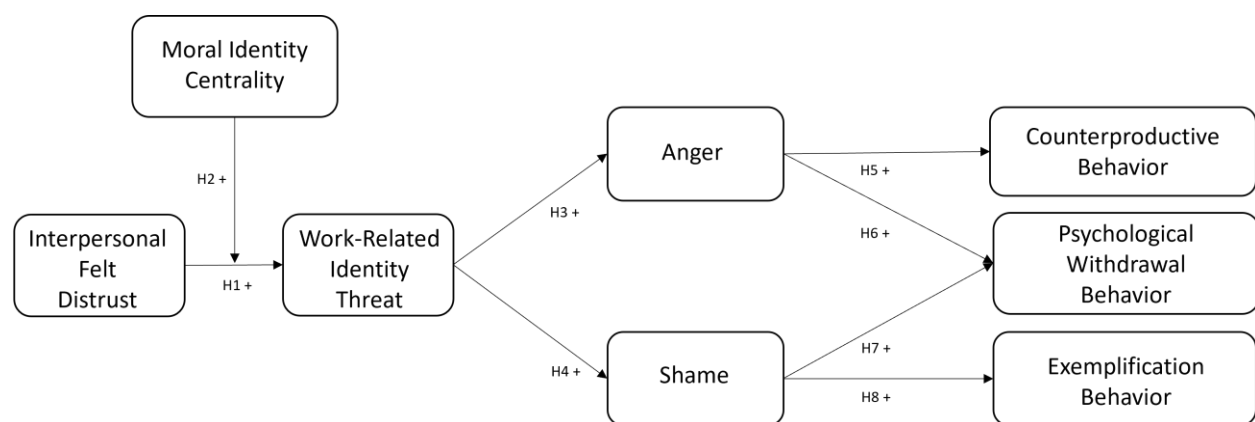
The research design involved a critical incident technique (CIT) (Flanagan, 1954) integrated into a survey. The CIT involved having participants recall the event of feeling distrusted by a supervisor with the intent to immerse them in that psychological state prior to answering the subsequent survey questions. This combined approach has been recommended by organizational scholars as an effective method of capturing individuals' responses to significant discrete workplace events, such as stress-coping (O'Driscoll & Cooper, 1994) and mistreatment (Hershcovis, 2011). It has been widely used to study individual reactions to sensitive interpersonal workplace experiences, particularly in the contexts of trust and conflict (Korsgaard et al., 2002; Lind et al., 1997; Lissak & Sheppard, 1983), personal offense (Aquino et al., 2001), abusive supervision (Priesemuth & Bigelow, 2020), unethical behaviors (Mitchell et al., 2015; Wellman et al., 2016), and gender threats (Leavitt et al., 2022).

As a feature of CIT, Study 2 participants were first asked to recall and reflect on their most recent experience of felt distrust by a supervisor through a series of structured scale-continuous

and open-ended questions. Upon having clearly recalled this event in their memories, they were asked to respond to a series of survey questions that assessed their emotional and behavioral reactions to the event through the lens of the time of its occurrence. Incorporating CIT into the survey design ensured that responses were grounded in actual workplace occurrences rather than hypothetical scenarios, which strengthened the ecological validity of the research. The hypothesized relationships and the hypotheses under study are depicted in Figure 9 below.

Figure 9

Conceptual Model Analyzed in Study 2



5.1.1 Sample and Procedure

5.1.1.1 Sample size determination

Sample size was determined a priori in light of my plan to use path analysis within the Structural Equation Modeling (SEM) technique in R to analyze the data from Study 2. Since all variables in the structural model were treated as observed variables, the model included an estimated 75 free parameters. Given that SEM sample size recommendations vary, I followed the

N:q rule, wherein N represents the sample size and q denotes the number of free parameters (Bentler & Chou, 1987). While a ratio of 10:1 has been widely recommended for complex models with latent variables (Kline, 2016), research suggests that a minimum ratio of 5:1 is acceptable for path analysis models with observed variables (Bentler & Chou, 1987; Schreiber et al., 2006). Additionally, some sources suggest that a minimum sample size of 300 is required for SEM models to ensure stable parameter estimation and convergence (Comrey & Lee, 2013; Tabachnick & Fidell, 2013). Based on these criteria, a minimum of 375 participants was required for the purposes of this study. To enhance model stability and account for potential data loss or nonresponse, I opted to oversample and recruited 400 participants with a final usable sample of 398 respondents, thereby ensuring sufficient power for reliable estimation.

5.1.1.2 Sample characteristics

The sample was a convenience non-probability sample of individuals who had experienced interpersonal felt distrust by a supervisor in the workplace rather than a general population. My choice to target employees who had felt distrusted by a supervisor was driven by both theoretical and methodological considerations. Given that the study focuses on examining the emotional and behavioral consequences of interpersonal felt distrust, it was essential to target participants with direct and lived experience of felt distrust. Using this purposive sampling approach enhanced the ecological validity of my findings, as responses are grounded in actual workplace interactions rather than in mere hypothetical scenarios (Hershcovis, 2011; O'Driscoll & Cooper, 1994). Prior research demonstrates that studies examining mistreatment, unethical behaviors, trust, and conflict benefit from such experience-based samples, allowing for deeper insights into the phenomena under investigation (Aquino et al., 2001; Korsgaard et al., 2002; Lind et al., 1997; Wellman et al.,

2016). Data was collected through an online panel on Prolific from respondents who are employed, have experience being supervised, and have experience feeling distrusted by a supervisor at work.

Out of the 398 participants, 218 were male (54.8%) and 180 were female (45.2%). The age distribution in the dataset ranged between 20 to 75 years. The mean age was 35.72 years; the median age was 33. The largest racial group surveyed was White, making up 49.5% of the sample (197 participants), followed by Black, comprising 145 individuals (36.4%), Asian, with 27 participants (6.8%), Mixed-race, with 14 individuals (3.5%), and 15 participants (3.8%) in the "Other" category.

Additional data on employment and location was captured through the survey. Three hundred and thirteen respondents (78.6%) were employed on a full-time basis at the time of the survey's completion and 85 (21.4%) were working part-time. The core survey respondents reported being employed across a wide range of industries. The largest number of respondents, being 67 participants (16.8%), were employed in Information Technology (IT). In addition, 62 participants (15.6%) were in Healthcare & Medicine, 44 individuals (11.1%) in Education & Training, 38 participants (9.5%) in Manufacturing & Engineering, 35 individuals (8.8%) in Retail, Sales & Customer Service, 26 participants (6.5%) in the Finance & Banking sector, 15 respondents (3.8%) in Business, 19 participants (4.8%) in Marketing & Consulting, 11 participants (2.8%) in Hospitality, Food & Tourism, Transportation & Logistics, 11 participants (2.8%) in Government & Public Sector, and 9 participants (2.3%) in Legal & Security services. The remaining 40 participants (10.1%) fall into miscellaneous or unclassified industries. This breakdown illustrates a diverse professional landscape among respondents.

The data also provides valuable insights on individual experiences of interpersonal felt distrust across cultures. The 398 respondents were all English-speaking and were located in 34

distinct countries: South Africa (98 participants, 24.6%), the United States (79 participants, 19.8%), the United Kingdom (56 participants, 14.1%), Canada (23 participants, 5.8%), Kenya (15 participants, 3.8%), Italy (13 participants, 3.3%), Portugal (11 participants, 2.8%), and Poland (10 participants, 2.5%). Smaller groups of country-based respondents ranging between 1-8 participants represented an even more diverse geographic coverage, including Australia, Austria, Belgium, Brazil, Chile, Denmark, Finland, France, Greece, Hungary, India, Latvia, Mexico, Morocco, Saudi Arabia, Scotland, Singapore, Slovenia, Spain, and The Netherlands. Despite these smaller numbers, this geographic spread indicates diverse respondent origins, contributing to a varied and globally distributed dataset.

5.1.1.3 Survey design

Study 2 was a cross-sectional study set up on Qualtrics with the CIT integrated into a survey. It aimed to elicit a memory of the most recent incident of interpersonal felt distrust by a supervisor in study participants through CIT and then measure their emotional and behavioral reactions to the incident by helping them recall their experience at the time of the distrust occurrence. The survey started with CIT, which included a mix of scale-continuous and open-ended questions to help participants recall their most recent experience of felt distrust by a supervisor. For example, the participants were asked about the recency, frequency and severity of felt distrust by this supervisor, specific behaviors of the supervisor that made them feel distrusted, and how feeling distrusted affected their emotions and behaviors at work at the time of the occurrence. Such a detailed reflection on the past experience of felt distrust aimed to refresh individual experience of felt distrust in participants' memories with the goal to increase the degree of sincerity and accuracy of their responses to the actual survey questions that followed right after the memory recall. It was followed by Likert-scale multiple-item measures of the core variables

of interest, namely interpersonal felt distrust, work-related identity threat, anger-experience, shame-experience, CWB, psychological withdrawal, exemplification, anger-trait, shame-trait, narcissism, neuroticism, sex, race, and age. Notably, the instructions to the survey questions consistently linked each of the survey questions to the memory recall shared by the participant. For example, before the participants could respond on the behavioral measures, they were asked “[r]ecollecting the memory of your experience of having felt distrusted by your supervisor at work, to what extent did you engage in the following behaviors at work after you felt distrusted?”

5.1.1.4 Survey procedure

Study 2 consisted of two Qualtrics surveys, a pre-screening and a core survey, which were sent to participants a few hours apart. Prolific identified 73,127 eligible participants from all over the world who met the three initial screening criteria to participate in the pre-screening survey (these criteria being aged between 20-85 years old, employment experience, and experience being supervised at work). From this pool, I requested 800 participants for a pre-screening survey that consisted of five questions, including, “Have you ever felt distrusted by a supervisor at work?”. The participants were ultimately screened for two criteria, the first being currently employed on a part-time or full-time basis and the second being having had experience being supervised at work. Out of the 800 participants, 487 responded that they had experienced feeling distrusted by a supervisor at work.

Next, I invited 487 eligible participants to take part in the core survey of Study 2 and requested a 400-sample size. Of the 487 participants in the target population, 447 participated. 400 completed the survey and 47 returned the survey due to no consent, no prior experience, or not passing the validation checks in place. Of the 400 who completed the survey, 398 met all validation checks. The final sample (n=398) had no missing data and has been confirmed by Qualtrics to have

been completed by humans and not by bots. It took participants approximately 32 minutes (median time) to fill out the survey.

I approved the responses and sent a follow-up message to all 398 participants thanking them for their input. 46 participants (12%) voluntarily responded to me with individual messages reflecting on the opportunity to participate in the study, noting that it is an important topic, that the survey allowed them to reflect and process something for themselves, and that they found the survey interesting and engaging.

5.1.2 Ensuring data quality

The Study 2 Qualtrics survey incorporated security measures and data quality controls to ensure that responses thereto were both legitimate and reliable. Participants' Prolific IDs were captured via embedded data, restricting survey access to registered panel members. QCapture scores, which assess interaction patterns to detect bots, were collected for each response. Metadata included participants' IP addresses to prevent duplicate submissions. To further enhance response validity, the survey featured two attention checks and a seriousness check, aligning with best practices for online panel studies (Aguinis & Bradley, 2014; Atzmüller & Steiner, 2010; Palan & Schitter, 2018; Tristram et al., 2012b, 2012c, 2012a). Questions were enforced, meaning that the participants could not skip a question before proceeding to the next, and participants were informed in the consent form that they could voluntarily withdraw from the study at any time and still be fully compensated. 398 participants passed all validation, attention, and seriousness checks; as such, their responses were used in the data analysis.

Following an informed consent, the Qualtrics survey for the core part of the study provided the respondents with a definition of distrust³ and a question regarding whether they felt distrusted by a supervisor based on this description. Only participants who answered this question affirmatively were allowed to continue with the rest of the survey. These participants were then asked to recall their most recent experience of felt distrust by a supervisor and respond to a set of structured questions designed to help them immerse themselves in the memory of this incident. This methodological technique has been widely used in organizational science to recall survey participants' experiences and reactions to discrete workplace events before assessing their individual reactions through established psychological scales (Aquino et al., 2001; Korsgaard et al., 2002; Leavitt et al., 2022; Mitchell et al., 2015; Wellman et al., 2016). Following this recall exercise of felt distrust, participants were asked to respond to a set of survey questions reporting how feeling distrusted by a supervisor made them feel and act at the time of the incident's occurrence.

5.1.3 Study 2 Measures

Study 2 included the same measures of Interpersonal Felt Distrust, Work-related Identity Threat, Anger-Experience, Shame-Experience, Moral Identity Centrality, Anger-Trait, Shame-Trait, Narcissism, and Neuroticism as Study 1. Three new constructs were also measured in Study 2, namely the behavioral outcomes: Counterproductive Work Behavior (CWB) (toward one's organization), Psychological Withdrawal, and Exemplification.

³ The definition provided was the following: "When someone distrusts you, they think you have bad intentions towards them. They think your words, decisions or behaviors may harm them in some way." This definition is based on the literature, and particularly on Lewicki et al., 1998, and Sitkin & Bijlsma-Frankema, 2018.

5.1.3.1 Dependent Variables

Counterproductive Work Behavior (CWB). CWB was measured through a sub-scale of organizational deviance from Bennett & Robinson's (2000) measure of workplace deviance, excluding the four items of psychological withdrawal that are measured as a separate construct in this dissertation. The scale included eight items, including such items as: “Taken property from work without permission” and “Neglected to follow your boss’s instructions”. The Cronbach’s alpha in this study was .82. The Omega 3 (Flora, 2020; Fornell & Larcker, 1981) composite reliability was 0.80.

Psychological Withdrawal Behavior (PWB). I used Lehman & Simpson's (1992) 8-item scale to assess psychological withdrawal. The participants were asked to respond to what extent they engaged in the following behaviors at work after experiencing felt distrust, with scale anchors ranging from 1 = never to 7 = almost always (defined as “in more than 90% of the chances I could have”). Sample items included “Thoughts of being absent” and “Chat with co-workers about nonwork topics”. Cronbach’s alpha in this study was .86 and the Omega 3 composite reliability was 0.85.

Exemplification. This construct was measured using Bolino & Turnley's (1999) 4-item measure of exemplification on a 7-item Likert scale, with scale anchors ranging from 1 = never to 7 = almost always (defined as “in more than 90% of the chances I could have”). I have revised two of the scale items to be more appropriately suited to the online or hybrid work environment, for instance, “Start work early to look dedicated [used to be “Arrive at ...”]” and “Work at night or on weekends to show that you are dedicated [used to be “Come to the office...”]”. Other sample item questions include: “Stay at work late so people will know you are hard working” and “Try to

appear busy, even at times when things are slower”. Cronbach’s alpha in this study was .83. Omega 3 composite reliability was 0.84.

For CWB, PWB, and Exemplification, the original scale anchors—ranging from "never" to "daily"—were adjusted to improve interpretability and response accuracy. The initial frequency-based anchors (e.g., "twice a year," "monthly") could be less practical, as behavioral occurrences related to these constructs may not follow rigid time-based patterns. Instead, a probability-based scale was introduced, which reflects how often participants engaged in a behavior when they had the opportunity (e.g., "rarely" for less than 10% of possible instances, "frequently" for about 70%). This modification ensures that responses align with actual behavioral tendencies rather than predetermined timeframes, making them more reflective of individual experiences.

5.1.3.2 Mediating Variables

Work-Related Identity Threat. This variable was measured in the same way as in Study 1, with a revised 8-item scale of identity threat from Kroeper et al. (2022) that asked participants: “Thinking about the experience of feeling distrusted by a supervisor at work, which you have just described, how much did you worry back then that: [...]” followed by the item questions on a 5-point Likert type scale (1 = Not at all; 5 = Extremely). Cronbach’s alpha in this study was .85. Omega 3 composite reliability was .81.

Anger-Experience. Just like in Study 1, Anger-Experience was measured with a 6-item Hostility subscale from Watson & Clark (1994). Participants were asked: “Please indicate to what extent you felt these different feelings and emotions when you felt distrusted by your supervisor at work”, followed by the item questions on a 5-point Likert type scale (1 = Not at all; 5 = Extremely). Cronbach’s alpha in this study was .83. Omega 3 composite reliability was .84.

Shame-Experience. This variable was measured with a 5-item Shame subscale of the State Shame and Guilt Scale (SSGS; Marschall et al. (1994), just like in Study 1. The participants were told: “The following are some statements which may or may not describe how you were feeling when you felt distrusted by your supervisor at work. Please rate each statement on a scale below”, followed by the item questions on a 5-point Likert type scale (1 = Not at all; 5 = Extremely). Cronbach’s alpha in this study was .88. Omega 3 composite reliability was also .88.

5.1.3.3 Moderating Variable

Moral Identity Centrality (MIC). Just like in Study 1, this variable was measured using the 5-item internalization subscale of Aquino & Reed's (2002) self-importance of moral identity measure (responses on a 5-point Likert scale ranging from 1 = Strongly disagree to 5 = Strongly agree). Cronbach’s alpha in this study was .76. Omega 3 composite reliability was .77.

5.1.3.4 Independent Variable

Interpersonal Felt Distrust. Interpersonal felt distrust by a supervisor at work was recalled through a critical incident technique (CIT) (Flanagan, 1954) and then measured by a 4-item scale of interpersonal felt distrust by a supervisor, which was a revised version of the feeling of distrust scale from Mooijman et al. (2017), the same as the one used in Study 1.

The Likert scale ranged from 1 = Not at all to 7 = Extremely. The Cronbach’s alpha in this study was .78, Omega 3 composite reliability was .81.

5.1.3.5 Control Variables

Following the rationale explained in section 4.2.3.5 Control Variables and for consistency of model assessment across the two studies, the same set of controls was included and operationalized in the same way in Study 2 as in Study 1: In Study 2, Shame-Trait, Anger-Trait, Narcissism, Neuroticism, Sex, Race, and Age.

Shame-Trait. This variable was measured through a 4-item Shame-Negative Self-Evaluation (NSE) subscale of the Guilt and Shame Proneness Scale (GASP) (T. R. Cohen et al., 2011) with responses on a 5-point Likert scale from 1 = Extremely unlikely to 5 = Extremely likely. Cronbach's alpha in this study was .78. Omega 3 composite reliability was also .78.

Anger-Trait. Disposition to experience anger was measured using a 10-item, two-dimensional STAXI Trait Anger (T-Anger) Scale (Forgays et al., 1997) on a 4-point Likert scale, ranging from 1 = Almost never to 4 = Almost always. Cronbach's alpha in this study was .83. Omega 3 composite reliability was .71.

Neuroticism. This variable was measured using a 12-item neuroticism subscale of the Eysenck Personality Questionnaire–Brief Version (EPQ-BV) (Sato, 2005), a shortened and validated version of the Eysenck Personality Questionnaire–Revised (EPQ-R), with responses on a 5-point Likert-type scale ranging from 1 = Not at all and 5 = Extremely. Cronbach's alpha in this study was .93. Omega 3 composite reliability was also .93.

Narcissism. Narcissism was measured through the 16-item The Narcissistic Personality Inventory-16 (NPI-16) (Ames et al., 2006), a shortened form of the NPI-40 (Raskin & Hall, 1979). Cronbach's alpha for Narcissism in this study was .78. Omega 3 composite reliability was .83. Since Narcissism is a construct with categorical indicators, Zumbo et al.'s (2007) ordinal alpha was also calculated and was .88.

Sex. Participants' sex has been operationalized as sex at birth (male or female).

Age. Age was measured as a continuous variable by asking respondents the following question: "Please indicate your age (in number of full years, e.g. 26 or 35):18-25, 26-35, 36-45, 46-55, 56-65, over 65".

Race. Race has been measured as a categorical variable by asking the question: “Please indicate your race/ethnicity: [...]”. Response options were in line with Prolific categories for Ethnicity (being White, Black, Asian, Mixed-race, and Other).

5.2 Study 2 Data Analysis Strategy

Data was analyzed in R in a series of steps. First, it was explored for quality, univariate, bivariate, and multivariate normality, and outliers to ensure appropriate assumptions for structural equation modeling (SEM). The reliability of measures was assessed through Cronbach’s alpha and Omega 3 composite coefficients (Flora, 2020). To assess Common Method Variance (CMV), Harman’s single-factor test was conducted using principal component analysis (PCA) in R, ensuring that no single factor accounted for the majority of any variance. Additionally, a nested model comparison using ANOVA evaluated whether a one-factor model provided a comparable fit to the full model, with results indicating that the full model fit significantly better. Next, the hypotheses were tested using path analysis within SEM in R with the *lavaan* package. All core variables were treated as composite averages of their respective indicators. Finally, a post hoc power analysis was performed using the *semPower* package to verify that the model had achieved sufficient statistical power to detect meaningful effects.

5.3 Study 2 Results

5.3.1 Assessing Data Quality

Survey participants’ mean Q-captcha score was .94, which indicates that the survey was completed by humans and not by bots. 398 respondents responded affirmatively to the question of whether they had ever felt distrusted by a supervisor at work, and passed attention, honesty, and seriousness checks. Each of the 398 respondents provided detailed and meaningful written responses to all the open-ended questions within the memory recall portion of the survey. Since all

the survey questions were enforced, none of the survey questions were skipped. As such, Study 2 data set does not have missing data.

5.3.2 Assessing Data for Normality

Following the guidelines for assessing data normality (Oppong & Agbedra, 2016), I assessed 13 core continuous variables for normality using both visual and formal statistical methods. While boxplots and histograms appeared normally distributed, formal statistical tests for univariate normality, such as the Shapiro-Wilk test, Lilliefors test (Kolmogorov-Smirnov with estimated parameters), Anderson-Darling test, and Royston's test, consistently indicated deviations from normality across most variables. Bivariate scatterplots revealed some non-linear patterns and heteroscedasticity, while Mardia's test for multivariate normality showed significant skewness and kurtosis, confirming violations of multivariate normality assumptions.

Given these results, I selected the Maximum Likelihood Robust (MLR) estimator for SEM to account for non-normality in both parameter estimation and model fit statistics. This estimator accounts for deviations from multivariate normality, providing more reliable parameter estimates and model fit indices in the presence of outliers or non-normal data distributions.

5.3.3 Assessing Data for Linearity

Three regression models at the core of the SEM structural model, which respectively predicted CWB, PWB, and Exemplification, were assessed for linearity. Visual diagnostics of multivariate relationships, including residual plots and scatterplots with loess lines, suggested the relationships between predictors and outcomes to be approximately linear across all three models. Ramsey's RESET test was applied to each of the three outcome variables to evaluate the assumption of linearity in the regression models. The test for the CWB model was statistically significant ($\text{RESET} = 7.99, p < .001$), indicating potential non-linearity or model misspecification.

In contrast, the tests for the PWB and Exemplification models were not significant ($p = .438$ and $p = .540$, respectively), suggesting that the linearity assumption is reasonably met for these models. These results imply that while the PWB and Exemplification models appear to be appropriately specified, the CWB model may benefit from further exploration of non-linear terms or alternative model specifications. While Ramsey's RESET test indicated potential non-linearity in the CWB model, the absence of strong visual patterns and the non-significant RESET results for the PWB and Exemplification models support the conclusion that the linearity assumption is reasonably met overall.

5.3.4 Assessing Data for Outliers and Influential Cases

To identify potential outliers and influential observations, Mahalanobis distance was calculated using all 13 continuous variables in the dataset. Based on a chi-square cut-off at the 0.001 significance level with 13 degrees of freedom, 30 observations out of 398 (or approximately 7.5%) were flagged as potential multivariate outliers and the chi-square critical values for the lower and upper tails were 24.50601 and 60.59215, respectively. This proportion is slightly higher than expected under normal distribution assumptions, suggesting the presence of some atypical cases. As a second test, Cook's distance was computed for three separate regression models predicting behavioral outcomes. For the model predicting CWB, 24 observations exceeded the common influence threshold ($4/n$). Similarly, 23 influential cases were identified in the PWB regression and 17 in the Exemplification regression. These findings indicate that a subset of cases may exert disproportionate influence on model estimates and thereby warrant further examination.

To evaluate the influence of outliers and high-leverage observations on the regression predicting CWB, three models were compared: the full model, a model excluding multivariate outliers identified via Mahalanobis distance, and a model excluding influential cases identified

through Cook's distance. In the full model, the predictors explained approximately 21.2% of the variance in CWB ($R^2 = 0.212$), with significant direct effects from Moral Identity Centrality (MIC), Anger-Trait, Shame-Trait, and Narcissism. After removing 30 Mahalanobis outliers, model fit decreased slightly ($R^2 = 0.179$), though the residual standard error improved, and Shame-Experience became a significant predictor. A third model excluding 24 influential cases identified by Cook's distance showed a similar pattern ($R^2 = 0.175$), with consistent significance for MIC, Shame-Experience, Anger-Trait, Shame-Trait, and Narcissism. Across the three models, the core effects remained relatively stable, though effect sizes and significance levels shifted slightly, suggesting that while outliers and influential cases had some impact, the overall model was robust.

Similarly, three versions of the model were compared to assess the robustness of the regression model predicting Psychological Withdrawal Behavior (PWB): the full model, a model excluding multivariate outliers identified via Mahalanobis distance, and a model excluding influential observations identified through Cook's distance. The full model explained approximately 18.3% of the variance in PWB ($R^2 = 0.183$), with significant direct effects from Identity Threat, Anger-Trait, and Age. After removing 30 Mahalanobis outliers, model fit slightly decreased ($R^2 = 0.166$); and while Identity Threat and Anger-Trait remained significant, Age and Neuroticism approached significance. In contrast, the model excluding 23 influential cases showed an improved fit ($R^2 = 0.204$), with Identity Threat and Anger-Trait again emerging as significant predictors and with Age nearing significance. These results suggest that while Mahalanobis outliers slightly reduced model performance, the influential cases identified by Cook's distance may have been suppressing the model's explanatory power. Overall, the findings for Identity Threat and Anger-Trait were consistent across models, supporting their robustness as predictors of PWB.

The same approach was applied to evaluate the stability of the regression model predicting Exemplification, comparing the full model, a model excluding multivariate outliers identified via Mahalanobis distance, and a model excluding influential observations identified through Cook's distance. The full model explained approximately 14.4% of the variance in Exemplification ($R^2 = 0.144$), with Identity Threat and Narcissism emerging as significant predictors. After removing 30 Mahalanobis outliers, model fit remained nearly unchanged ($R^2 = 0.141$), though the significance of Narcissism slightly decreased and Anger-Trait lost marginal significance. In contrast, the model excluding 17 influential cases showed an improved fit ($R^2 = 0.181$), with Identity Threat and Narcissism remaining significant and with Anger Experience and Neuroticism approaching significance. These results suggest that while Mahalanobis outliers had a minimal impact on the model, the influential cases identified by Cook's distance may have been suppressing meaningful variance, particularly in relation to emotional traits. Overall, Identity Threat and Narcissism were consistently robust predictors of Exemplification across all models.

The presence of multivariate outliers and influential cases identified through Mahalanobis distance and Cook's distance diagnostics supported my choice to conduct the SEM using robust Maximum Likelihood estimator (MLR). The use of MLR ensures that the conclusions drawn from the SEM are not unduly influenced by extreme observations or violations of distributional assumptions, thereby enhancing the validity of the findings.

5.3.5 Descriptives

To evaluate the distribution and characteristics of the dataset, descriptive statistics were computed for all continuous variables (see Table 14). Several key variables exhibited near-normal distributions, including Individual Felt Distrust ($M = 4.55$, $SD = 1.27$) and Anger-Experience ($M = 3.30$, $SD = 0.85$), which both showed minimal skewness. Conversely, Counterproductive Work

Behavior (CWB; $M = 1.68$, $SD = 0.77$) had high positive skewness (2.04) and kurtosis (5.23), indicating a strong right-tail distribution. Similarly, Moral Identity Centrality ($M = 4.54$, $SD = 0.59$) and Shame-Trait ($M = 4.26$, $SD = 0.78$) exhibited negative skewness (-1.90 and -1.70, respectively), suggesting higher concentration of responses at the upper end of the scale.

Males were twice as likely to be distrusted by Males (149) than by Females (69), while Females were more likely to be distrusted by Females (103) than by Males (77). Among White participants (197/398), 91.4% (180 out of 197) reported feeling distrusted by White supervisors compared to 4.1% by Black supervisors, 1.5% by Asian supervisors, 1.0% by Mixed-race supervisors, and 2.0% by Other supervisors. Among Black participants (145/398), 48.3% (70 out of 145) reported felt distrust from Black supervisors and 39.3% (57 out of 145) from White supervisors; in contrast, felt distrust from Asian (2.1%), Mixed-race (6.9%), and Other (3.4%) supervisors was less common. Asian participants (27/398) most frequently reported distrust from Asian supervisors (40.7%), followed by White supervisors (51.9%), while distrust from other groups was minimal. Mixed-race participants (14/398) reported distrust from Mixed-race supervisors most often (50.0%), though White supervisors (35.7%) also accounted for a notable portion of distrusters. Finally, Other participants (15/398) displayed a relatively even distribution of distrust across White (53.3%) and Asian (20.0%) supervisors, with remaining categories contributing less.

Table 14*Study 2. Means, Standard Deviations, and Other Descriptive Statistics for Continuous Variables*

#	Variable	Mean	Standard Deviation (sd)	Median	Minimum	Maximum	Range	Standard Error (se)
1	Interpersonal Felt Distrust	4.55	1.27	4.50	1.25	7.00	5.75	0.06
2	Work-related Identity Threat	3.61	0.81	3.75	1.00	5.00	4.00	0.04
3	Anger-Experience	3.30	0.85	3.33	1.17	5.00	3.83	0.04
4	Shame-Experience	3.35	1.06	3.60	1.00	5.00	4.00	0.05
5	Counterproductive Work Behavior	1.68	0.77	1.45	1.00	5.70	4.70	0.04
6	Psychological Withdrawal	3.47	1.18	3.38	1.00	6.62	5.62	0.06
7	Exemplification	2.83	1.40	2.50	1.00	7.00	6.00	0.07
8	Moral Identity Centrality (MIC)	4.54	0.59	4.80	1.00	5.00	4.00	0.03
9	Shame-Trait	4.26	0.78	4.50	1.00	5.00	4.00	0.04
10	Anger-Trait	1.96	0.53	2.00	1.00	3.80	2.80	0.03
11	Neuroticism	2.68	0.90	2.62	1.00	5.00	4.00	0.05
12	Narcissism	0.25	0.20	0.19	0.00	0.88	0.88	0.01
13	Age	35.72	10.98	33.00	20.00	75.00	55.00	0.55

Note: n = 398

A Wilcoxon rank-sum test was conducted to examine whether the recency of interpersonal felt distrust experience (not reported in the table of Descriptives but provided by the participants in the CIT section), categorized as “within the past year” versus “more than a year ago”—was associated with differences in reported interpersonal felt distrust scores. The test revealed no statistically significant difference between groups, $W = 14,988$, $p = .108$, suggesting that the timing of the distrust experience did not influence the intensity of interpersonal felt distrust recalled. However, when examining perceived work-related identity threat, the Wilcoxon test indicated a statistically significant difference by recency, $W = 14,480$, $p = .036$. This finding suggests that individuals who experienced distrust more recently may perceive higher levels of identity threat in the workplace than those whose experiences occurred further in the past. One interpretation

aligns with cognitive-affective processing theories, which suggest that emotionally charged events are more cognitively accessible when temporally close, leading to heightened threat sensitivity (Forgas, 1995; Robinson & Clore, 2002). The salience of the recent experience may reactivate threat appraisals, amplifying work-related identity threat. Alternatively, research on psychological distancing and meaning reconstruction suggests that as time passes, individuals may engage in reframing and self-protective reinterpretation strategies, thereby reducing the psychological intensity of the experience (Wilson & Gilbert, 2005; Park, 2010).

5.3.6 Validity and Reliability of Measures

Study 2 employs path analysis, wherein individual scale items are aggregated into a single composite score for each variable rather than treated as latent constructs. As a result, the individual factor loadings of each item are less critical when compared to a traditional Confirmatory Factor Analysis (CFA), wherein items serve as reflective indicators of an underlying latent factor (L. Deng & Yuan, 2023). However, since each variable is measured using multiple survey items, a formal CFA was conducted using the *lavaan* package in R using a Maximum Likelihood Robust (MLR) estimator to assess the reliability and validity of the measurement scales in the survey. The CFA results are documented in Appendix D, Tables D1 (for the core variables) and D2 (for control variables).

Overall, the standardized factor loadings for all variables were moderate and statistically significant, supporting the relevance of each item in capturing the intended construct. While Composite Reliability (CR) and Average Variance Extracted (AVE) fell below conventional thresholds, the Cronbach's alpha and omega reliability coefficients were sufficiently high, indicating acceptable internal consistency across variables. Given that each construct is analyzed as a composite score rather than as a latent factor, these reliability considerations do not undermine

the validity of the path analysis approach and the results remain methodologically sound for the purposes of the intended research objectives.

5.3.7 Correlational Analysis

Table 15 presents a correlation matrix summarizing relationships between the continuous variables in the model, with Cronbach's alpha values displayed along the diagonal to indicate internal consistency. As is visible from Table 15, Interpersonal Felt Distrust and Work-Related Identity Threat exhibit a moderate positive correlation, indicating that individuals experiencing interpersonal distrust also perceive threats to their identity. Additionally, Work-Related Identity Threat is significantly correlated with both Anger-Experience and Shame-Experience.

Further examination of emotional experiences and behavioral tendencies reveals other meaningful associations. Anger-Experience has a weak positive correlation with Counterproductive Work Behavior (CWB), Psychological Withdrawal (PWB), and Exemplification. Similarly, Shame-Experience shows a weak positive correlation with CWB, Psychological Withdrawal, and Exemplification. CWB and Psychological Withdrawal are moderately associated, suggesting shared patterns of disengagement. The Cronbach alpha values, ranging from 0.76 to 0.93, confirm good internal consistency across measures, supporting their reliability in assessing core constructs.

Table 15*Study 2. Intercorrelations of Continuous Variables*

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Interpersonal Felt Distrust (.78)												
2. Work-Related Identity Threat	.56***	(.85)										
3. Anger-Experience	.52***	.54***	(.83)									
4. Shame-Experience	.43***	.70***	.51***	(.88)								
5. Counterproductive Work Behavior	.05	.08	.14**	.12*	(.82)							
6. Psychological Withdrawal	.14**	.29***	.20***	.26***	.48***	(.86)						
7. Exemplification	.21***	.30***	.18***	.25***	.32***	.24***	(.83)					
8. Moral Identity Centrality	.13**	.14**	.11*	.16***	-.21***	-.05	.00	(.76)				
9. Shame-Trait	.12*	.25***	.11*	.24***	-.19***	.06	.00	.38***	(.78)			
10. Anger-Trait	.15**	.18***	.36***	.21***	.34***	.32***	.21***	-.11*	-.01	(.83)		
11. Neuroticism	.07	.22***	.28***	.31***	.19***	.28***	.18***	-.06	.04	.6***	(.93)	
12. Narcissism	.11*	-.02	.04	-.09	.25***	.09	.14**	-.12*	-.15**	.13**	-.10	(.78)
13. Age	.00	-.01	-.04	-.15**	-.12**	-.16**	-.08	.08	.03	-.13**	-.19***	-.09

Note: n = 398. Entries on the main diagonal are Chronbach's alpha. Age is a one-item scales, internal consistency was not computed. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.00$

Since the current sample includes only individuals who felt distrusted by a supervisor (IFD), the sample of Study 2 represents a range-restricted sample where correlations of directly range-restricted interpersonal felt distrust variable with core variables in the model are potentially attenuated (Sackett et al., 2007; Sackett & Yang, 2000). To account for attenuation due to direct range restriction on the IFD variable, I applied Thorndike's Case 2 correction (Carretta & Ree, 2022; Sackett et al., 2007; Sackett & Yang, 2000) to all observed correlations involving IFD. This procedure is in line with best practice recommendations by Aguinis et al. (2021).

First, I determined that the pre-screening sample of 800 participants who were initially surveyed about whether they had ever experienced interpersonal distrust by a supervisor at work is essentially the full sample that could be used for correction for attenuation. Of these, 487 individuals (approximately 61%) reported having experienced such distrust, while 313 individuals (approximately 39%) reported no such experience. In the core phase of Study 2, only those who

endorsed having experienced distrust were recontacted and invited to complete additional measures. From this subsample, 400 participants responded, and 398 were retained for final analysis after quality screening. As a result, the analytic dataset was restricted to individuals with prior distrust experiences, introducing direct range restriction on the IFD variable.

Second, to address the potential attenuation of observed correlations caused by this restriction, a range restriction correction (Thorndike's Case 2) was applied correction (Carretta & Ree, 2022; Sackett et al., 2007; Sackett & Yang, 2000). The unrestricted standard deviation of IFD was estimated by combining the real IFD scores from the 398 participants with an assumed score of 1 (Not at all distrusted) for the 313 participants who had not experienced distrust, yielding a full-sample estimate based on 711 individuals. This procedure enabled more accurate estimation of IFD's associations with other variables in the model by accounting for the reduced variance inherent in the restricted analytic sample.

As shown in Table 16, corrected correlations are consistently stronger than observed values, suggesting that restriction to participants with prior distrust experiences substantially underestimated the strength of IFD's relationships with other variables. This adjustment provides a more accurate estimate of the effect sizes expected in the full population.

Table 16*Study 2. Comparison Table: Observed vs. Corrected Correlations*

Path	Observed <i>r</i>	Corrected <i>r</i>
IFD --> WRIT	0.56	0.73
IFD --> AngerExp	0.52	0.69
IFD --> ShameExp	0.43	0.60
IFD --> CWB	0.05	0.08
IFD --> PWB	0.14	0.22
IFD --> Exemp	0.21	0.32
IFD --> MIC	0.13	0.20
IFD --> ShameTr	0.12	0.19
IFD --> AngerTr	0.15	0.23
IFD --> Neuro	0.07	0.11
IFD --> Narci	0.11	0.17
IFD --> Age	0.00	0.00

Note: Corrected correlation values are based on Thorndike's Case 2 formula using an estimated unrestricted SD of 2.00 and restricted SD of 1.27 for predictor IFD. Abbreviations: IFD - Interpersonal Felt Distrust; WRIT - Work-Related Identity Threat; AngerExp - Anger-Experience; ShameExp - Shame-Experience; CWB - Counterproductive Work Behavior; PWB - Psychological Withdrawal Behavior; Exemp - Exemplification; MIC - Moral Identity Centrality; ShameTr - Shame-Trait; AngerTr - Anger-Trait; Neuro - Neuroticism; Narci - Narcissism.

5.3.8 Path Analysis within Structural Equation Modeling (SEM)

The structural equation model estimated a total of 75 free parameters, including regression paths, variances, and covariances. With a sample size of 398 participants, this yields a case-to-parameter ratio of approximately 5.3:1. This ratio meets the commonly recommended minimum threshold of 5:1 for SEM, indicating that the model is adequately powered for the purposes of reliable estimation. To further account for potential non-normality and the presence of outliers identified in preliminary diagnostics, the model was estimated using robust maximum likelihood (MLR), which provides adjusted standard errors and test statistics. This approach enhances the robustness and validity of the parameter estimates and model fit indices.

To prepare the data for moderation analysis within the SEM framework, the moderator variable (MIC) was mean-centered (i.e., the sample mean was subtracted from each individual score), resulting in a variable with a mean of zero. Mean-centering the moderator is a common

practice when creating interaction terms, as it helps to reduce multicollinearity between the moderator and the interaction term. Additionally, it enhances the interpretability of the model by allowing its main effects to be interpreted at the average level of the moderator. Importantly, this transformation does not alter the underlying relationships among variables, but rather improves the stability and clarity of the parameter estimates in models involving moderation.

5.3.8.1 Assessing Common Method Variance

To assess Common Method Variance (CMV), Harman's single-Factor test was conducted following the procedure outlined by Tehseen et al. (2017). Principal Component Analysis (PCA) was performed on all study variables to determine whether a single factor accounted for the majority of variance, which would indicate potential CMV. The first factor explained 15.04% of the total variance, which falls well below the 50% threshold for concern. Additionally, cumulative variance across multiple factors showed a more distributed variance structure, suggesting that CMV was not a pervasive issue in the dataset. Based on these results, the presence of CMV was deemed unlikely, supporting the validity of the measurement model (Tehseen et al., 2017).

As an additional test to examine CMV, a nested model comparison using ANOVA was conducted. This approach assessed whether a one-factor model—which would indicate significant CMV—provided a comparable fit to the full measurement model. The analysis involved comparing models with different numbers of latent factors using chi-square tests, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Root Mean Square Error of Approximation (RMSEA). The results showed that the full 16-factor model provided a significantly better fit than the one-factor alternative ($\chi^2 = 7,552.23$, RMSEA = 0.049) when compared to the one-factor model ($\chi^2 = 15,865.18$, RMSEA = 1.626). Since collapsing all factors

into a single latent construct led to a substantial deterioration in model fit, this analysis indicated that CMV was not a significant concern, supporting the validity of the measurement model.

5.3.8.2 Hypotheses Testing Results (original SEM model)

A structural equation model (SEM) was estimated using maximum likelihood with robust standard errors (MLR) to examine the relationships among interpersonal felt distrust, work-related identity threat, emotional reactions, and workplace behaviors in a sample of 398 participants. The model demonstrated acceptable fit to the data ($\chi^2 = 16.53$, $df = 6$, $p = .011$; CFI = 0.990, TLI = 0.877, RMSEA = 0.065, SRMR = 0.008). These indices suggest that the hypothesized model adequately represents the observed data.

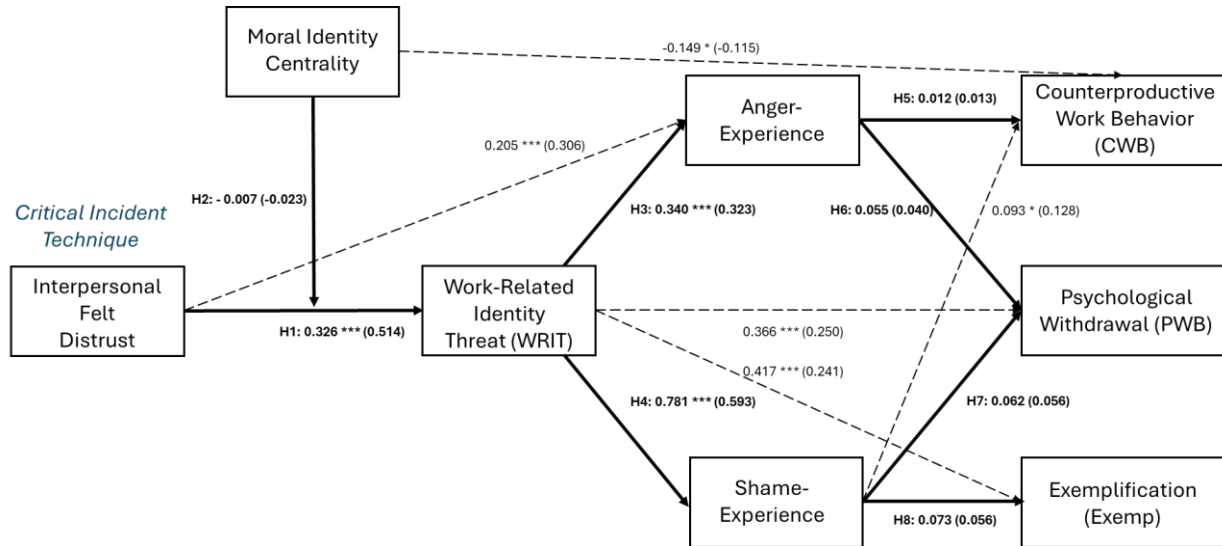
To evaluate the robustness of the structural equation model, a comparative analysis was conducted between the full model and a version excluding influential cases identified through Cook's distance across three preliminary regression models. All three models estimated an identical number of parameters (75), allowing for a direct comparison of fit indices. The full model demonstrated excellent fit (CFI = 0.990, RMSEA = 0.064, SRMR = 0.008). The model excluding influential cases yielded nearly identical results (CFI = 0.989, RMSEA = 0.069, SRMR = 0.008), indicating minimal impact on overall model fit. These findings suggest that the structural model is stable and not unduly influenced by the identified cases. As such, the full dataset was retained for the final SEM analysis to preserve statistical power and generalizability.

The results of the regressions are presented in Figure 10. The path analysis of a full structural equation model provided partial support for the hypothesized relationships among interpersonal felt distrust, work-related identity threat, emotional responses, and workplace behaviors. The results of hypotheses testing are presented in Table 17, with unstandardized and standardized coefficients of the hypothesized direct and indirect effects.

As is visible in Figure 10 and Table 17, **H1** was supported: Interpersonal Felt Distrust significantly predicted Work-Related Identity Threat ($\beta = 0.514, p < .001$), indicating that an experience of interpersonal felt distrust by a supervisor is associated with an increase in threat to one's work-related identity. **H2** was not supported: The interaction between Interpersonal Felt Distrust and moral identity centrality did not significantly predict work-related identity threat ($\beta = -0.023, p = .905$), suggesting no moderating effect. **H3** and **H4** were supported: Work-Related Identity Threat significantly predicted both Anger-Expression ($\beta = 0.323, p < .001$) and Shame-Experience ($\beta = 0.593, p < .001$). Considering the finding in H1, the mediation pathways from Interpersonal Felt Distrust through Work-Related Identity Threat to Emotional responses has been supported. **H5** was not supported: Anger-Experience did not significantly predict Counterproductive Work Behavior (CWB) ($\beta = 0.013, p = .795$). **H6** and **H7** were not supported: Neither Anger-Experience ($\beta = -0.040, p = .521$) nor Shame-Experience ($\beta = 0.056, p = .432$) significantly predicted Psychological Withdrawal Behavior, indicating no mediation effects in these pathways. **H8** was not supported: Shame-Experience did not significantly predict Exemplification behavior ($\beta = 0.056, p = .397$), suggesting no mediation effect in this pathway either.

Figure 10

Study 2. SEM Hypothesis Testing Results (Original SEM Model)



Note: n=398; all variables observed. Bold paths are hypothesized relationships. Dashed paths are direct effects (not hypothesized but examined to test mediation). Values in each path: Hypothesis, unstandardized estimate, significance, standardized estimate in brackets. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$

Table 17

Study 2. Hypotheses Testing Results (Original SEM model)

Hypothesis	Path	Unstandardized Coefficient	Standardized Coefficient	P-value	Supported
H1	Interpersonal Felt Distrust → Work-Related Identity Threat [direct effect]	0.326	0.514	0.0	Yes
H2	Interaction between Interpersonal Felt Distrust and Moral Identity Centrality → Work-Related Identity Threat [moderation]	-0.007	-0.023	.905	No
H3	Interpersonal Felt Distrust → Work-Related Identity Threat → Anger-Experience [indirect effect; mediation]	0.111	0.166	0.0	Yes

H4	Interpersonal Felt Distrust → Work-Related Identity Threat → Shame-Experience [indirect effect; mediation]	0.255	0.305	0.0	Yes
H5	Work-Related Identity Threat → Anger-Experience → Counterproductive Work Behavior [indirect effect; mediation]	0.004	0.004	0.824	No
H6	Work-Related Identity Threat → Anger-Experience → Psychological Withdrawal [indirect effect; mediation]	-0.019	-0.013	0.508	No
H7	Work-Related Identity Threat → Shame-Experience → Psychological Withdrawal [indirect effect; mediation]	0.048	0.033	0.404	No
H8	Work-Related Identity Threat → Shame-Experience → Exemplification [indirect effect; mediation]	0.057	0.033	0.421	No

Mediation analysis was conducted using the *lavaan* package in R to examine indirect and direct effects. The model tested whether Work-Related Identity Threat, Anger-Experience, and Shame-Experience mediated the effects of Interpersonal Felt Distrust on workplace behaviors, namely CWB, Psychological Withdrawal, and Exemplification. Indirect effects were analyzed to determine mediation significance. The results are consistent with the results of Study 1 in that interpersonal felt distrust is positively associated with work-related identity threat (H1) and work-related identity threat is positively associated with anger and shame and mediates the relationship between interpersonal felt distrust and anger (H3) and shame (H4). In contrast to Study 1, the moderating effect of moral identity centrality was not significant in Study 2 (H2), and anger and

shame were not associated with increased levels of CWB, PWB, or Exemplification as hypothesized (H5-8).

5.3.8.3. Additional Findings (original SEM model)

As part of the overall SEM testing, direct effects were subsequently examined, revealing several significant relationships, which were not hypothesized in the Study. CWB was significantly associated with Anger-Trait (Unstandardized $\beta = 0.413$, $p < 0.001$) (i.e., a construct measuring dispositional anger as a control variable), Shame-Trait (Unstandardized $\beta = -0.142$, $p = 0.009$), and Narcissism (Unstandardized $\beta = 0.692$, $p = 0.001$). A positive and significant direct effect of Shame-Experience on CWB (Unstandardized $\beta = 0.093$, $p = 0.037$) suggests an alternative emotional pathway between Interpersonal Felt Distrust, Work-Related Identity Threat, Shame-Experience and CWB. Psychological Withdrawal was significantly associated with Work-Related Identity Threat (Unstandardized $\beta = 0.366$, $p < 0.001$), Race (being Non-White) (Unstandardized $\beta = -0.236$, $p = 0.049$), and Anger-Trait (Unstandardized $\beta = 0.458$, $p = 0.001$). Exemplification was significantly associated with Work-Related Identity Threat (Unstandardized $\beta = 0.417$, $p < 0.001$) and Narcissism (Unstandardized $\beta = 0.866$, $p = 0.013$). Having a significant positive association of narcissism with both CWB and Exemplification suggests that individuals high in narcissistic traits may simultaneously engage in both self-serving and image-enhancing behaviors.

These additional findings suggest that direct effects of work-related identity threat drive workplace behaviors more prominently than emotional mediation pathways. In other words, mediated by work-related identity threat, interpersonal felt distrust affects PWB (Unstandardized $\beta = 0.019$, $p < 0.001$) and Exemplification (Unstandardized $\beta = 0.136$, $p < 0.001$), but not CWB (Unstandardized $\beta = 0.008$, n.s.). This indirect effect on behavioral outcomes is not mediated by incidental anger or shame that arise following interpersonal felt distrust. Instead, the results suggest

that anger, shame, PWB, and Exemplification all emerge as outcomes of work-related identity threat following interpersonal felt distrust by one's supervisor. Interestingly, the proportional indirect effect of interpersonal felt distrust on psychological withdrawal through work-related identity threat was -6.33 , indicating that the indirect effect was over six times larger than the total effect, but in the opposite direction. This suggests a suppression effect, where the mediator reveals a relationship that is otherwise obscured in the total effect (MacKinnon et al., 2000). The effects are summarized in Table 18.

Table 18

Study 2. SEM: Direct, Indirect, Total and Proportional Indirect Effects in the Mediation Analysis (Original SEM Model)

Path	Direct Effect	Significance of Direct Effect	Indirect Effect	Significance of Indirect Effect	Total Effect	Proportional Indirect Effect
1 IFD → WRIT → Anger-Experience	0.205	***	0.111	***	0.316	0.351
2 IFD → WRIT → Shame-Experience	0.056		0.255	***	0.311	0.820
3 WRIT → Anger-Experience → CWB	0.026		0.004		0.030	0.133
4 WRIT → Anger-Experience → PWB	0.366	***	-0.019		0.347	0.055
5 WRIT → Shame-Experience → PWB	0.366	***	0.048		0.414	0.116
6 WRIT → Shame-Experience → Exemplification	0.417	***	0.057		0.474	0.120
7 IFD → WRIT → Anger-Experience → CWB	-0.037		0.005		-0.032	0.156
8 IFD → WRIT → Anger-Experience → PWB	-0.022		0.003		-0.019	0.158
9 IFD → WRIT → Shame-Experience → PWB	-0.022		0.003		-0.019	0.158
10 IFD → WRIT → Shame-Experience → Exemplification	0.059		0.004		0.063	0.063
11 IFD → WRIT → CWB [without emotion in mediation analysis]	-0.037		0.008		-0.029	-0.276
12 IFD → WRIT → PWB [without emotion in mediation analysis]	-0.022		0.019	***	-0.003	-6.330
13 IFD → WRIT → Exemplification [without emotion in mediation analysis]	0.059		0.136	***	0.195	0.697

Note: $n = 398$. Calculations of effects is based on unstandardized coefficients from the SEM model output. IFD - Interpersonal Felt Distrust; WRIT - Work-Related Identity; Threat; CWB - Counterproductive Work Behavior; PWB - Psychological Withdrawal

5.3.8.4 Post-Hoc Analysis (original SEM model)

To assess the adequacy of the sample size for detecting meaningful effects in the SEM, a post hoc power analysis was conducted using the *semPower* package in R (Moshagen & Bader, 2024). This analysis was based on the most conservative standardized path coefficients (std.all) detected in the model, ensuring a realistic estimate of power for smaller effect sizes. The power analysis employed F_0 as the effect size measure, representing global model fit discrepancy in SEM.

The effect size was set to 0.10, aligning with the smallest standardized coefficient observed in the model. The results indicated a statistical power of 85.56% (power = 0.8556; β = 0.1443), exceeding the conventional 0.80 threshold for sufficient power in SEM. The Root Mean Square Error of Approximation (RMSEA = 0.0365) indicates strong model fit (n = 398, df = 75, critical χ^2 = 96.2167, alpha/beta ratio = 0.3464). The implied alpha/beta ratio suggests a well-balanced model regarding Type I and Type II errors. These results confirm that the selected sample size was sufficient for detecting even the smallest significant effects, reinforcing the robustness of the analytical approach.

5.3.9 SEM Sensitivity Analysis

5.3.9.1 Accounting for Range Restriction

To account for the effects of range restriction in the measurement of Interpersonal Felt Distrust (IFD), a sensitivity analysis was conducted using a corrected covariance matrix. Specifically, correlations involving IFD were adjusted based on its unrestricted standard deviation (SD = 2.00) obtained from the full Study 2 sample (n = 711), following Thorndike's Case II formula (Caretta & Ree, 2022). The corrected correlation matrix was constructed by replacing attenuated IFD correlations with deattenuated values and then transformed into a covariance matrix using empirical standard deviations of all SEM variables. This corrected covariance matrix was used as input for a structural equation model identical in structure to the original, allowing for direct comparison of parameter estimates and significance patterns across models.

While the original SEM model (n = 398) demonstrated excellent statistical fit: $\chi^2(6)$ = 15.84, p = .015, RMSEA = 0.064, CFI = 0.990, TLI = 0.881—the corrected SEM model (n = 711), adjusted for range restriction in Interpersonal Felt Distrust (IFD), yielded a poorer overall fit: $\chi^2(6)$ = 93.46, p < .001, RMSEA = 0.143, CFI = 0.965, TLI = 0.563. However, this decline in statistical

fit must be weighed against gains in predictive validity of IFD and generalizability of the study findings to broader population of employees (Carretta & Ree, 2022). The original model was based solely on participants who reported experiencing IFD, introducing range restriction that attenuated key correlations. The corrected SEM addressed this by incorporating deattenuated correlations using the full sample's unrestricted variance in IFD, thereby capturing population-level dynamics. As a result, the corrected model unveiled theoretically predicted relationships that were previously masked—most notably, a significant interaction between IFD and moral identity centrality. Despite its lower fit indices, the corrected SEM offers a more accurate and externally valid representation of how interpersonal distrust functions within broader organizational contexts. Below I report the corrected standardized coefficients and compare them to the original (restricted) SEM model output. The results of hypotheses testing of the corrected SEM model are presented in Table 19 and Figure 11, with unstandardized and standardized coefficients of the hypothesized direct and indirect effects.

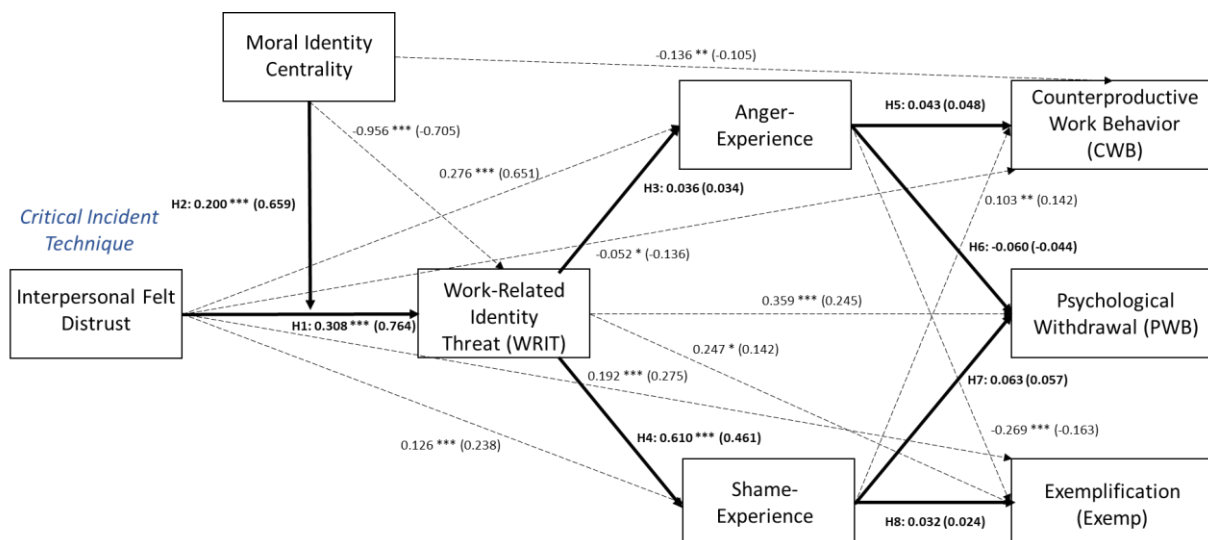
5.3.9.2 Testing Hypotheses Results (corrected SEM model)

Consistent with the original analysis, **Hypothesis 1 was supported**: IFD continued to significantly predict Work-Related Identity Threat (WRIT), and the path coefficient increased from $\beta = .514$ ($p < .001$) to $\beta = .764$ ($p < .001$), reinforcing the central theoretical claim that interpersonal felt distrust by supervisor strongly threatens an individual's work-related identity. Importantly, **Hypothesis 2 was now supported**: the interaction between IFD and Moral Identity Centrality (MIC) significantly predicted Work-Related Identity Threat ($\beta = .659$, $p < .001$), suggesting that individuals with higher moral identity centrality are more vulnerable to identity threat when experiencing interpersonal distrust. This interaction, previously suppressed under range-restricted conditions, became visible after correction.

The emotional outcomes proposed in Hypotheses 3 and 4 showed differential shifts. While **H4 remained supported**, with Interpersonal Felt Distrust (IDT) significantly predicting Shame-Experience ($\beta = .463, p < .001$), **H3 was no longer supported**: the path from Interpersonal Felt Distrust (IDT) to Anger-Experience became non-significant ($\beta = .034, p = .383$), suggesting that shame may be the more robust emotional consequence of identity threat under corrected conditions. Hypotheses 5–8 examined the direct effects of emotional experiences on behavioral outcomes. **All four hypotheses remained unsupported in the corrected model**: H5: Anger-Experience did not significantly predict Counterproductive Work Behavior (CWB) ($\beta = .048, p = .333$); H6: Anger-Experience was not associated with Psychological Withdrawal Behavior (PWB) ($\beta = -.044, p = .382$); H7: Shame-Experience did not predict PWB ($\beta = .057, p = .272$); H8: Shame-Experience did not significantly predict Exemplification ($\beta = .024, p = .643$).

Figure 11

Study 2. SEM Hypothesis Testing Results (Corrected SEM Model)



Note: $n=711$; all variables observed. Bold paths are hypothesized relationships. Dashed paths are not hypothesized but statistically significant relationships. Values in each path: Hypothesis #, unstandardized estimate, significance, standardized estimate in brackets. $***p < 0.001$, $**p < 0.01$, $*p < 0.05$, $p < 0.1$

Table 19*Study 2. Hypotheses Testing Results (Corrected SEM Model)*

Hypothesis	Path	Unstandardized Coefficient	Standardized Coefficient	P-value	Supported
H1	Interpersonal Felt Distrust → Work-Related Identity Threat [direct effect]	0.308	0.764	0.0	Yes
H2	Interaction between Interpersonal Felt Distrust and Moral Identity Centrality → Work-Related Identity Threat [moderation]	0.200	0.659	0.0	Yes
H3	Interpersonal Felt Distrust → Work-Related Identity Threat → Anger-Experience [indirect effect; mediation]	0.036	0.034	0.383	No
H4	Interpersonal Felt Distrust → Work-Related Identity Threat → Shame-Experience [indirect effect; mediation]	0.610	0.461	0.0	Yes
H5	Work-Related Identity Threat → Anger-Experience → Counterproductive Work Behavior [indirect effect; mediation]	0.043	0.048	0.333	No
H6	Work-Related Identity Threat → Anger-Experience → Psychological Withdrawal [indirect effect; mediation]	-0.060	-0.044	0.382	No
H7	Work-Related Identity Threat → Shame-Experience → Psychological Withdrawal [indirect effect; mediation]	0.063	0.057	0.272	No
H8	Work-Related Identity Threat → Shame-Experience → Exemplification [indirect effect; mediation]	0.032	0.024	0.643	No

5.3.9.3 Additional Findings (corrected SEM model)

Although the hypothesized emotion-to-behavior paths were non-significant, the corrected SEM model revealed new direct effects of IFD that bypass emotional mediators. Specifically, IFD significantly predicted CWB ($\beta = -.136, p = .031$) and Exemplification ($\beta = .275, p < .001$), suggesting direct behavioral responses to interpersonal felt distrust—retaliatory and image-enhancing—that are not channeled through emotional appraisal alone. Work-Related Identity Threat (WRIT) also demonstrated several significant direct effects on behavioral outcomes in the corrected model. Specifically, WRIT significantly predicted higher levels of Exemplification ($\beta = .142, p = .015$), suggesting that individuals experiencing identity threat may engage in image-enhancing impression management behaviors as a self-protective strategy following interpersonal felt distrust. Furthermore, IDT positively predicted Psychological Withdrawal Behavior ($\beta = .245, p < .001$), indicating that threat to work identity may also lead to disengagement from organizational demands. While IDT's prediction of Counterproductive Work Behavior (CWB) remained non-significant, these significant pathways to both withdrawal and strategic impression management highlight the behavioral manifestations of identity threat beyond its emotional consequences.

Altogether, these findings underscore the analytic value of correcting for range restriction when constructs are conditionally measured. The corrected model provided a more sensitive test of hypothesized relationships, revealing stronger identity threat effects, a previously masked moderation by moral identity centrality, and new direct behavioral pathways originating from IFD.

A mediation analysis conducted with the corrected SEM model ($n = 711$) revealed important shifts in the strength and structure of key pathways relative to the original model ($n = 398$). In both models, work-related identity threat (WRIT) significantly mediated the effect of

interpersonal felt distrust (IFD) on shame-experience, but this indirect effect was markedly stronger in the original model ($\beta = .255, p < .001$; proportional indirect = .820) compared to the corrected one ($\beta = .188$; proportional indirect = .599). However, the IFD $\rightarrow$ WRIT $\rightarrow$ anger-experience pathway lost significance and coefficient magnitude in the corrected model ($\beta = .011$; proportional indirect = .038) versus the original ($\beta = .111$; proportional indirect = .351), suggesting shame, not anger, emerged more clearly as the primary emotional pathway once measurement bias was addressed. Notably, the indirect effects from IFD to workplace behaviors (e.g., CWB, PWB, exemplification) through WRIT and emotions were small and non-significant in both models. Importantly, the corrected model revealed new significant direct effects of IFD on behaviors, specifically, on Exemplification ($\beta = .192, p < .001$) and CWB ($\beta = -.052, p = .031$). The indirect paths from Interpersonal Felt Distrust to Psychological Withdrawal and Exemplification via Work-Related Identity Threat alone were significant in both models (e.g., IFD $\rightarrow$ WRIT $\rightarrow$ PWB: $\beta = .111$; proportional indirect = 1.040; IFD $\rightarrow$ WRIT $\rightarrow$ Exemplification: $\beta = .076$; proportional indirect = 0.284 in the corrected model). The results of exploration of these direct and indirect effects across two models (range-restricted/original and unrestricted/full) suggest that, when modeling a less range-restricted population, identity threat strongly appears to be a driver of both emotional and behavioral responses to interpersonal felt distrust by supervisor, while emotion-driven behavioral effects appear attenuated and some of the direct effects are concealed. The results of the direct and indirect paths for the corrected SEM model are presented in Table 20.

Table 20

Study 2. SEM: Direct, Indirect, Total and Proportional Indirect Effects in the Mediation Analysis (Corrected SEM Model)

Path	Direct Effect	Significance of Direct Effect	Indirect Effect	Significance of Indirect Effect	Total Effect	Proportional Indirect Effect
1 IFD → WRIT → Anger-Experience	0.276	***	0.011		0.287	0.038
2 IFD → WRIT → Shame-Experience	0.126	***	0.188	***	0.314	0.599
3 WRIT → Anger-Experience → CWB	0.060		0.002		0.062	0.032
4 WRIT → Anger-Experience → PWB	0.359	***	-0.002		0.357	0.006
5 WRIT → Shame-Experience → PWB	0.359	***	0.038		0.397	0.096
6 WRIT → Shame-Experience → Exemplification	0.247	***	0.020		0.267	0.075
7 IFD → WRIT → Anger-Experience → CWB	-0.052	*	0.001		-0.051	0.020
8 IFD → WRIT → Anger-Experience → PWB	-0.005		-0.001		-0.006	0.167
9 IFD → WRIT → Shame-Experience → PWB	-0.005		0.012		0.007	1.690
10 IFD → WRIT → Shame-Experience → Exemplification	0.192	***	0.006		0.198	0.030
11 IFD → WRIT → CWB [without emotion in mediation analysis]	-0.052	*	0.018		-0.034	-0.544
12 IFD → WRIT → PWB [without emotion in mediation analysis]	-0.005		0.111	***	0.106	1.040
13 IFD → WRIT → Exemplification [without emotion in mediation analysis]	0.192	***	0.076	***	0.268	0.284

Note: n = 711. Calculations of effects is based on unstandardized coefficients from the SEM model output. IFD - Interpersonal Felt Distrust; WRIT - Work-Related Identity Threat; CWB - Counterproductive Work Behavior; PWB - Psychological Withdrawal

5.3.9.4 Post-Hoc Analysis (corrected SEM model)

Post-hoc power analysis using RMSEA = 0.05 as the target fit index revealed 67.7% power, indicating somewhat limited ability to detect close model fit given the low degrees of freedom (df = 6) and sample size (N = 711). In contrast, power to detect model misfit at RMSEA = 0.143—the value observed in the corrected SEM—exceeded 99%, with a near-zero Type II error rate ($\beta \approx 0$). This indicates that the poor model fit was not the result of sampling error or insufficient power but rather reflects genuine structural misfit in the model as specified. In other words, the significant chi-square test and elevated RMSEA are statistically trustworthy signals, not artifacts of underpowered testing. Despite the model's weaker fit indices, the corrected SEM is theoretically justified and analytically valuable: it corrects for range restriction in IFD, incorporates a more representative sample, and reveals several theoretically meaningful relationships—such as the moderating effect of moral identity centrality and direct effects of IFD on behavioral outcomes—

that were either attenuated or undetected in the original model. On this basis, interpretation of Study 2 findings focuses on the results of the corrected model. However, it is important to interpret the findings from the corrected model, i.e., the path coefficients, indirect effects, and mediation results, with caution, as the poor model fit indicates that the specified structure does not adequately capture the underlying data patterns, even after accounting for range restriction.

5.4 Study 2 Discussion

These findings bring into focus key differences between the original SEM model (based on the Interpersonal Felt Distrust (IFD) -experienced subsample, $N = 398$) and the corrected SEM using a covariance matrix adjusted for range restriction ($N = 711$). The original model captured patterns among individuals who had experienced interpersonal felt distrust, but this conditional sample likely constrained variance in IFD, attenuating its correlations with other variables. To address this, a deattenuated covariance matrix was generated using the full sample's unrestricted standard deviation for IFD, thereby modeling relationships as they might unfold in a broader workplace population. The corrected SEM revealed stronger predictive effects of IFD on identity threat and behavioral outcomes and uncovered a previously masked moderation by moral identity centrality, however, the corrected model has a poor model fit, hence, the findings below should be interpreted with caution.

The results of hypothesis testing provide valuable insights into the dynamics linking interpersonal felt distrust, work-related identity threat, emotional reactions, and behavioral responses. As hypothesized, interpersonal felt distrust significantly predicted work-related identity threat (H1), and this relationship became stronger in the corrected SEM, reinforcing the foundational claim that being distrusted by a supervisor powerfully threatens individuals' sense of workplace identity. Notably, the interaction between interpersonal felt distrust and moral identity

centrality (MIC)—previously non-significant—became robustly significant in the corrected model, supporting H2. This finding suggests that individuals with strong moral self-concepts are particularly vulnerable to perceived interpersonal distrust, aligning with theoretical expectations.

The affective reactions to identity threat shifted with the corrected model. While shame-experience remained a significant outcome of work-related identity threat (supporting H4), the path from work-related identity threat to anger-experience (H3) was no longer significant. These patterns suggest that shame may be the more robust and reliable emotional consequence of identity threat, at least under conditions of corrected IFD variance. However, none of the hypothesized direct paths from emotional experiences to behavioral outcomes (H5–H8) were supported, either before or after correction, suggesting that affective states may not independently account for behavioral responses, or that additional regulatory mechanisms are involved.

Importantly, a number of previously undetected direct behavioral effects emerged in the corrected model. Interpersonal Felt Distrust significantly predicted both counterproductive work behavior and exemplification, bypassing emotional mediators altogether. Similarly, work-related identity threat showed new direct predictive power for exemplification and psychological withdrawal. These findings expand the theoretical landscape by suggesting that perceived distrust and identity threat can lead to both disengagement and impression management—representing potentially divergent adaptive strategies.

Additional dispositional predictors maintained their influence. Narcissism remained a significant positive predictor of both CWB and exemplification, suggesting that individuals high in narcissism may engage simultaneously in self-serving and image-enhancing behaviors. Anger-trait and shame-trait also predicted CWB, underscoring the contribution of personality in shaping behavioral responses to perceived threat. Notably, racial minority participants were more likely to

report elevated identity threat, yet less likely to engage in psychological withdrawal — an asymmetry that warrants further exploration. Age was negatively associated with both Psychological Withdrawal Behavior and Shame-Experience, indicating that older individuals were less likely to disengage or report self-conscious emotions in response to identity threat.

Taken together, these results suggest that the corrected SEM provides a more accurate and generalizable depiction of the theoretical processes under study; even though its results should be interpreted with caution due to poor model fit. By accounting for measurement limitations imposed by conditional sampling, the analysis revealed clearer affective and behavioral consequences of interpersonal felt distrust by supervisor at work. The results of hypotheses testing are presented in Appendix C Table A3.

Chapter 6: General Discussion

6.1 Overview

This dissertation is the first to empirically examine the impact of felt distrust from a supervisor on subordinates' emotions and behaviors in the workplace. Importantly, interpersonal felt distrust is conceptualized and measured as a stand-alone construct distinct from felt trust, which is in line with growing research evidence that trust and distrust are unique and separate constructs (i.e., Bijlsma-Frankema et al., 2015; Casado-Aranda et al., 2019; Cho, 2006; Dimoka, 2010; Lee & Lee, 2024; Reimann et al., 2017; Saunders et al., 2014; Saunders & Thornhill, 2004; Sitkin & Roth, 1993). Building on the definition of distrust as “confident negative expectations of another’s conduct” (Lewicki et al., 1998, p. 439), this dissertation conceptualizes felt distrust as a social cognition that another individual thinks one has ill intentions toward them and fears one intends to harm them. This dissertation introduces and empirically examines a novel theoretical mechanism to explain individual reactions to felt distrust, namely the threat that a supervisor’s distrust poses to the subordinate’s work-related identity.

The results of this dissertation indicate that interpersonal felt distrust by a supervisor is experienced by the subordinate as a significant negative evaluation of their self. Felt distrust threatens one’s work-related identity, evokes strong negative moral emotions of anger and shame, and is associated with both negative workplace behaviors (CWB and psychological withdrawal) and positive workplace behaviors (exemplification). Employees who feel distrusted by a supervisor engage in these behaviors because they experience felt distrust as an attack on their personal identity at work. Moreover, the threat to identity is even higher for individuals who place a high value on their moral identity. These findings make several significant contributions to the

complex literature on distrust in the workplace and the emerging literature on interpersonal felt distrust within the organizational sciences. Below I discuss my findings and contributions further.

6.2 Interpersonal felt distrust as a cause of work-related identity threat

This dissertation extends the distrust literature by adopting and empirically examining a novel theoretical perspective on the outcomes of interpersonal felt distrust from a supervisor. It offers a new theoretical mechanism that captures the detrimental effects felt distrust can have on an individual's work-related identity (Miscenko & Day, 2016). Through the lens of identity theory (Stryker, 1980, 1987; Stryker & Burke, 2000), I found that interpersonal felt distrust by a supervisor evokes a strong sense of work-related identity threat. In other words, the experience of being distrusted by one's supervisor constitutes an attack on the subordinate's personal identity at work. Importantly, the results are robust across two studies incorporating different research designs and samples, thereby enhancing external validity. Internal validity is also strengthened due to the experimental design of the first study.

These findings extend the current literature on felt distrust. Past research on reactions to felt distrust predominantly focused on impersonal distrust and proposed retaliation (Kruglanski, 1970), reactance (Brehm, 1966) and negative reciprocity (Frey, 1993) as underlying mechanisms. Fundamentally, it was speculated (but not empirically tested) that employees who feel distrusted will experience lack of autonomy and react with negative workplace behaviors. This dissertation, which focuses on the outcomes of interpersonal felt distrust in the supervisor-subordinate dyad, introduces identity threat as a novel underlying mechanism. In doing so, it complements the very sparse research on outcomes of interpersonal felt distrust. In this limited research, emotional exhaustion was examined as the underlying mechanism in responses to felt distrust from colleagues (Lanaj et al., 2018), and self-regulation was examined as the underlying mechanism in

response to ability-distrust from a supervisor (Liu et al., 2024). Importantly, these studies operationalized felt distrust as low trustworthiness. Aligning the conceptualization and operationalization of interpersonal felt distrust with contemporary perspectives on distrust as a stand-alone construct, this dissertation finds that identity threat perspective captures the destructive and, potentially, long-lasting effect that interpersonal felt distrust from a supervisor has on one's sense of self at work. Since identity threat impacts the value, meaning and/or enactment of one's identity (George et al., 2023; Petriglieri, 2011), this dissertation suggests that the effect of felt distrust constitutes a more profound harm on the individual than previously considered.

By identifying interpersonal felt distrust by a supervisor as an antecedent to identity threat at work, this dissertation also contribute to the work-related identity literature (Dutton et al., 2010; Miscenko & Day, 2016; Selenko et al., 2018) and identity threat literature (Aquino & Douglas, 2003; Kroeper et al., 2022; Petriglieri, 2011). Past research on triggers of identity threat in the context of social interactions identified such critical factors as workplace violence and aggression (Baron & Neuman, 1996; Neuman & Baron, 1998), incivility (Andersson & Pearson, 1999), insults on religious grounds, harsh criticism of one's abilities, deceit, and public humiliation (Bies, 1999). Importantly, this body of research conceptualizes the triggers of identity threat in the context of social interaction as coercive action (Tedeschi & Felson, 1994), meaning "action taken with the intent of imposing harm on another person" (Andersson & Pearson, 1999, p. 457). The findings of this dissertation research contribute to this body of knowledge by identifying interpersonal felt distrust as a cause of identity threat in the context of supervisor-subordinate social interaction. What distinguishes this cause from other established causes of identity threat at work identified in prior research is that interpersonal felt distrust by one's supervisor does not constitute a coercive action. Perceived distrust by one's supervisor instead connotes a defensive action on the part of

the supervisor (Lewicki et al., 1998) that still serves as a trigger of identity threat in the individual on the receiving end of distrust.

6.3 The impact of interpersonal felt distrust on emotions

This dissertation provides critically needed theory and evidence on the individual emotional reactions to felt distrust in the workplace, responding to multiple calls to examine felt distrust and associated emotions (Bies et al., 2018; Mooijman & Graham, 2018; Schoorman et al., 2007). Specifically, it provides empirical evidence that interpersonal felt distrust by a supervisor is associated with negative moral emotions of anger and shame.

To date, only one study examined the impact of felt distrust on individuals' emotions (Liu et al., 2024). This dissertation complements Liu et al.'s (2024) findings and extends the distrust literature by identifying an important indirect effect of felt distrust on shame through identity threat. Whereas Liu et al. (2024) found a direct link between felt distrust and shame, this dissertation reveals a more nuanced picture. The primary pathway through which felt distrust leads to shame is indirect, operating through the experience of identity threat at work. This pattern holds consistently across both experimental and survey studies, suggesting that threats to one's work-related identity play a central role in shaping emotional responses to feeling distrusted by a supervisor. These findings then confirm existing research that interpersonal felt distrust by one's supervisor may evoke shame (Liu et al., 2024), but also extend these findings by providing empirical evidence for a strong mediating effect of work-related identity threat in the relationship between interpersonal felt distrust and shame.

This dissertation is also the first to empirically examine anger as an emotional reaction to felt distrust in the workplace. While it has been suggested that felt distrust can increase negative

emotions, such as anger (Mooijman & Graham, 2018), this was not empirically tested in prior research. This dissertation finds in Study 1 that interpersonal felt distrust is associated with anger through the mediating effect of work-related identity threat. In contrast, Study 2 finds that interpersonal felt distrust is associated with anger directly, and not through the mediating effect of work-related identity threat. The different results across the two studies, i.e., finding support only for indirect effect in Study 1 and only for direct effect in Study 2, may be due to the different research design employed in each study – experiment in Study 1 and CIT with survey, results corrected for range restriction, in Study 2. Since both studies show a consistent effect of interpersonal felt distrust on anger, whether directly or indirectly, this dissertation extends the current literature on distrust and emotions by providing empirical support for association between interpersonal felt distrust by one’s supervisor and anger.

Finally, this dissertation’s findings on the impact of interpersonal felt distrust by one’s supervisor on employee emotions complements parallel research on the emotional outcomes of felt trust. While research on felt trust found that felt trust is associated with pride (Baer et al., 2015), this dissertation found that felt distrust is associated with anger and shame.

6.4 Interpersonal felt distrust by a supervisor

This research contributes to our understanding of the relational nature of distrust in the supervisor-subordinate context. It complements the budding literature on interpersonal felt distrust (Lanaj et al., 2018; Liu et al., 2024; Rudorf et al., 2018) by identifying the emotional and behavioral outcomes of feeling distrusted by a supervisor at work. The interpersonal context of felt distrust is unique because the employee who feels distrusted by a supervisor is singled out from the rest of the colleagues in this experience. Unlike the impersonal nature of felt distrust, where a group of employees feels distrusted by management or organization in general (Cha & Wong-On-

Wing, 2020; Christ, 2013; Mooijman et al., 2017; Schweitzer et al., 2018), interpersonal felt distrust is experienced in a more intimate way and has the potential to affect the quality of the relationship with the supervisor. In this manner, this dissertation research also contributes to the Leader-Member Exchange (LMX) literature. While high quality relationships are characterized by mutual influence, trust and respect (Anand et al., 2011; Dansereau et al., 1975; Liden & Maslyn, 1998), employees feeling distrusted by a supervisor may never be able to attain such a level of relationship with their supervisor because their positive sense of self in the workplace is attacked by the leader on whom they depend for this positive identity validation and support. This experience shakes the very foundation necessary for the emergence of mutual trust, respect or liking of one another. Research also suggests that the LMX development process is heavily influenced by affect in the leader-follower relationship (Anand et al., 2011). Since this dissertation finds that interpersonal felt distrust is often accompanied by anger and shame, it provides additional evidence that employees who feel distrusted by a supervisor may struggle to attain the desired quality of LMX, suggesting interpersonal felt distrust as a stumbling block on the way of development of high-quality leader-follower relationships.

6.5 Moral identity centrality as a moderator in the relationship between interpersonal felt distrust and work-related identity threat

Moral identity centrality, a dispositional factor that captures the importance of moral identity within the hierarchy of personal identities (Aquino et al., 2009; Aquino & Reed, 2002; Boegershausen et al., 2015), was found to strengthen the relationship between interpersonal felt distrust and work-related identity threat (in Study 1 and Study 2 when correcting for range restriction). Individuals with high moral identity centrality (MIC) experienced a higher level of work-related identity threat following interpersonal felt distrust than individuals who scored low

on MIC. This finding further supports the theoretical claim that interpersonal felt distrust takes a toll on one's personal identity at work, particularly because it targets its moral dimensions, such as being seen as a kind and caring human being. This finding makes an important contribution to our understanding of personality factors that may magnify the negative effects of interpersonal felt distrust by a supervisor at work. The findings also suggest that when employees with strong moral values feel distrusted by their supervisors, their moral identity may actually heighten, not protect against, the experience of identity threat at work.

Interestingly, the moderating effect of MIC contrasts with prior research findings. Previous studies that examined the moderation effect of psychological identity centrality (Huang Johnson, 2015) and moral identity centrality (A. Chen, 2020) did not find a positive (hypothesized) statistically significant effect of identity centrality on the positive relationship between the trigger of identity threat (the event itself) and perceived identity threat (the experience of that event). These authors offered several potential explanations for not seeing the hypothesized interaction effect and suggested future investigation into the moderating effect of MIC given the strong theoretical argumentation for its potential role in strengthening the positive relationship between identity threat-inducing incidents and perceived identity threat. This dissertation responds to those calls and provides empirical evidence for the moderating effect of moral identity centrality on the relationship between interpersonal felt distrust and work-related identity threat across two studies.

6.6 Behavioral reactions to interpersonal felt distrust

This dissertation finds that interpersonal felt distrust is associated with negative workplace behaviors, such as CWB and psychological withdrawal, and positive behavior, such as exemplification. Even though the hypothesized relationships between emotional reactions and behavioral responses to felt distrust didn't receive support, the results of a holistic assessment of

Study 2 SEM regression results revealed other statistically significant direct and indirect effects between interpersonal felt distrust and the behavioral outcomes in the model.

First, this dissertation finds support for both negative (direct) and positive (indirect, through work-related identity and shame) effects of interpersonal felt distrust on CWB. In other words, when employees felt distrusted, they both directly decreased their engagement in CWB and indirectly increased their engagement in CWB via shame. These findings are consistent with the idea that individuals may regulate their behavioral response to feeling distrusted by a supervisor, but when experiencing a moral negative emotion of shame following felt distrust, they may be less likely to regulate their behavior and engage in CWB. The finding of positive association between shame and CWB is consistent with prior research (Bauer & Spector, 2015). In this way, this study complements our understanding of the important role of shame in shaping employees' experiences at work and its potential to increase the instances of CWB, even when the direct effect is negative. This finding also extends the current literature on emotions in the workplace by identifying interpersonal felt distrust as a possible antecedent of shame in employees. It also contributes to the literature on felt distrust by identifying CWB directed at the organization as a possible behavioral outcome of interpersonal felt distrust by one's supervisor, through the mediating mechanisms of work-related identity threat and shame.

Second, the findings that interpersonal felt distrust affects psychological withdrawal through work-related identity threat is consistent with research on the association between interpersonal felt distrust by coworkers and withdrawal at work (Lanaj et al., 2018). This dissertation identifies an alternative theoretical mechanism that explains this relationship -- threat to the distrusted employee's work-related identity, rather than emotional exhaustion. As such, it extends the literature on the outcomes of interpersonal felt distrust by offering an alternative

explanation, as well as examining an alternative referent (i.e., felt distrust by one's supervisor rather than by colleagues) to explain psychological withdrawal when an employee feels distrusted by their supervisor.

Third, with the exception deviant innovation (Liu et al., 2024), past research has not examined positive behavioral outcomes to felt distrust. The findings in this dissertation about the positive association between interpersonal felt distrust and exemplification, through the mediating role of work-related identity threat and directly, is a novel contribution to the distrust literature.

Contrary to the hypothesized relationships, this dissertation did not find direct effects between (1) anger and CWB; (2) anger and psychological withdrawal; (3) shame and psychological withdrawal; and (4) shame and exemplification. This comes in contrast with prior research that suggests an association between anger and counterproductive behaviors (Fox et al., 2001), shame and psychological withdrawal (de Hooze et al., 2010), and shame and exemplification (Bonner et al., 2017; de Hooze et al., 2008, 2018; Molines & Perrier, 2024). There may be several reasons as to why the hypothesized relationships were not supported in Study 2, even after the model was corrected for range restriction. First, the lack of support may be due to the floor effects of the behavioral measures (i.e., most of the participants responded on the lower end of the scales) (Beal & Dawson, 2007). However, a review of the descriptive statistics (mean, standard deviation) of CWB, psychological withdrawal, and exemplification in past studies (Barclay & Kiefer, 2014; Bauer & Spector, 2015; Bonner et al., 2017; Khan et al., 2013, 2017) reveals similar ranges to those of the descriptives of these variables in Study 2, which suggests that the truncated distributions (Beal & Dawson, 2007) of these variables is a normal occurrence in organizational science research. Second, it is possible that emotions didn't translate into behaviors due to individuals' ability to self-regulate (Heatherton & Baumeister, 1996) to achieve

their career goals. Engaging in counterproductive or withdrawal behaviors could cost someone their career or result in further complications in one's interpersonal relationship with their supervisor. Finally, engaging in exemplification behavior may be challenging in the context of a threatened work-related identity because it may seem insincere or drawing unwanted attention.

6.7 Practical Implications

This dissertation offers significant insights into the psychological and organizational consequences of employees feeling distrusted by their supervisors. While trust has long been understood as a cornerstone of effective leadership, this research reveals that *felt distrust* is not merely the absence of trust — it is an interpersonal experience that can seriously undermine employees' self-concept at work. The results across both experimental and survey studies indicate that interpersonal felt distrust by a supervisor leads to work-related identity threat and evokes strong negative emotional responses, particularly anger and shame. For organizational leaders and HR professionals, this highlights the need to view trust dynamics not only in terms of fostering positive climates but also in terms of addressing and alleviating communication of distrust by supervisors.

One important practical implication of this research is the explanation it offers for why rebuilding trust may fail in supervisor–subordinate relationships where felt distrust has taken root. Drawing on the work of Bijlsma-Frankema et al. (2015), this dissertation supports the idea that distrust is not a neutral or passive condition—rather, it actively disrupts the foundation necessary for trust to exist. When employees perceive that they are distrusted, their sense of value, meaning and enactment in their role becomes threatened, creating a psychological barrier to reestablishing mutual respect and trust. This finding is particularly relevant for HR professionals involved in conflict resolution and leadership development as it highlights the need to look more closely at

what's causing ongoing trust issues. Another key implication is the importance of recognizing emotions —especially anger and shame —as signs of deeper psychological strain rooted in work-related identity threat. In this case, these emotional responses are not necessarily irrational or disruptive by nature, rather, they may be adaptive reactions to one's damaged sense of self. Supervisors and HR practitioners should therefore pay attention to the underlying causes of workplace emotions, rather than dismissing them as overreactions. If left unacknowledged, these emotional experiences can weigh heavily on the individual, continue to grow, shaping how employees feel about their role and the organization.

Also, the dissertation findings point to downstream effects, such as increased likelihood of counterproductive work behaviors (CWBs), psychological withdrawal, and efforts to repair one's self-image through impression management behaviors like exemplification. Engaging in deviant behaviors at work may negatively affect the morale of other workers (Robinson et al., 2014). Psychological withdrawal may lead to decreased productivity and harm organizational performance in the long run (Sagie et al., 2002). Exemplification behaviors, while seemingly harmless strategic impression management behaviors (D. M. Long, 2017), can put a great toll on the employee's emotional and physical well-being, prove to be exhausting, and decrease overall job satisfaction (Baer et al., 2015; De Cuyper et al., 2014; Leary et al., 1994; LePine et al., 2005). These behavioral outcomes may appear paradoxical — for instance, an employee working harder to prove their worth while simultaneously feeling alienated. For organizational decision-makers, this suggests that performance evaluations alone may overlook deeper issues affecting employees. A closer understanding of the relational context, particularly felt distrust by supervisor, will help managers and HR practitioners respond with actions that effectively support employees and repair damaged working relationships.

Finally, this research emphasizes the need for leaders to be trained to recognize how subtle signals in their attitudes, emotional expressions or actions — such as weariness, cynicism, lack of transparency, micromanagement — may inadvertently convey distrust. Managers should be reminded that relational misperceptions are commonplace in the workplace (Byron & Landis, 2020). Since felt distrust is a subjective individual psychological experience, it may be felt by a subordinate even when a supervisor has no intention of distrusting them, if their attitudes, words, mood or actions signal otherwise. Therefore, it is important for supervisors to build proactive and transparent communication and relational rapport with employees to minimize opportunities for miscommunication and misperceptions. In addition, organizations should establish formal and informal channels where employees feel safe discussing their experiences of feeling distrusted. These channels can help HR professionals detect early signs of distress before they translate into employee disengagement, burnout or harm to the organization. By recognizing that the communication of distrust by managers can cause such emotions and/or behaviors, leaders are better equipped to consider the unintended signals of distrust in their communication. This heightened awareness is required by organizational leaders to practice greater emotional intelligence, maintain a culture of psychological safety, and cultivate relationships with subordinates that are grounded in transparency, mutual respect, and shared purpose—laying the groundwork for a healthier and more effective organizational culture.

Chapter 7: Limitations and Future Research

7.1 Limitations

As is the case with any research, this dissertation research contains several limitations. I will first discuss two possible theoretical limitations, followed by several methodological limitations of this dissertation research.

The theoretical limitations center around the role of attributions in the experience of felt distrust. First, the present research does not account for variation in how employees perceive their supervisor. Specifically, the study did not assess whether the supervisor was perceived to have a generalized tendency to distrust others. Prior research suggests that disposition to distrust (McKnight & Chervany, 2001a, 2001b) affects one's distrust attitude. It is possible that when a supervisor is seen as generally distrustful of everyone—rather than singling out a particular employee—the experience of felt distrust may be less personalized, and thus less likely to threaten one's work-related identity or evoke shame or anger. In such cases, employees may attribute felt distrust to the supervisor's character rather than to a deficiency in their own self, thereby buffering the emotional impact. This contextual nuance was not captured in the current study design and presents a valuable direction for future research.

Another limitation worth noting is that this research did not examine whether an individual considers their experience of interpersonal felt distrust to be warranted or not. It well may be that in some situations or contexts, when an individual feels distrusted, they take blame for it and justify their supervisor's distrust. However, there may be other situations in which an individual who feels distrusted thinks it is unwarranted that they have been distrusted and does not see themselves as the source of the problem. It is possible that when the attribution of blame is made toward oneself,

the individual is more likely to feel shame than anger; however, when attribution of blame is made towards the supervisor, the individual is more likely to feel anger than shame. More research is needed to understand these dynamics.

In Study 1, one potential methodological limitation concerns internal validity, specifically the decision not to exclude participants who “failed” the manipulation check for their assigned condition. Although the manipulation check revealed a significant difference in felt distrust between the experimental and control groups, individual-level variability remained within each condition. That is, some participants in the control group reported feeling distrusted, while some in the experimental group did not. Participants were retained regardless of manipulation check scores to preserve the natural variability in responses and reflect the reality that felt distrust is a subjective social cognition that may not be uniformly triggered by identical stimuli. While this approach may introduce noise, it also enhances ecological validity by acknowledging that individuals interpret contextual cues differently. Nonetheless, this variability may have attenuated the observed effects and should be considered when interpreting the findings (Shadish et al., 2002).

Another methodological limitation concerns potential confounds in the experimental manipulation of felt distrust. While the scenarios were intentionally designed to reflect a stark contrast between high and low distrust, drawing on established literature regarding the nature of distrust (e.g., Lewicki et al., 1998) and supervisor behaviors that signal distrust (e.g., Bijlsma-Frankema et al., 2015), the inclusion of morally charged language in the experimental condition may have inadvertently aligned with specific items added to the modified work-related identity threat measure. This overlap raises the possibility that observed differences in work-related identity threat could be partially driven by shared linguistic cues rather than the manipulation alone. Additionally, the scenarios differed in length and in the degree of supervisor involvement,

which may have introduced unintended inequivalence across conditions. Although these design choices were made to enhance psychological realism and clearly differentiate high distrust from the neutral condition, they may have contributed to alternative explanations for the observed effects. Additionally, the use of generic organizational description and limited contextual detail in the scenarios may have constrained psychological realism of the vignette-based manipulation of felt distrust. A statistical limitation of Study 2 may be the choice of a targeted purposive sampling strategy therein whereby only employees who felt distrusted by a supervisor were surveyed, which could have introduced range restriction (Carretta & Ree, 2022; Sackett & Yang, 2000). However, such a choice of sampling strategy does not pose a threat to the external validity of the study (Landers & Behrend, 2015) since the external validity of convenience samples in I – O psychology research is typically judged by sample relevance and prototypicality rather than whether they are representative of the intended target population (Sackett & Larson, 1990). Since the sample of employees who felt distrusted by a supervisor reflects a recognizable and relevant subset of the global workforce and allows for focused examination of an underrepresented employee experience, its use is both appropriate and informative. Therefore, the choice of a targeted convenience sample in Study 2 is justified and does not compromise the external validity of the research. A related limitation in Study 2 concerns a poor overall fit of the model corrected for range restriction. Although the corrected model accounts for attenuation and offers theoretically meaningful paths, its poor fit suggests that the specified structure does not fully capture the underlying relationships in the data. Therefore, findings should be interpreted with caution, and future research is needed to refine the model and validate its pathways using improved fit indices.

Another limitation of this dissertation is reliance on participant memory for recall of their behavioral reactions to felt distrust by supervisor at work. Given that many of the recalled

experiences in Study 2 took place over a year before participating in the study, participants may not be able to accurately recall whether they engaged in certain behaviors or not at the time of the incident. Therefore, scholars recommend a much more immediate assessment of counterproductive workplace behaviors for more accurate memory recall, such as at the end of the workday (Spector, Fox & Spector, 2005). Moreover, research suggests that memory recall tasks are prone to false memories (Toglia, Neuschatz, & Goodwin, 1999) and accuracy of self-reported behavioral assessment following an incident is affected by the ease of incident retrieval from long-term memory (Aarts & Dijksterhuis, 1999). This suggests that reliance on individuals' retrospective assessment of their behavioral responses to felt distrust may be vulnerable to distortion. Future research would benefit from incorporating experience sampling methodology, i.e. research technique that captures people's thoughts, emotions, and behaviors in real time and in their natural environments (Myin-Germeys & Kuppens, 2021), to triangulate self-report data and reduce the influence of memory-based biases.

A methodological limitation present in both studies is related to being able to draw firm causal conclusions about the proposed relationships. Because both studies' designs involved serial mediation within complex cognition-emotion and cognition-emotion-behavior relationships, there's a possibility of reverse causality or alternative explanations for the effects among the variables in every relationship. Without temporal separation or experimental manipulation of each step in the conceptual model, it is difficult to determine the directionality of effects among the variables in each pathway or rule out the effect of third variables on the outcomes of interest.

One more possible methodological limitation is that the findings may not be generalizable to certain organizational contexts or cultures. This research implicitly assumes that employees expect a collaborative, trusting, and positive supervisor-subordinate relationship. In such a type of

relationship, it is more likely that employees who feel distrusted by their supervisors will experience a threat to their work-related identity. However, in certain organizations (i.e., secret service organizations, mafia cartels, and even some layers of government officials) distrust is intentionally built and maintained, and the nature of interpersonal relationships between a supervisor and a subordinate is naturally grounded in suspicion and cautiousness owing to the nature of the trade (Božič & Keston-Siebert, 2024; Champeyrache, 2022; Gambetta, 2000; Rohde, 2020; Siebert & Czarniawska, 2020; Wintrobe, 2018). It is likely that because interpersonal felt distrust may be normalized and expected between supervisors and subordinates in these distrustful organizations, it may not elicit as much of a threat to one's work-related identity. Therefore, attention should be paid to the nuances of the organizational context in which distrust toward employees and felt distrust by a supervisor may be "normalized" as a feature of a trade, since this may affect employees' responses thereto in ways that cannot be predicted by this research's findings.

Another methodological limitation of this research, common method variance (Podsakoff et al., 2003, 2024), relates to cross-sectional and self-reported data in both studies. In both studies, concerns about common method variance were minimized by using multiple item measures of the core constructs with high reliability of the scales and statistically dissipated by conducting confirmatory factor analysis (CFA) on the overall model. Importantly, the specifics of the research design of each study (i.e., an experiment in Study 1 and a CIT integrated in survey in Study 2) did not lend themselves to a longitudinal approach because such techniques require a lack of temporal break between the manipulation technique and the survey in order to most accurately capture individual emotions and/or behaviors related to the induced (in Study 1) or recalled (in Study 2) experience of interpersonal felt distrust.

Moreover, the self-reported nature of both surveys allows to capture the sensitive and subjective nature of interpersonal felt distrust. Since feeling distrusted is a subjective experience, it is best understood by the people who experience it, not their observers or supervisors. Likewise, reporting the behaviors that took place after feeling distrusted is also best captured by the individuals who experienced and felt distrust. It would not be feasible or reliable to ask others (i.e., supervisors or colleagues) about emotions expressed by another person at work who feels distrusted. Moreover, recent organizational psychology research provides an evidence-based understanding of the merits of using self-reported data rating the source of one's personality, emotions, and behaviors (Mercado et al., 2018), grounded in the results of meta-analyses demonstrating convergence across rating sources for CWB, OCB, and other related constructs (e.g., Berry et al., 2012; Carpenter et al., 2014).

A more general methodological limitation may be due to the reliance on a paid online participant panel for conducting both studies. There has been much criticism of employing paid online panels as study participants, especially with respect to data quality and sample representability (Boas et al., 2020; Landers & Behrend, 2015; Porter et al., 2019). At the same time, online panels prove to be a convenient, accessible, and popular source of data for organization science research (Aguinis et al., 2020; Palan & Schitter, 2018; Walter et al., 2018). Online panel participants were a suitable sample for both studies in this dissertation because they allowed me to pre-select participants who have been employed and who had experience being supervised at work. Special efforts have been made to follow best practices in utilizing online surveys and online panels for research (Aguinis & Bradley, 2014; Tristram et al., 2012c, 2012a, 2012b) to increase the quality of participant responses. Specifically, each study's online survey contained several attention checks, an instructional check, and a seriousness check that had to be

passed to qualify for the study. All questions were made mandatory to respond, which took care of potential issues with missing data. The online set-up of the survey verified that the survey was completed from one IP address only, limiting participant responses to one each. Qcaptcha score was generated by Qualtrics identify any instance in which the survey has been completed by a bot rather than a human. As a result, all of the participant responses that were included in the final datasets passed all validation checks and Qualtrics confirmation that the survey was completed by humans. With all of the required mechanisms in place to secure the quality of the data, an online panel provides a unique opportunity of access to diverse participant pools and to address sensitive research questions that people may find difficulty discussing openly in a field study in an organization or even in a paper survey.

7.2 Future Research

This dissertation on the impact of felt distrust in the workplace provides a fruitful arena for future research.

Future research should aim to strengthen causal claims about the impact of felt distrust by incorporating longitudinal designs that allow for temporal separation, such as experience sampling methodology, or experimental designs that avoid serial mediation, such as those that examine a single specific pathway in the conceptual model at a time. Surveying the same participants over time could help establish the directionality of effects within cognition–emotion and cognition–emotion–behavior pathways. Experiments that focus on a specific relationship in the model at a time, i.e., the effect of felt distrust on emotions, the effect of emotions on felt distrust, etc., could offer a more rigorous manipulation of the main variable direct effect and evaluate the potential bidirectional or recursive relationships among variables, thereby clarifying whether certain emotional states influence cognitive appraisals or vice versa. Including variables that control for

alternative explanations in the outcomes of interest would further enhance the robustness of findings and reduce the likelihood of confounding due to third-variable influences.

Future research could extend this dissertation work by refining experimental manipulations to balance psychological realism with structural equivalence across conditions. While Study 1 vignettes deliberately employed a sharp contrast in distrust signaling by the supervisor to evoke strong responses, future studies might benefit from a more nuanced scenario design that ensures comparable length, tone, and supervisor involvement across conditions. Incorporating richer contextual detail and more vivid organizational background in the scenarios could also enhance participant engagement and emotional immersion, addressing concerns about realism in vignette-based designs. Drawing on recommendations such as those outlined by Fischer et al. (2021), researchers may consider pretesting scenarios for perceived authenticity and emotional impact or employing immersive methods such as role-play or video-based manipulations to better capture how felt distrust unfolds in real-world supervisory interactions.

Conceptually, future research could examine how individual traits of both employees and supervisors influence subordinates' experiences of interpersonal felt distrust and its related outcomes. For example, employees with low core self-evaluations (Judge, Erez, Bono, & Thoresen, 2003)—marked by low self-esteem, low emotional stability, feelings of inefficacy and external locus of control—may be more likely to interpret ambiguous supervisory behaviors as signals of distrust. Conversely, supervisor traits, such as low agreeableness, high neuroticism, or low emotional intelligence (Goleman, 1998; Bar-On, 2007), may make some supervisors more prone to communicating distrust, even unintentionally. These dynamics warrant further empirical examination.

Another promising direction involves examining the supervisor's generalized reputation for distrustfulness. According to McKnight and Chervany (2001a, 2001b), people develop expectations of others' trustworthiness based on past experiences. If an employee perceives their supervisor as chronically distrustful of everyone, the resulting distrust may feel less personal, potentially dampening its impact on identity threat or shifting emotional responses from shame to frustration. Additionally, the degree of respect and legitimacy the employee attributes to the supervisor likely affects the value they attach to felt distrust by supervisor. When a highly respected leader communicates distrust, it may carry greater weight, intensifying work-related identity threat and leading to deeper emotional distress. Past research suggests that employee perceptions of leader trustworthiness impact employee's internalization of leader expectations into self-expectations and affect performance (Karakowsky, DeGama & McBey, 2012). In contrast, when felt distrust comes from a supervisor viewed as incompetent or immoral, the employee may dismiss the negative evaluation altogether.

The setting in which distrust is communicated—whether in private or in public—may also influence outcomes. Public expression of interpersonal distrust by supervisor could increase employee's feeling of shame due to reputational damage and social disapproval from others, while private comments may increase anger. Understanding under which circumstances felt distrust may result in shame or anger would also allow us to better predict the behavioral outcomes, such as whether interpersonal felt distrust may result in counterproductive work behavior (following anger) as opposed to exemplification behavior (following shame).

Future research could also examine how employees' expectations about the future consequences of distrust by supervisors, such as potential damage to their promotion or career in the organization, may affect the severity of both the identity threat and the emotional reaction.

These future-oriented concerns may amplify the experience of identity threat and intensify emotional reactions. Anticipating lasting damage to one's professional standing and career prospects can make the experience of felt distrust especially impactful.

An additional interesting future research stream could examine antecedents of felt distrust. While we know about some individual behaviors that signal distrust in general, such as observable defensiveness, watchfulness, vigilance, diminished cooperation and avoidance of interaction (Bijlsma-Frankema et al., 2015; Lewicki et al., 1998), it is less clear which managerial behaviors may trigger interpersonal felt distrust. Moreover, given that felt distrust involves a subjective perception, it is quite possible that some managerial behaviors may have unintended consequences on their subordinates' felt distrust. Since individual perceptions are not always reflective of objective reality (Byron & Landis, 2020), individuals may interpret managerial behaviors and decisions as signs of distrust while managers may not even realize it (Litzky et al., 2006).

Another area that requires further attention is whether an individual considers their experience of interpersonal felt distrust to be warranted or not. That is, more research is needed to understand whether employees' attribution of responsibility or blame for interpersonal felt distrust, affects work-related identity threat and associated emotional and behavioral outcomes. Current research alerts us that little attention has been paid to understanding the moral risks of "distrusting the trustworthy" (D'Cruz, 2019, 2020) or to the emotional and behavioral consequences of unwarranted felt distrust.

Future research can also investigate practical solutions to the outcomes of interpersonal felt distrust by a supervisor in the workplace. While the organizational literature provides some insights on ways to address counterproductive behaviors (Flores, 2023; London & Smither, 2009) psychological withdrawal (Alexander, 2016; Koslowsky & Krausz, 2002) and alerts to possible

stresses associated with excessive exemplification behaviors at work (Baer et al., 2015; Leary, 1995; LePine et al., 2005) in general, there is a need to develop empirically-supported best managerial practices to address the issue and consequences of interpersonal felt distrust by a supervisor at work. Future action-oriented research can engage in investigating practical solutions to managing interpersonal felt distrust by a supervisor at work to provide a toolkit for HR managers and organizational leaders to address this issue.

It would be valuable to further explore the impact of interpersonal felt distrust on representatives of visible minority groups in light of organizational foci on equity, diversity and inclusion in the workplace. In this dissertation research, racial minority participants were more likely to report elevated identity threat in Study 2, yet less likely to engage in psychological withdrawal — an asymmetry that warrants further exploration.

Finally, future research may also investigate the outcomes of felt distrust among leaders (i.e., between a low-level and a middle-level manager) or felt distrust among top management teams. While this dissertation contributes to the nascent literature on felt distrust in the supervisor-subordinate context, no research has been conducted on interpersonal felt distrust among organizational leaders. At the same time, leader relations appear to be precisely prone to issues of distrust (Fukuyama, 1995; Hardin, 2004), and hence issues of felt distrust, due to a high level of power and ambition held by both interaction partners, and the high risks and stakes involved in the relationship.

References

- Aarts, H., & Dijksterhuis, A. (1999). How often did I do it? Experienced ease of retrieval and frequency estimates of past behavior. *Acta Psychologica*, 103(1-2), 77-89.
- Adeyeye, J. O., Adeniji, A. A., Osinbanjo, A. O., & Oludayo, O. A. (2015). Effects of workplace ethics on employees and organisational productivity in Nigeria. *International Conference on African Development Issues*, 267–273.
- Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, 17 (4), 351–371. <https://doi.org/10.1177/1094428114547952>
- Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, 39(6), 1490–1528. <https://doi.org/10.1177/0149206313478188>
- Aguinis, H., Hill, N. S., & Bailey, J. R. (2021). Best practices in data collection and preparation: Recommendations for reviewers, editors, and authors. *Organizational Research Methods*, 24(4), 678–693. <https://doi.org/10.1177/1094428119836485>
- Aguinis, H., Villamor, I., & Ramani, R. S. (2020). MTurk research: Review and recommendations. *Journal of Management*, 47(4), 823–837. <https://doi.org/10.1177/0149206320969787>
- Alexander, J. F. (2016). Mitigating the effects of withdrawal behavior on organizations [Doctoral dissertation, Walden University].
- Allred, K. G. (1999). Anger and retaliation: Toward an understanding of impassioned conflict in organizations. In R. J. Bies, R. J. Lewicki, & B. H. Sheppard (Eds.), *Research in negotiation in organizations* (Vol. 7, pp. 27–58). Elsevier Science/JAI Press.
- Allred, K. G. (2000). Anger and retaliation in conflict: The role of attribution. In M. Deutsch & P. T. Coleman (Eds.), *The handbook of conflict resolution: Theory and practice* (pp. 236–255). Jossey-Bass.
- Allred, K. G., Mallozzi, J. S., Matsui, F., & Raia, C. P. (1997). The influence of anger and compassion on negotiation performance. *Organizational Behavior and Human Decision Processes*, 70(3), 175–187. <https://doi.org/10.1006/OBHD.1997.2705>
- Alpers, L. M. (2018). Distrust and patients in intercultural healthcare: A qualitative interview study. *Nursing Ethics*, 25(3), 313–323. <https://doi.org/10.1177/0969733016652449>
- Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality*, 40(4), 440–450. <https://doi.org/10.1016/j.jrp.2005.03.002>

- Anand, S., Hu, J., Liden, R. C., & Vidyarthi, P. R. (2011). Leader-member exchange: Recent research findings and prospects for the future. In *The Sage handbook of leadership* (pp. 311–325).
- Andersson, L. M., & Pearson, C. M. (1999). Tit for tat? The spiraling effect of incivility in the workplace. *Academy of Management Review*, 24(3), 452–471.
<https://doi.org/10.5465/amr.1999.2202131>
- Andreescu, V., & Keeling, D. G. (2012). Explaining the public distrust of police in the newest European Union countries. *International Journal of Police Science and Management*, 14(3), 219–245. <https://doi.org/10.1350/ijps.2012.14.3.278>
- Aquino, K., & Douglas, S. C. (2003). Identity threat and antisocial behavior in organizations: The moderating effects of individual differences, aggressive modeling, and hierarchical status. *Organizational Behavior and Human Decision Processes*, 90(1), 195–208.
[https://doi.org/10.1016/S0749-5978\(02\)00517-4](https://doi.org/10.1016/S0749-5978(02)00517-4)
- Aquino, K., Freeman, D., Reed, A. I., Lim, V. K. G., & Felps, W. (2009). Testing a social-cognitive model of moral behavior: The interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology*, 97(1), 123–141.
- Aquino, K., & Reed, A. I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037/0022-3514.83.6.1423>
- Aquino, K., Tripp, T. M., & Bies, R. J. (2001). How employees respond to personal offense: The effects of blame attribution, victim status, and offender status on revenge and reconciliation in the workplace. *Journal of Applied Psychology*, 86(1), 52–59.
<https://doi.org/10.1037/0021-9010.86.1.52>
- Aquino, K., Douglas, S. C., & Martinko, M. J. (2004). Overt anger in response to victimization: Attributional style and organizational norms as moderators. *Journal of Occupational Health Psychology*, 9(2), 152–164. <https://doi.org/10.1037/1076-8998.9.2.152>
- Aquino, K., Tripp, T. M., & Bies, R. J. (2006). Getting even or moving on? Power, procedural justice, and types of offense as predictors of revenge, forgiveness, reconciliation, and avoidance in organizations. *Journal of Applied Psychology*, 91(3), 653–668.
<https://doi.org/10.1037/0021-9010.91.3.653>
- Armstrong, K., Rose, A., Peters, N., Long, J. A., McMurphy, S., & Shea, J. A. (2006). Distrust of the health care system and self-reported health in the United States. *Journal of General Internal Medicine*, 21(4), 292–297. <https://doi.org/10.1111/j.1525-1497.2006.00396.x>
- Ashforth, B. E., & Johnson, S. A. (2001). Which hat to wear? The relative salience of multiple identities in organizational contexts. In M. A. Hogg & D. J. Terry (Eds.), *Social Identity Processes in Organizational Contexts* (pp. 31–48). Psychology Press.
- Ashkanasy, N. M., Härtel, C. E. J., & Zerbe, W. J. (2000). Emotions in the workplace: Research, theory, and practice. In N. M. Ashkanasy, C. E. J. Härtel, & W. J. Zerbe (Eds.), *Emotions in*

- the workplace: Research, theory, and practice* (pp. 3–18). Quorum Books/Greenwood Publishing Group.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(3), 128–138. <https://doi.org/10.1027/1614-2241/A000014>
- Aubé, C., Rousseau, V., Mama, C., & Morin, E. M. (2009). Counterproductive behaviors and psychological well-being: The moderating effect of task interdependence. *Journal of Business and Psychology*, 24(3), 351–361. <https://doi.org/10.1007/S10869-009-9113-5/FIGURES/2>
- Averill, J. R. (1982). *Anger and aggression: An essay on emotion*. Springer.
- Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, 45(2), 527–535. <https://doi.org/10.3758/s13428-012-0265-2>
- Baba, M. L. (1999). Dangerous liaisons: Trust, distrust, and information technology in American work organizations. *Human Organization*, 58(3), 331–346. <https://doi.org/10.17730/HUMO.58.3.HT622PK6L41L35M1>
- Baer, M. D., Dhensa-Kahlon, R. K., Colquitt, J. A., Rodell, J. B., Outlaw, R., & Long, D. (2015). Uneasy lies the head that bears the trust: The effects of feeling trusted on emotional exhaustion. *Academy of Management Journal*, 58(6), 1637–1657. <https://doi.org/10.5465/amj.2014.0246>
- Baer, M. D., Frank, E. L., Matta, F. K., Luciano, M. M., & Wellman, N. (2021). Under trusted, over trusted, or just right? The fairness of (in)congruence between trust wanted and trust received. *Academy of Management Journal*, 64(1), 180–206. <https://doi.org/https://doi.org/10.5465/amj.2018.0334>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Barclay, L. J., & Kiefer, T. (2014). Approach or avoid? Exploring overall justice and the differential effects of positive and negative emotions. *Journal of Management*, 40(7), 1857–1898. <https://doi.org/10.1177/0149206312441833>
- Barclay, L. J., Skarlicki, D. P., & Pugh, S. D. (2005). Exploring the role of emotions in injustice perceptions and retaliation. *Journal of Applied Psychology*, 90(4), 629–643. <https://doi.org/10.1037/0021-9010.90.4.629>
- Baron, R. A., & Neuman, J. H. (1996). Workplace violence and workplace aggression: Evidence of their relative frequency and potential causes. *Aggressive Behavior*, 22, 161–173.
- Bar-On, R. (2007). The Bar-On model of emotional intelligence: A valid, robust and applicable EI model. In *Organisations and People* (2nd ed., Vol. 14, pp. 27–34).

- Bauer, J. A., & Spector, P. E. (2015). Discrete negative emotions and counterproductive work behavior. *Human Performance*, 28(4), 307–331. <https://doi.org/10.1080/08959285.2015.1021040>
- Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin*, 91(1), 3–26. <https://doi.org/10.1037/0033-2909.91.1.3>
- Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review*, 103(1), 5–33.
- Beal, D. J., & Dawson, J. F. (2007). On the use of Likert-type scales in multilevel data: Influence on aggregate variables. *Organizational Research Methods*, 10(4), 657–672. <https://doi.org/10.1177/1094428106295492>
- Belot, M., & Schröder, M. (2016). The spillover effects of monitoring: A field experiment. *Management Science*, 62(1), 37–45. <https://doi.org/10.1287/mnsc.2014.2089>
- Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85(3), 349–360. <https://doi.org/10.1037/0021-9010.85.3.349>
- Berinsky, A. J., Margolis, M. F., Sances, M. W., & Warshaw, C. (2021). Using screeners to measure respondent attention on self-administered surveys: Which items and how many? *Political Science Research and Methods*, 9(2), 430–437. <https://doi.org/10.1017/psrm.2019.53>
- Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology*, 97(3), 613–636.
- Benamati, J., Serva, M. A., & Fuller, M. A. (2006). Are trust and distrust distinct constructs? An empirical study of the effects of trust and distrust among online banking users. *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, 121b–121b. <https://doi.org/10.1109/HICSS.2006.63>
- Bies, R. J. (1999). Interactional (in)justice: The sacred and the profane. In J. Greenberg & R. Cropanzano (Eds.), *Advances in Organizational Justice*. The New Lexington Press.
- Bies, R. J., Barclay, L. J., Saldanha, M. F., Kay, A. A., & Tripp, T. M. (2018). Trust and distrust. In *The Routledge Companion to Trust* (1st ed., pp. 302–325). Taylor and Francis Inc. <https://doi.org/10.4324/9781315745572-21>
- Bies, R. J., & Tripp, T. M. (2002). “Hot flashes, open wounds”: Injustice and the tyranny of its emotions. In S. W. Gilliland, D. D. Steiner, & D. P. Skarlicki (Eds.), *Emerging perspectives on managing organizational justice* (pp. 203–224). Information Age.
- Bies, R. J., Tripp, T. M., & Kramer, R. M. (1997). At the breaking point: Cognitive and social dynamics of revenge in organizations. In R. Giacalone & J. Greenberg (Eds.), *Antisocial behavior in organizations* (pp. 18–36). Sage.

- Bijlsma-Frankema, K. M., Sitkin, S. B., & Weibel, A. (2015). Distrust in the balance: The emergence and development of intergroup distrust in a court of law. *Organization Science*, 26(4), 1018–1039. <https://doi.org/10.1287/orsc.2015.0977>
- Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. Kurtines & J. Gewirtz (Eds.), *Morality, Moral Behavior and Moral Development* (pp. 128–139). Wiley.
- Blasi, A. (1993). The development of identity: Some implications for moral functioning. In G. G. Naom & T. E. Wren (Eds.), *The Moral Self* (pp. 99–122). MIT Press.
- Boegershausen, J., Aquino, K., & Reed, A. (2015). Moral identity. *Current Opinion in Psychology*, 6, 162–166. <https://doi.org/10.1016/J.COPSYC.2015.07.017>
- Bolger, N., & Schilling, E. A. (1991). Personality and the problems of everyday life: The role of neuroticism in exposure and reactivity to daily stressors. *Journal of Personality*, 59(3), 355–386. <https://doi.org/10.1111/J.1467-6494.1991.TB00253.X>
- Bolger, N., & Zuckerman, A. (1995). A framework for studying personality in the stress process. *Journal of Personality and Social Psychology*, 69(5), 890. <https://doi.org/10.1037/0022-3514.69.5.890>
- Bolino, M. C. (1999). Citizenship and impression management: Good soldiers or good actors? *Academy of Management Review*, 24(1), 82–98. <https://doi.org/10.5465/AMR.1999.1580442>
- Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, 34(6), 1080–1109. <https://doi.org/10.1177/0149206308324325>
- Bolino, M. C., & Turnley, W. H. (1999). Measuring impression management in organizations: A scale development based on the Jones and Pittman taxonomy. *Organizational Research Methods*, 2(2), 187–206. <https://doi.org/10.1177/109442819922005>
- Bolino, M. C., & Turnley, W. H. (2003). More than one way to make an impression: Exploring profiles of impression management. *Journal of Management*, 29(2), 141–160. <https://doi.org/10.1177/014920630302900202>
- Bonner, J. M., Greenbaum, R. L., & Quade, M. J. (2017). Employee unethical behavior to shame as an indicator of self-image threat and exemplification as a form of self-image protection: The exacerbating role of supervisor bottom-line mentality. *Journal of Applied Psychology*, 102(8), 12–13.
- Braithwaite, V., Ahmed, E., & Braithwaite, J. (2008). Workplace bullying and victimization: The influence of organizational context, shame and pride. *International Journal of Organisational Behavior*, 13(2), 71–94.
- Breakwell, G. M. (1986). *Coping with Threatened Identities*. Psychology Press.

- Breitsohl, H. (2019). Beyond ANOVA: An introduction to structural equation models for experimental designs. *Organizational Research Methods*, 22(3), 1–29. <https://doi.org/10.1177/1094428118754988>
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Breugelmans, S. M., & Poortinga, Y. H. (2006). Emotion without a word: Shame and guilt among Rarámuri Indians and rural Javanese. *Journal of Personality and Social Psychology*, 91(6), 1111–1122. <https://doi.org/10.1037/0022-3514.91.6.1111>
- Brief, A. P., & Weiss, H. M. (2002). Organizational behavior: Affect in the workplace. *Annual Review of Psychology*, 53, 279–307. <https://doi.org/10.1146/ANNUREV.PSYCH.53.100901.135156>
- Brockner, J. (1988). The effects of work layoffs on survivors: Research, theory, and practice. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior (Vol. 10)* (pp. 213–255). JAI Press.
- Brower, H. H., Lester, S. W., Korsgaard, M. A., & Dineen, B. R. (2009). A closer look at trust between managers and subordinates: Understanding the effects of both trusting and being trusted on subordinate outcomes. *Journal of Management*, 35(2), 327–347. <https://doi.org/10.1177/0149206307312511>
- Brown, B. R. (1970). Face-saving following experimentally induced embarrassment. *Journal of Experimental Social Psychology*, 6(3), 255–271. [https://doi.org/10.1016/0022-1031\(70\)90061-2](https://doi.org/10.1016/0022-1031(70)90061-2)
- Brown, J. D. (2012). Understanding the better than average effect: Motives (still) matter. *Personality and Social Psychology Bulletin*, 38(2), 209–219. <https://doi.org/10.1177/0146167211432763>
- Brühlmann, F., Petralito, S., Aeschbach, L. F., & Opwis, K. (2020). The quality of data collected online: An investigation of careless responding in a crowdsourced sample. *Methods in Psychology*, 2. <https://doi.org/10.1016/j.metip.2020.100022>
- Burke, P. J., & Stets, J. E. (1999). Trust and commitment through self-verification. *Social Psychology Quarterly*, 62(4), 347. <https://doi.org/10.2307/2695833>
- Burke, P. J., & Stets, J. E. (2009). *Identity theory*. Oxford University Press.
- Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, 75(1), 219–229. <https://doi.org/10.1037/0022-3514.75.1.219>
- Campagna, R. L., Mislin, A. A., & Bottom, W. P. (2019). Motivated by guilt and low felt trust: The impact of negotiators' anger expressions on the implementation of negotiated agreements. *Journal of Behavioral Decision Making*, 32(4), 450–470. <https://doi.org/10.1002/bdm.2119>

- Capiola, A., Alarcon, G. M., Gibson, A. M., Jessup, S. A., & Hamdan, I. A. (2022). The same or different? Investigating whether trust and distrust are orthogonal constructs or span a continuum. *Proceedings of the 55th Hawaii International Conference on System Sciences*, 5891–5900. <http://hdl.handle.net/10125/80056>
- Carpenter, N. C., Berry, C. M., & Houston, L. (2014). A meta-analytic comparison of self-reported and other-reported organizational citizenship behavior. *Journal of Organizational Behavior*, 35(4), 547–574.
- Carretta, T. R., & Ree, M. J. (2022). Correction for range restriction: Lessons from 20 research scenarios. *Military Psychology*, 34(5), 551–569. <https://doi.org/10.1080/08995605.2021.2022067>
- Casado-Aranda, L. A., Dimoka, A., & Sánchez-Fernández, J. (2019). Consumer processing of online trust signals: A neuroimaging study. *Journal of Interactive Marketing*, 47, 159–180. <https://doi.org/10.1016/j.intmar.2019.02.006>
- Cervone, D. (2004). The architecture of personality. *Psychological Review*, 111, 183–204.
- Cervone, D., & Shoda, Y. (1999). *The coherence of personality: Social-cognitive bases of consistency, variability, and organization* (D. Cervone & Y. Shoda, Eds.). Guilford Press.
- Cervone, D., & Tripathi, R. (2009). The moral functioning of the person as a whole: On moral psychology and personality science. In D. Narvaez & D. Lapsley (Eds.), *Personality, Identity, and Character: Explorations in Moral Psychology* (pp. 30–51).
- Cha, Y., & Wong-On-Wing, B. (2020). The effect of formal control justification on dysfunctional behavior. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3616529>
- Champeyrache, C. (2022). Institutional mistrust, instrumental trust, and the privatization of law: The mafia as a territorial ruler. *Journal of Economic Issues*, 56(4), 945–958. <https://doi.org/10.1080/00213624.2022.2111139>
- Chang, Y.-Sh., & Fang, S.-R. (2013). Antecedents and distinctions between online trust and distrust: Predicting high- and low-risk internet behaviors. *Journal of Electronic Commerce Research*, 14(2), 149–166. <http://web.csulb.edu/journals/jecr/issues/20132/paper2.pdf>
- Chen, A. (2020). Moral threat or moral elevation? A regulatory focus perspective on observer reactions to ethical voice. [Doctoral Dissertation]
- Chen, X., Zhu, Z., & Liu, J. (2019). Does a trusted leader always behave better? The relationship between leader feeling trusted by employees and benevolent and laissez-faire leadership behaviors. *Journal of Business Ethics*, 1, 3. <https://doi.org/10.1007/s10551-019-04390-7>
- Chiu, H. C., & Chiang, P. H. (2019). A trickle-down effect of subordinates' felt trust. *Personnel Review*, 48(4), 957–976. <https://doi.org/10.1108/PR-01-2018-0036>
- Cho, J. (2006). The mechanism of trust and distrust formation and their relational outcomes. *Journal of Retailing*, 82(1), 25–35. <https://doi.org/10.1016/J.JRETAI.2005.11.002>

- Cho, Y., & Jiang, W. Y. (2022). How work orientation impacts objective career outcomes via managerial (mis)perceptions. *Academy of Management Journal*, 65(4), 1353–1382. <https://doi.org/10.5465/amj.2020.0841>
- Christ, M. H. (2013). An experimental investigation of the interactions among intentions, reciprocity, and control. *Journal of Management Accounting Research*, 25(1), 169–197. <https://doi.org/10.2308/jmar-50443>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum.
- Cohen, T. R., Panter, A. T., Turan, N., Morse, L., & Kim, Y. (2014). Moral character in the workplace. *Journal of Personality and Social Psychology*, 107(5), 943.
- Cohen, T. R., Wolf, S. T., Panter, A. T., & Insko, C. A. (2011). Introducing the GASP Scale: A new measure of guilt and shame proneness. *Journal of Personality and Social Psychology*, 100(5), 947–966. <https://doi.org/10.1037/a0022641>
- Coleman, S. R., Pincus, A. L., & Smyth, J. M. (2019). Narcissism and stress-reactivity: A biobehavioural health perspective. *Health Psychology Review*, 13(1), 35-72.
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.
- Connelly, B. L., Miller, T., & Devers, C. E. (2012). Under a cloud of suspicion: trust, distrust, and their interactive effect in interorganizational contracting. *Strategic Management Journal*, 33(7), 820–833. <https://doi.org/10.1002/smj.974>
- Cooley, C. H. (1902). *Human nature and the social order*. Scribner.
- Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, 130(3), 392–414. <https://doi.org/10.1037/0033-2909.130.3.392>
- Deffenbacher, J. L. (2011). Cognitive-behavioral conceptualization and treatment of anger. *Cognitive and Behavioral Practice*, 18, 212–221. www.elsevier.com/locate/cabp
- Daniels, M. A., & Robinson, S. L. (2019). The shame of it all: A review of shame in organizational life. *Journal of Management*, 45(6), 2448–2473. <https://doi.org/10.1177/0149206318817604>
- Dansereau, F., Graen, G., & Haga, W.J. (1975). A vertical dyad linkage approach to leadership within formal organizations: A longitudinal investigation of the role making process. *Organizational Behavior and Human Performance*, 13(1), 46-78
- Darwin, C. (1872). *The expression of emotion in men and animals*. Scribners.
- D’Cruz, J. (2019). Humble trust. *Philosophical Studies*, 176(4), 933–953. <https://doi.org/10.1007/s11098-018-1220-6>
- D’Cruz, J. (2020). Trust and distrust. In *The Routledge handbook of trust and philosophy* (pp. 41–51). Routledge. <https://doi.org/10.4324/9781315542294-3>

- De Cuyper, N., Schreurs, B., Vander Elst, T., Baillien, E., & De Witte, H. (2014). Exemplification and perceived job insecurity: Associations with self-rated performance and emotional exhaustion. *Journal of Personnel Psychology*, 13(1 A), 1–10. <https://doi.org/10.1027/1866-5888/a000099>
- de Hooge, I. E., Breugelmans, S. M., Wagemans, F. M. A., & Zeelenberg, M. (2018). The social side of shame: Approach versus withdrawal. *Cognition and Emotion*, 32(8), 1671–1677. <https://doi.org/10.1080/02699931.2017.1422696>
- de Hooge, I. E., Breugelmans, S. M., & Zeelenberg, M. (2008). Not so ugly after all: When shame acts as a commitment device. *Journal of Personality and Social Psychology*, 95(4), 933–943. <https://doi.org/10.1037/a0011991>
- de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2010). Restore and protect motivations following shame. *Cognition and Emotion*, 24(1), 111–127. <https://doi.org/10.1080/02699930802584466>
- de Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2011). A functionalist account of shame-induced behaviour. *Cognition and Emotion*, 25(5), 939–946. <https://doi.org/10.1080/02699931.2010.516909>
- Deci, E. L., Connell, J. P., & Ryan, R. M. (1989). Self-determination in a work organization. In *Journal of Applied Psychology* (Vol. 74, Issue 4).
- Deng, J., & Wang, K. Y. (2009). Feeling trusted and loyalty: Modeling supervisor-subordinate interaction from a trustee perspective. *The International Employment Relations Review*, 15(1), 16–38. <https://opus.lib.uts.edu.au/handle/10453/10405>
- Deng, L., & Yuan, K. H. (2023). Which method is more powerful in testing the relationship of theoretical constructs? A meta comparison of structural equation modeling and path analysis with weighted composites. *Behavior Research Methods*, 55(3), 1460–1479. <https://doi.org/10.3758/s13428-022-01838-z>
- DeSimone, J. A., & Harms, P. D. (2018). Dirty data: The effects of screening respondents who provide low-quality data in survey research. *Journal of Business and Psychology*, 33(5), 559–577. <https://doi.org/10.1007/s10869-017-9514-9>
- Devine, D., Gaskell, J., Jennings, W., & Stoker, G. (2020). *Exploring trust, mistrust and distrust* (Discussion paper for the 1st Digital Workshop of the ESRC “TrustGov” project; Working Paper Series).
- Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly*, 34(2), 373–396.
- Douglas, S. C., Kiewitz, C., Martinko, M. J., Harvey, P., Kim, Y., & Uhm, J. (2008). Cognitions, emotions, and evaluations: An elaboration likelihood model for workplace aggression. *Academy of Management Review*, 33(2), 425–451. <https://doi.org/10.5465/amr.2008.31193490>

- Douglas, S. C., & Martinko, M. J. (2001). Exploring the role of individual differences in the prediction of workplace aggression. *Journal of Applied Psychology*, 86, 547–559. doi:10.1037/0021-9010.86.4.547
- Durik, A. M., Hyde, J. S., Marks, A. C., Roy, A. L., Anaya, D., & Schultz, G. (2006). Ethnicity and gender stereotypes of emotion. *Sex Roles*, 54(7), 429–445.
- Dutton, J. E., Roberts, L. M., & Bednar, J. (2010). Pathways for positive identity construction at work: Four types of positive identity and the building of social resources. *Academy of Management Review*, 35(2), 265–293. <https://doi.org/10.5465/amr.35.2.zok265>
- Eerola, T., Armitage, J., Lavan, N., & Knight, S. (2021). Online data collection in auditory perception and cognition research: Recruitment, testing, data quality and ethical considerations. *Auditory Perception & Cognition*, 4(3–4), 251–280. <https://doi.org/10.1080/25742442.2021.2007718>
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.
- Ellemers, N. (2012). The group self. *Science*, 336(6083), 848–852. <https://doi.org/10.1126/SCIENCE.1220987>
- Ellemers, N., Doosje, B., & Spears, R. (2004). Sources of respect: The effects of being liked by ingroups and outgroups. *European Journal of Social Psychology*, 34(2), 155–172. <https://doi.org/10.1002/ejsp.196>
- Else-Quest, N. M., Higgins, A., Allison, C., & Morton, L. C. (2012). Gender differences in self-conscious emotional experience: a meta-analysis. *Psychological bulletin*, 138(5), 947.
- Enzle, M. E., & Anderson, S. C. (1993). Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology*, 64(2), 257–266. <https://doi.org/10.1037/0022-3514.64.2.257>
- Erez, M., & Early, P. (1993). *Culture, self-identity, and work*. Oxford University Press.
- Erikson, E. H. (1968). *Identity: Youth and crisis*. Norton.
- Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Erikson, E. H. (1968). *Identity, youth and crises*. Norton.
- Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5), 1611–1630. <https://doi.org/10.1257/aer.96.5.1611>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>

- Felson, R. B. (1992). "Kick 'em when they are down": Explanations of the relationship between stress and interpersonal aggression and violence. *Sociological Quarterly*, 33(1), 1–16. <https://doi.org/10.1111/j.1533-8525.1992.tb00360.x>
- Felson, R. B., & Steadman, H. J. (1983). Situational factors in disputes leading to criminal violence. *Criminology*, 21(1), 59–74. <https://doi.org/10.1111/J.1745-9125.1983.TB00251.X>
- Ferris, D. L., Yan, M., Lim, V. K., Chen, Y., & Fatimah, S. (2016). An approach–avoidance framework of workplace aggression. *Academy of management journal*, 59(5), 1777–1800
- Fischer, T., Tian, A. W., Lee, A., & Hughes, D. J. (2021). Abusive supervision: A systematic review and fundamental rethink. *The Leadership Quarterly*, 32(6), 101540.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327–358. <https://doi.org/10.1037/h0061470>
- Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, 3(4), 484–501.
- Flores, A. (2023). Addressing counterproductive work behaviors in a behavioral health organization [*Doctoral dissertation, Walden University*].
- Fontaine, J. R. J., Luyten, P., De Boeck, P., Corveleyn, J., Fernandez, M., Herrera, D., Itzès, A., & Tomcsányi, T. (2006). Untying the Gordian knot of guilt and shame: The structure of guilt and shame reactions based on situation and person variation in Belgium, Hungary, and Peru. *Journal of Cross-Cultural Psychology*, 37(3), 273–292. <https://doi.org/10.1177/0022022105284493>
- Forgas, J. P. (1995). Mood and judgment: The affect infusion model (AIM). *Psychological Bulletin*, 117(1), 39–66. <https://doi.org/10.1037/0033-2909.117.1.39>
- Forgas, J. P., & George, J. M. (2001). Affective influences on judgments and behavior in organizations: An information processing perspective. *Organizational Behavior and Human Decision Processes*, 86(1), 3–34. <https://doi.org/10.1006/obhd.2001.2971>
- Forgas, J. P., & Smith, C. A. (2003). Affect and emotion. In M. A. Hogg & J. Cooper (Eds.), *The Sage handbook of social psychology* (pp. 161–189). Sage.
- Forgays, D. G., Forgays, D. K., & Spielberger, C. D. (1997). Factor structure of the state-trait anger expression inventory. *Journal of Personality Assessment*, 69(3), 497–507. https://doi.org/10.1207/s15327752jpa6903_5
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39–50.
- Fox, S., & Spector, P. E. (1999). A model of work frustration-aggression. *Journal of Organizational Behavior*, 20, 915–931. doi:10.1002/(ISSN)1099-1379

- Fox, S., Spector, P. E., & Miles, D. (2001). Counterproductive work behavior (CWB) in response to job stressors and organizational justice: Some mediator and moderator tests for autonomy and emotions. *Journal of Vocational Behavior*, 59(3), 291–309. <https://doi.org/10.1006/jvbe.2001.1803>
- Frey, B. S. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, 31(4), 663–670. <https://doi.org/10.1111/j.1465-7295.1993.tb00897.x>
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
- Fukuyama, F. (1995). *Trust. The social virtues and the creation of prosperity*. The Penguin Group.
- Fuqua, D. R., Leonard, E., Masters, M. A., Smith, R. J., Campbell, J. L., & Rscher, P. C. (1991). A structural analysis of the State-Trait Anger Expression Inventory. *Educational and Psychological Measurement*, 51, 439–446.
- Gausel, N., & Leach, C. W. (2011). Concern for self-image and social image in the management of moral failure: Rethinking shame. *European Journal of Social Psychology*, 41(4), 468–478. <https://doi.org/10.1002/EJSP.803>
- Gambetta, D. (2000). Mafia: The price of distrust. In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 158–175). Department of Sociology, University of Oxford. <http://www.sociology.ox.ac.uk/papers/gambetta158-175.pdf>
- Garofalo, C., Giammarioli, L., Aiolfi, I., Delvecchio, E., & Mazzeschi, C. (2025). Conceptualization and assessment of shame experience and regulation: An umbrella review of synthesis studies. *Clinical Psychology & Psychotherapy*, 32(4), e70136.
- Gecas, V. (1982). The self-concept. *Annual Review of Sociology*, 8, 1–33. <https://doi.org/10.2307/2945986>
- George, M. M., Strauss, K., Mell, J. N., & Vough, H. C. (2023). When “who I am” is under threat: Measures of threat to identity value, meanings, and enactment. *Journal of Applied Psychology*. <https://doi.org/10.1037/APL0001114>
- Gibson, D. E., & Callister, R. R. (2009). Anger in organizations: Review and integration. *Journal of Management*, 36(1), 66–93. <https://doi.org/10.1177/0149206309348060>
- Gilbert, D. T., & Jones, E. E. (1986). Exemplification: The self-presentation of moral character. *Journal of Personality*, 54(3), 593–615. <https://doi.org/10.1111/J.1467-6494.1986.TB00414.X>
- Gilbert, P. (1997). The evolution of social attractiveness and its role in shame, humiliation, guilt and therapy. *British Journal of Medical Psychology*, 70(2), 113–147. <https://doi.org/10.1111/J.2044-8341.1997.TB01893.X>
- Gill, H., Cassidy, S. A., Cragg, C., Algate, P., Weijjs, C. A., & Finegan, J. E. (2019). Beyond reciprocity: The role of empowerment in understanding felt trust. *European Journal of*

- Work and Organizational Psychology*, 28(6), 845–858.
<https://doi.org/10.1080/1359432X.2019.1678586>
- Gilligan, J. (1996). *Violence: Our deadly epidemic and its causes*. G.P. Putnam.
- Goffman, E. (1955). On face-work. *Psychiatry: Journal for the Study of Interpersonal Processes*, 18, 213–231.
- Goffman, E. (1967). *Interaction ritual*. Pantheon.
- Goldman, B. M. (2003). The application of referent cognitions theory to legal-claiming by terminated workers: The role of organizational justice and anger. *Journal of Management*, 29, 705–728. doi:10.1016/S0149-2063(03)00032-1
- Goleman, D. (1998). The emotional intelligence of leaders. *Leader to Leader*, 1998(10), 20–26.
<https://doi.org/10.1002/ltl.40619981008>
- Gong, B., & Sims, R. L. (2021). The Covid 19 pandemic and psychological contract breach: The disruption of a work from home schedule. *Academy of Management Proceedings*, 2021(1).
<https://doi.org/10.5465/AMBPP.2021.7>
- Grace, C. (2022). *A cycle of distrust: Key findings from the 2022 Edelman Trust Barometer*.
- Greenidge, C. A. (2009). *Patients' trust and distrust of physicians: Polar or separate constructs?* Michigan State University.
- Gruenewald, T. L., Dickerson, S. S., & Kemeny, M. E. (2007). A social function for self-conscious emotions: The social self-preservation theory. In J. L. Tracy, R. W. Robins, & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 68–87). Guilford Press.
- Gunter, L. (2024, September 17). Why don't all Canadians mistrust our government? *Toronto Sun*.
- Guo, S.-L., Lumineau, F., & Lewicki, R. J. (2017). Revisiting the foundations of organizational distrust. *Foundations and Trends in Strategic Management*, 1(1), 1–88.
- Haidt, J. (2003). The moral emotions. In R. J. Davison, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 852–870). Oxford University Press.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- Harari, D., Parke, M. R., & Marr, J. C. (2022). When helping hurts helpers: Anticipatory versus reactive helping, helper's relative status, and recipient self-threat. *Academy of Management Journal*, 65(6), 1954–1983.
- Hardin, R. (2004). *Distrust: Manifestations and Management*. Russell Sage Foundation.

- Harding, N., Azim, M. I., Jidin, R., & Muir, J. P. (2016). A consideration of literature on trust and distrust as they relate to auditor professional scepticism. *Australian Accounting Review*, 26(3), 243–254. <https://doi.org/10.1111/AUAR.12126>
- Hardy, S., & Carlo, G. (2011). Moral identity. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 495–513). Springer. https://doi.org/http://dx.doi.org/10.1007/978-1-4419-7988-9_19
- Hart, D. (2005). The development of moral identity. *Nebraska Symposium on Motivation*, 51, 165–196.
- Hart, D., Atkins, R., & Ford, D. (1998). Urban America as a context for the development of moral identity in adolescence. *Journal of Social Issues*, 54, 513–530.
- Heatherton, T. F., & Baumeister, R. F. (1996). Self-regulation failure: Past, present, and future. *Psychological Inquiry*, 7(1), 90–98. https://doi.org/10.1207/s15327965pli0701_20
- Heese, J., & Pérez-Cavazos, G. (2021). The effect of retaliation costs on employee whistleblowing. *Journal of Accounting and Economics*, 71(2–3). <https://doi.org/10.1016/j.jacceco.2020.101385>
- Henry, J. D., von Hippel, W., Nangle, M. R., & Waters, M. (2018). Age and the experience of strong self-conscious emotion. *Aging & mental health*, 22(4), 497–502.
- Hershcovis, M. S. (2011). “Incivility, social undermining, bullying... oh my!”: A call to reconcile constructs within workplace aggression research. *Journal of Organizational Behavior*, 32(3), 499–519. <https://doi.org/10.1002/job.689>
- Hertz, S. G., & Krettenauer, T. (2016). Does moral identity effectively predict moral behavior?: A meta-analysis. *Review of General Psychology*, 20(2), 129–140. <https://doi.org/10.1037/GPR0000062>
- Hobfoll, S. E. (2001). The influence of culture, community, and the nested-self in the stress process: Advancing Conservation of Resources theory. *Applied Psychology*, 50(3), 337–421. <https://doi.org/10.1111/1464-0597.00062>
- Hogg, M. A., & Terry, D. I. (2000). Social identity and self-categorization processes in organizational contexts. *Academy of Management Review*, 25(1), 121–140. <https://doi.org/10.5465/amr.2000.2791606>
- Homans G. C. (1961). *Social behavior: Its elementary forms*. New York: Harcourt, Brace & World.
- Huang, H. J., & Dastmalchian, A. (2006). Implications of trust and distrust for organizations. *Personnel Review*, 35(4), 361–377. <https://doi.org/10.1108/00483480610670553>
- Huang Johnson, H. (2015). When is a threat more or less of a threat? The sensitivity of psychologically central identities to threat and the resulting impact on work. [Doctoral dissertation, University of Washington].

- Huo, Y. J., Smith, H. J., Tyler, T. R., & Lind, E. A. (1996). Superordinate identification, subgroup identification, and justice concerns: Is separatism the problem; is assimilation the answer? *Psychological Science*, 7(1), 40–45.
- Huo, Y. J., & Tyler, T. R. (2000). How different ethnic groups react to legal authorities. *Law and Society Review*, 34(3), 449–490.
- Ickes, W., & Simpson, J. A. (2001). Motivational goals and the regulation of interpersonal relationships. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell handbook of social psychology: Interpersonal processes* (pp. 273–293). Blackwell Publishing.
- Ilgen, D. R., & Pulakos, E. D. (1999). The changing nature of performance: Implications for staffing, motivation, and development. Jossey-Bass.
- Ilies, R., Schwind, K. M., & Heller, D. (2007). Employee well-being: A multilevel model linking work and nonwork domains. *European Journal of Work and Organizational Psychology*, 16(3), 326–341. <https://doi.org/10.1080/13594320701363712>
- Inesi, M. E., Gruenfeld, D. H., & Galinsky, A. D. (2012). How power corrupts relationships: Cynical attributions for others' generous acts. *Journal of Experimental Social Psychology*, 48(4), 795–803. <https://doi.org/10.1016/j.jesp.2012.01.008>
- Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision-making, and attention. *Journal of Personality and Social Psychology*, 99(3), 467–481. <https://doi.org/10.1037/a0018951>
- Isaeva, N., Hughes, C., & Saunders, M. N. K. (2019). Trust, distrust and human resource management. In *Elgar Introduction to Theories of Human Resources and Employment Relations* (pp. 247–263). Edward Elgar Publishing. <https://doi.org/10.4337/9781786439017.00025>
- Izard, C. E. (1991). *The psychology of emotions*. Plenum Press.
- James, W. (1890). *Principles of Psychology*. Holt.
- Jason Huang, H., & Dastmalchian, A. (2006). Implications of trust and distrust for organizations: Role of customer orientation in a four-nation study. *Personnel Review*, 35(4), 361–377. <https://doi.org/10.1108/00483480610670553>
- Jensen, J. M., Opland, R. A., & Ryan, A. M. (2010). Psychological contracts and counterproductive work behaviors: Employee responses to transactional and relational breach. *Journal of Business and Psychology*, 25(4), 555–568. <https://doi.org/10.1007/S10869-009-9148-7/TABLES/3>
- Jermier, J. M., Knights, D., & Nord, W. (1994). *Resistance and power in organizations*. Routledge.

- Jiang, J. J., & Klein, G. (1999). Project risks and the practice of software development risk management. *Journal of Systems and Software*, 52(1), 3–10. [https://doi.org/10.1016/S0164-1212\(99\)00002-9](https://doi.org/10.1016/S0164-1212(99)00002-9)
- Jones, D. A. (2009). Getting even with one's supervisor and one's organization: Relationships among types of injustice, desires for revenge, and counterproductive work behaviors. *Journal of Organizational Behavior*, 30(4), 525–542. <https://doi.org/10.1002/job.563>
- Jones, E. E., & Pittman, T. S. (1982). Toward a general theory of strategic self-presentation. In J. Suls (Ed.), *Psychological perspectives on the Self (Vol. 1)* (pp. 231–262). Erlbaum.
- Jones, D. A., & Skarlicki, D. P. (2013). How perceptions of inconsistency in organizational justice climate influence work outcomes: A longitudinal perspective. *Journal of Organizational Behavior*, 34(4), 514–535. <https://doi.org/10.1002/job.1814>
- Jordan, J., Mullen, E., & Murnighan, J. K. (2011). Striving for the moral self: The effects of recalling past moral actions on future moral behavior. *Personality and Social Psychology Bulletin*, 37(5), 701–713. <https://doi.org/10.1177/0146167211400208>
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The core self-evaluations scale: Development of a measure. *Personnel Psychology*, 56(2), 303–331. <https://doi.org/10.1111/j.1744-6570.2003.tb00152.x>
- Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, 89(1), 36–51. <https://doi.org/10.1037/0021-9010.89.1.36>
- Karakowsky, L., Degama, N., & Mcbey, K. (2012). Facilitating the Pygmalion effect: The overlooked role of subordinate perceptions of the leader. *Journal of Occupational and Organizational Psychology*, 85(4), 579–599. <https://doi.org/10.1111/j.2044-8325.2012.02056.x>
- Keltner, D., & Buswell, B. N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion*, 10(2), 155–172. <https://doi.org/10.1080/026999396380312>
- Keltner, D., & Harker, L. A. (1998). The forms and functions of the nonverbal signal of shame. In P. Gilbert & B. Andrews (Eds.), *Shame: Interpersonal behavior, psychopathology, and culture* (pp. 78–98). Oxford University Press.
- Kerekes, J., Chow, J., Lemak, A., & Perhan, Z. (2013). Trust or betrayal: Immigrant engineers' employment-seeking experiences in Canada. In *Discourses of trust* (pp. 269–284). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-137-29556-9_17
- Kernis, M. H., Grannemann, B. D., & Barclay, L. C. (1989). Stability and level of self-esteem as predictors of anger arousal and hostility. In *Journal of Personality and Social Psychology* (Vol. 6, Issue 6).

- Khan, A. K., Bell, C. M., & Quratulain, S. (2017). The two faces of envy: Perceived opportunity to perform as a moderator of envy manifestation. *Personnel Review*, 46(3), 490–511. <https://doi.org/10.1108/PR-12-2014-0279>
- Khan, A. K., Quratulain, S., & Crawshaw, J. R. (2013). The mediating role of discrete emotions in the relationship between injustice and counterproductive work behaviors: A study in Pakistan. *Journal of Business and Psychology*, 28(1), 49–61. <https://doi.org/10.1007/s10869-012-9269-2>
- Kim, J., Lee, H. W., & Johnson, R. E. (2021). When CEOs are all about themselves: Perceived CEO narcissism and middle managers' workplace behaviors amid the COVID-19 pandemic. *Journal of Applied Psychology*, 106(9), 1283–1298. <https://doi.org/10.1037/apl0000965>
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Kim, S. H., Smith, R. H., & Brigham, N. L. (1998). Effects of power imbalance and the presence of third parties on reactions to harm: Upward and downward revenge. *Personality and Social Psychology Bulletin*, 24(4), 353–361. <https://doi.org/10.1177/0146167298244002>
- Komiak, S., & Benbasat, I. (2008). A two-process view of trust and distrust building in recommendation agents: A process-tracing study. *Journal of the Association for Information Systems*, 9(12), 727–747. <https://doi.org/10.17705/1jais.00180>
- Korsgaard, M. A. (2018). Reciprocal trust: A self-reinforcing dynamic process. In R. H. Searle, A.-M. I. Nienaber, & S. B. Sitkin (Eds.), *The Routledge companion to trust* (pp. 14–28). Taylor & Francis Group.
- Koslowsky, M., & Krausz, M. (Eds.). (2002). *Voluntary employee withdrawal and inattendance: A current perspective*. Springer Science + Business Media. <https://doi.org/10.1007/978-1-4615-0599-0>
- Kostis, A., & Näsholm, M. H. (2020). Towards a research agenda on how, when and why trust and distrust matter to competition. *Journal of Trust Research*, 10(1), 66–90. <https://doi.org/10.1080/21515581.2019.1692664>
- Kramer, R. M. (1994). The sinister attribution error: Paranoid cognition and collective distrust in organizations. *Motivation and Emotion*, 18(2), 199–230. <https://doi.org/10.1007/BF02249399>
- Kramer, R. M. (1995). Power, paranoia, and distrust in organizations: The distorted view from the top. In R. Bies, R. Lewicki, & B. Sheppard (Eds.), *Research on Negotiation in Organizations (vol.5)* (pp. 119–154). Emerald Group Publishing Limited.

- Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology*, 50, 569–598. <https://www.annualreviews.org/doi/pdf/10.1146/annurev.psych.50.1.569>
- Kramer, R. M. (2004). Collective paranoia: Distrust between social groups. In R. Hardin (Ed.), *Distrust* (pp. 136–167). Russell.
- Kramer, R. M., & Cook, K. (2004). *Trust and distrust in organizations: Dilemmas and approaches*. Russell Sage.
- Krehbiel, P. J., & Cropanzano, R. (2000). Procedural justice, outcome favorability and emotion. *Social Justice Research*, 13(4), 339–360. <https://doi.org/10.1023/A:1007670909889>
- Kroeper, K. M., Williams, H. E., & Murphy, M. C. (2022). Counterfeit diversity: How strategically misrepresenting gender diversity dampens organizations’ perceived sincerity and elevates women’s identity threat concerns. *Journal of Personality and Social Psychology*, 122(3), 399–426. <https://doi.org/10.1037/pspi0000348>
- Kruglanski, A. W. (1970). Attributing trustworthiness in supervisor-worker relations. *Journal of Experimental Social Psychology*, 6(2), 214–232. [https://doi.org/10.1016/0022-1031\(70\)90088-0](https://doi.org/10.1016/0022-1031(70)90088-0)
- Kujala, J., Lehtimäki, H., & Pučetaité, R. (2016). Trust and distrust constructing unity and fragmentation of organisational culture. *Journal of Business Ethics*, 139(4), 701–716. <https://doi.org/10.1007/s10551-015-2915-7>
- Kundro, T. G., & Nurmohamed, S. (2021). Understanding when and why cover-ups are punished less severely. *Academy of Management Journal*, 64(3), 873–900.
- Kuppens, P., & Tuerlinckx, F. (2007). Personality traits predicting anger in self-, ambiguous-, and other caused unpleasant situations. *Personality and Individual Differences*, 42(6), 1105–1115.
- Kurtines, W. M., & Gewirtz, J. L. (1984). *Morality, moral behavior and moral development*. John Wiley & Sons.
- Lanaj, K., Kim, P. H., Koopman, J., & Matta, F. K. (2018). Daily mistrust: A resource perspective and its implications for work and home. *Personnel Psychology*, 71(4), 545–570. <https://doi.org/10.1111/peps.12268>
- Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology*, 8(2), 142–164. <https://doi.org/10.1017/iop.2015.13>
- Lapsley, D. K., & Narvaez, D. (2004). A social-cognitive approach to the moral personality. In D. K. Lapsley & C. Power (Eds.), *Moral development, self and identity* (pp. 189–212). Lawrence Erlbaum Associates.

- Lau, D. C., & Lam, L. W. (2008). Effects of trusting and being trusted on team citizenship behaviours in chain stores. *Asian Journal of Social Psychology*, 11(2), 141–149. <https://doi.org/10.1111/j.1467-839X.2008.00251.x>
- Lau, D. C., Lam, L. W., & Wen, S. S. (2014). Examining the effects of feeling trusted by supervisors in the workplace: A self-evaluative perspective. *Journal of Organizational Behavior*, 35(1), 112–127. <https://doi.org/10.1002/job.1861>
- Lau, D. C., Liu, J., & Fu, P. P. (2007). Feeling trusted by business leaders in China: Antecedents and the mediating role of value congruence. *Asia Pacific Journal of Management*, 24(3), 321–340. <https://doi.org/10.1007/s10490-006-9026-z>
- Lazarus, R. S. (1999). *Stress and emotion: A new synthesis*. Springer.
- Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior*. Brown & Benchmark.
- Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, 58, 317–344. <https://doi.org/10.1146/ANNUREV.PSYCH.58.110405.085658>
- Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, 107(1), 34–47. <https://doi.org/10.1037/0033-2909.107.1.34>
- Leary, M. R., Landel, J. L., & Patton, K. M. (1996). The motivated expression of embarrassment following a self-presentational predicament. *Journal of Personality*, 64(3), 619–636. <https://doi.org/10.1111/J.1467-6494.1996.TB00524.X>
- Leary, M. R., & Miller, R. S. (2000). Self-presentational perspectives on personal relationships. In W. Ickes & S. Duck (Eds.), *The social psychology of personal relationships* (pp. 129–154). Wiley.
- Leary, M. R., Tchividjian, L. R., & Kraxberger, B. E. (1994). Self-presentation can be hazardous to your health: Impression management and health risk. *Health Psychology*, 13(6), 461.
- Leary, M. R., Twenge, J. M., & Quinlivan, E. (2006). Interpersonal rejection as a determinant of anger and aggression. *Personality and Social Psychology Review*, 10(2), 111–132. https://doi.org/10.1207/s15327957pspr1002_2
- Leavitt, K., Zhu, L. (L.), Klotz, A., & Kouchaki, M. (2022). Fragile or robust? Differential effects of gender threats in the workplace among men and women. *Organizational Behavior and Human Decision Processes*, 168, 104112. <https://doi.org/10.1016/J.OBHDP.2021.104112>
- Lee, J. (2021). When illusion met illusion: How interacting biases affect (dis)trust within coopetitive policy networks. *Public Administration Review*, 81(5), 962–972. <https://doi.org/10.1111/puar.13351>

- Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: The role of affect and cognitions. *Journal of Applied Psychology*, 87, 131–142. doi:10.1037/0021-9010.87.1.131
- Lee, J., & Dodge, J. (2019). Keeping your enemies close: The role of distrust in structuring a local hydraulic fracturing policy network in New York. *Journal of Public Administration Research and Theory*, 29(2), 175–192. <https://doi.org/10.1093/jopart/muy074>
- Lee, J., & Lee, J. (2024). Compatibility of the incompatible: How does asymmetric power lead to coexistence of trust and distrust in adversarial policy networks? *International Journal of Public Administration*, 47(5), 297–312. <https://doi.org/10.1080/01900692.2022.2094410>
- Lehman, W. E. K., & Simpson, D. D. (1992). Employee substance use and on-the-job behaviors. *Journal of Applied Psychology*, 77(3), 309–321. <https://doi.org/10.1037/0021-9010.77.3.309>
- LePine, J. A., Podsakoff, N. P., & LePine, M. A. (2005). A meta-analytic test of the challenge stressor–hindrance stressor framework: An explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal*, 48(5), 764–775. <https://doi.org/10.5465/amj.2005.18803921>
- Lester, S. W., & Brower, H. H. (2003). In the eyes of the beholder: The relationship between subordinates' felt trustworthiness and their work attitudes and behaviors. *Journal of Leadership & Organizational Studies*, 10(2), 17–33. <https://doi.org/10.1177/107179190301000203>
- Levi, M. (2000). When good defenses make good neighbors: A transaction cost approach to trust, the absence of trust and distrust. In C. Menard (Ed.), *Institutions, Contracts and Organizations. Perspectives from New Institutional Economics* (pp. 137–157). Edward Elgar.
- Lewicki, R. J. (2007). Trust and distrust. In A. K. Schneider & C. Honeyman (Eds.), *The negotiator's fieldbook* (pp. 191–202). American Bar Association.
- Lewicki, R. J. (2021). *Roy Lewicki on why trust matters (First Center for Trust Studies workshop) [video]*. The Center for Trust Studies. Eller College of Management, The University of Arizona. <https://eller.arizona.edu/departments-research/centers-labs-institutes/center-trust-studies/curricula>
- Lewicki, R. J., Elgoibar, P., & Euwema, M. (2016). Trust, distrust, and trust repair in organizational industrial relations. In *Building trust and constructive conflict management in organizations* (pp. 93–117).
- Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, 23(3), 438–458. <https://doi.org/10.5465/amr.1998.926620>
- Lewis, H. B. (1971). *Shame and guilt in neurosis*. International Universities Press.

- Lewis, M. (1992). *Shame: The exposed self*. Free Press.
- Liden, R.C. & Maslyn, J.M. (1998). Multidimensionality of leader-member exchange: An empirical assessment through scale development. *Journal of Management*, 24(1), 43-72.
- Lin, G. C., Wen, Z., Marsh, H. W., & Lin, H. S. (2010). *Structural equation models of latent interactions: Clarification of orthogonalizing and double-mean-centering strategies. Structural Equation Modeling*. 17(3), 374–391.
- Lin, J. (n.d.). *Introduction to Structural Equation Modeling (SEM) in R with lavaan*. UCLA: Statistical Consulting Group. <https://stats.oarc.ucla.edu/r/seminars/rsem/#s3b> (Accessed October 25, 2023).
- Lind, E. A., & Tyler, T. R. (1988). *The social psychology of procedural justice*. Plenum.
- Linde, J., & Vis, B. (2017). Do politicians take risks like the rest of us? An experimental test of prospect theory under MPs. *Political Psychology*, 38(1), 101–117. <https://doi.org/10.1111/pops.12335>
- Lindsay-Hartz, J. (1984). Contrasting experiences of shame and guilt. *American Behavioral Scientist*, 27, 689–704.
- Lissak, R. I., & Sheppard, B. H. (1983). Beyond fairness: The criterion problem in research on dispute intervention. *Journal of Applied Social Psychology*, 13(1), 45–65. <https://doi.org/10.1111/j.1559-1816.1983.tb00886.x>
- Litzky, B. E., Eddleston, K. A., & Kidder, D. L. (2006). The good, the bad, and the misguided: How managers inadvertently encourage deviant behaviors. *Academy of Management Perspectives*, 20, 91–103. <https://doi.org/10.5465/amp.2006.19873411>
- Liu, Y., Wang, X., Chen, Y., Zhang, R., & Zhang, J. (2024). Does the feeling of ability-distrusted motivate employees? Investigating its dual effects on employees' behaviors and the roles of shame and LMX. *Current Psychology*, 43(23), 20417–20430. <https://doi.org/10.1007/s12144-024-05840-w>
- London, M., & Smither, J. W. (2009). Diagnosing, understanding, and dealing with counterproductive work behavior. In *Performance management* (Vol. 21). Center for Creative Leadership.
- Long, D. M. (2017). A method to the martyrdom: Employee exemplification as an impression management strategy. *Organizational Psychology Review*, 7(1), 36–65. <https://doi.org/10.1177/2041386616663816>
- Long, C. P., & Sitkin, S. B. (2018). Control–trust dynamics in organizations: Identifying shared perspectives and charting conceptual fault lines. *Academy of Management Annals*, 12(2), 725–751. <https://doi.org/10.5465/annals.2016.0055>
- Luhmann, N. (1979). *Trust and Power*. Wiley.

- Lumineau, F. (2017). How contracts influence trust and distrust. *Journal of Management*, 43(5), 1553–1577. <https://doi.org/10.1177/0149206314556656>
- Lyubych, Z., Bozeman, J., Hershcovis, M. S., Turner, N., & Shan, J. V. (2022). Employee performance and abusive supervision: The role of supervisor over-attributions. *Journal of Organizational Behavior*, 43(1), 125–145. <https://doi.org/10.1002/JOB.2560>
- MacInnis, C. C., Boss, H. C. D., & Bourdage, J. S. (2020). More evidence of participant misrepresentation on Mturk and investigating who misrepresents. *Personality and Individual Differences*, 152. <https://doi.org/10.1016/j.paid.2019.109603>
- Mangione, T. W., & Quinn, R. P. (1975). Job satisfaction, counterproductive behavior, and drug use at work. *Journal of Applied Psychology*, 60(1), 114–116. <https://doi.org/10.1037/H0076355>
- Marineau, J. E. (2017). Trust and distrust network accuracy and career advancement in an organization. *Group & Organization Management*, 42(4), 487–520. <https://doi.org/10.1177/1059601115627529>
- Marschall, D., Saftner, J., & Tangney, J. P. (1994). *The State Shame and Guilt Scale*. George Mason University. <https://gospel-app.com/wp-content/uploads/2018/10/SSGS.pdf>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734. <https://doi.org/10.5465/amr.1995.9508080335>
- McCall, G. J., & Simmons, J. L. (1978). *Identities and Interactions: An Examination of Human Associations in Everyday Life*. Free Press.
- McFall, L. (1987). Integrity. *Ethics*, 98, 5–20.
- McKnight, D. H., & Chervany, N. L. (2001a). While trust is cool and collected, distrust is fiery and frenzied: A model of distrust concepts. *AMCIS 2001 Proceedings*, 883–888. <http://aisel.aisnet.org/amcis2001/17>
- McKnight, D. H., & Chervany, N. L. (2001b). Trust and distrust definitions: One bite at a time. In R. Falcone, M. Singh, & Y.-H. Tan (Eds.), *Trust in cyber-societies: Integrating the human and artificial perspectives* (pp. 27–54). Springer. <https://msu.edu/user/mcknig26/Trust%20and%20Distrust%20Definitions.pdf>
- McKnight, D. H., & Choudhury, V. (2006). Distrust and trust in B2C e-commerce: Do they differ? *Proceedings of the 8th International Conference on Electronic Commerce The New E-Commerce: Innovations for Conquering Current Barriers, Obstacles and Limitations to Conducting Successful Business on the Internet - ICEC '06*, 482–491. <https://doi.org/10.1145/1151454.1151527>
- McKnight, D. H., Kacmar, C. J., & Choudhury, V. (2004). Dispositional trust and distrust distinctions in predicting high- and low-risk internet expert advice site perceptions. *E-Service Journal*, 3(2), 35–55. <https://doi.org/10.2979/esj.2004.3.2.35>

- McKnight, D. H., Lankton, N. K., Nicolaou, A., & Price, J. (2017). Distinguishing the effects of B2B information quality, system quality, and service outcome quality on trust and distrust. *Journal of Strategic Information Systems*, 26(2), 118–141. <https://doi.org/10.1016/j.jsis.2017.01.001>
- Mead, G. H. (1934). *Mind, self, and society*. University of Chicago Press.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Mendoza, J. P., & Wielhouwer, J. L. (2024). Compliance under distrust: Do people comply less when they feel distrusted? *Journal of Public Administration Research and Theory*, 34(1), Article muae025.
- Mercado, B. K., Dilchert, S., Giordano, C., & Ones, D. S. (2018). Counterproductive work behaviors. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE handbook of industrial, work and organizational psychology: Personnel psychology and employee performance* (2nd ed., pp. 109–211). Sage. <https://doi.org/10.4135/9781473914940.n7>
- Meyers, L. S., Gamst, G., & Guarino, A. J. (2017). *Applied multivariate research: Design and interpretation* (3rd ed.). SAGE Publications.
- Miron, A. M., & Brehm, J. W. (2006). Reactance Theory - 40 Years Later. In *Zeitschrift fur Sozialpsychologie* (Vol. 37, Issue 1, pp. 9–18). Verlag Hans Huber. <https://doi.org/10.1024/0044-3514.37.1.9>
- Miscenko, D., & Day, D. V. (2016). Identity and identification at work. *Organizational Psychology Review*, 6(3), 215–247. <https://doi.org/10.1177/2041386615584009>
- Mitchell, M. S., Vogel, R. M., & Folger, R. (2015). Third parties' reactions to the abusive supervision of coworkers. *Journal of Applied Psychology*, 100(4), 1040–1055. <https://doi.org/10.1037/APL0000002>
- Mohr, C. D., Haverly, S. N., Froidevaux, A., & Wang, M. (2017). Escaping the self: Negative self-evaluations and employee alcohol misuse. In *The Self at work* (pp. 273–292). Routledge.
- Molines, M., & Perrier, A. (2024). Managing my shame: Investigating the effects of organizational identity threat and public service motivation on exemplification. *Public Administration Review*, 84(4), 726–747. <https://doi.org/10.1111/puar.13728>
- Moody, G. D., Galletta, D. F., & Lowry, P. J. (2014). When trust and distrust collide online: The engenderment and role of consumer ambivalence in online consumer behavior. *Electronic Commerce Research and Applications*, 13(4), 266–282. <https://doi.org/https://doi.org/10.1016/j.elerap.2014.05.001>

- Mooijman, M., & Graham, J. (2018). Unjust punishment in organizations. In B. Staw & A. P. Brief (Eds.), *Research in organizational behavior* (Vol. 38, pp. 95–106). JAI Press.
<https://doi.org/10.1016/j.riob.2018.10.001>
- Mooijman, M., van Dijk, W. W., Ellemers, N., & van Dijk, E. (2015). Why leaders punish: A power perspective. *Journal of Personality and Social Psychology*, 109(1), 75–89.
<https://doi.org/10.1037/pspi0000021>.
- Mooijman, M., van Dijk, W. W., van Dijk, E., & Ellemers, N. (2017). On sanction-goal justifications: How and why deterrence justifications undermine rule compliance. *Journal of Personality and Social Psychology*, 112(4), 577–588.
<https://doi.org/10.1037/pspi0000084>
- Mooijman, M., van Dijk, W. W., van Dijk, E., & Ellemers, N. (2019). Leader power, power stability, and interpersonal trust. *Organizational Behavior and Human Decision Processes*, 152, 1–10. <https://doi.org/10.1016/j.obhdp.2019.03.009>
- Mulder, L. B., & Aquino, K. (2013). The role of moral identity in the aftermath of dishonesty. *Organizational Behavior and Human Decision Processes*, 121(2), 219–230.
<https://doi.org/10.1016/J.OBHDP.2013.03.005>
- Myin-Germeys, I., & Kuppens, P. (2021). *The open handbook of experience sampling methodology*. The Center for Research on Experience Sampling and Ambulatory Methods Leuven (REAL) - Belgium.
- Nerstad, C. G. L., Searle, R., Černe, M., Dysvik, A., Škerlavaj, M., & Scherer, R. (2018). Perceived mastery climate, felt trust, and knowledge sharing. *Journal of Organizational Behavior*, 39(4), 429–447. <https://doi.org/10.1002/job.2241>
- Neuman, J. H., & Baron, R. A. (1998). Workplace violence and workplace aggression: Evidence concerning specific forms, potential causes, and preferred targets. *Journal of Management*, 24(3), 391–419. <https://doi.org/10.1177/014920639802400305>
- Norquay, G. (2022, November 30). Canada's growing problem with trust in government. *Policy. Canadian Politics and Public Policy*.
- Nurmohamed, S. (2020). The underdog effect: When low expectations increase performance. *Academy of Management Journal*, 63(4), 1106–1133.
<https://doi.org/10.5465/amj.2017.0181>
- Ohly, S., Sonnentag, S., Niessen, C., & Zapf, D. (2010). Diary studies in organizational research: An introduction and some practical recommendations. *Journal of Personnel Psychology*, 9(2), 79–93. <https://doi.org/10.1027/1866-5888/a000009>
- Onwuegbuzie, A. J., & Daniel, L. G. (2002). A framework for reporting and interpreting internal consistency reliability estimates. *Measurement and Evaluation in Counseling and Development*, 35(2), 89–103. <https://doi.org/10.1080/07481756.2002.12069052>

- Organ, D. W. (1990). The motivational basis of organizational citizenship behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 12, pp. 43–72). JAI Press.
- Organ, D. W., Podsakoff, P. M., & MacKenzie, S. B. (2006). *Organizational citizenship behavior: Its nature, antecedents, and consequences*. Sage.
- O'Reilly, C. A., Chatman, J., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal*, 34(3), 487–516. <https://doi.org/10.5465/256404>
- Orth, U., & Luciano, E. C. (2015). Self-esteem, narcissism, and stressful life events: Testing for selection and socialization. *Journal of Personality and Social Psychology*, 109(4), 707.
- Ou, C. X., & Sia, C. L. (2010). Consumer trust and distrust: An issue of website design. *International Journal of Human-Computer Studies*, 68(12), 913–934. <https://doi.org/10.1016/J.IJHCS.2010.08.003>
- Overbeck, J. R., Tiedens, L. Z., & Brion, S. (2006). The powerful grab attention in the inattentive: The psychological presence of power. *Journal of Personality and Social Psychology*, 91(4), 557–568. <https://doi.org/10.1037/0022-3514.91.4.557>
- Paddock, E. L., Hyde, S. T., & Matheson, K. (2022). The effects of distrust in authority and workplace cynicism on organizational citizenship behavior. *Journal of Business and Psychology*, 37(4), 623–639. <https://doi.org/10.1007/s10869-021-09756-1>
- Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22–27. <https://doi.org/10.1016/J.JBEF.2017.12.004>
- Palfai, T. P., & Hart, K. E. (1997). Anger coping styles and perceived social support. *Journal of Social Psychology*, 137(4), 405–411. <https://doi.org/10.1080/00224549709595455>
- Park, C. L. (2010). Making sense of the meaning literature: An integrative review of meaning making and its effects on adjustment to stressful life events. *Psychological Bulletin*, 136(2), 257–301. <https://doi.org/10.1037/a0018301>
- Park, C. L., & Folkman, S. (1997). Meaning in the context of stress and coping. *Review of General Psychology*, 1(2), 115–144. <https://doi.org/10.1037/1089-2680.1.2.115>
- Pearsall, M. J., Ellis, A. P. J., & Stein, J. H. (2009). Coping with challenge and hindrance stressors in teams: Behavioral, cognitive, and affective outcomes. *Organizational Behavior and Human Decision Processes*, 109(1), 18–28. <https://doi.org/10.1016/j.obhdp.2009.02.002>
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163. <https://doi.org/10.1016/J.JESP.2017.01.006>

- Penley, J. A., & Tomaka, J. (2002). Associations among the Big Five, emotional responses, and coping with acute stress. *Personality and Individual Differences*, 32(7), 1215–1228. [https://doi.org/10.1016/S0191-8869\(01\)00087-3](https://doi.org/10.1016/S0191-8869(01)00087-3)
- Penney, L. M., & Spector, P. E. (2005). Job stress, incivility, and counterproductive work behavior (CWB): the moderating role of negative affectivity. *Journal of Organizational Behavior*, 26(7), 777–796. <https://doi.org/10.1002/JOB.336>
- Petersen, M. B. (2010). Distinct emotions, distinct domains: Anger, anxiety and perceptions of intentionality. *The Journal of Politics*, 72(2), 357–365. <https://doi.org/10.1017/S002238160999079X>
- Petriglieri, J. L. (2011). Under threat: Responses to and the consequences of threats to individuals' identities. *Academy of Management Review*, 36(4), 641–662. <https://doi.org/10.5465/AMR.2009.0087>
- Plutchik, R. (1962). *The emotions: Facts, theories, and a new model*. Random House.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *The Leadership Quarterly*, 1(2), 107–142. [https://doi.org/10.1016/1048-9843\(90\)90009-7](https://doi.org/10.1016/1048-9843(90)90009-7)
- Porter, C. O. L. H., Outlaw, R., Gale, J. P., & Cho, T. S. (2019). The use of online panel data in management research: A review and recommendations. *Journal of Management*, 45(1), 319–344.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31(4), 437–448. <https://doi.org/10.3102/10769986031004437>
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. In *Behavior research methods, instruments, & computers* (Vol. 36, Issue 4). www.psychonomic.org/archive/.
- Quratulain, S., Khan, A. K., Crawshaw, J. R., & Arain, G. A. (2018). Linking ethical leadership to employee burnout, workplace deviance and performance: Testing the mediating roles of trust in leader and surface acting. *Journal of Business Ethics*, 148(4), 951–967. <https://doi.org/10.1007/s10551-015-2926-3>

- Raja, U., Azeem, M. U., Haq, I. U., & Naseer, S. (2020). Perceived threat of terrorism and employee outcomes: The moderating role of negative affectivity. *Journal of Business Research*, 111, 333–343. <https://doi.org/10.1016/j.jbusres.2019.09.044>
- Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*. <https://psycnet.apa.org/record/1981-08131-001>
- Raza-Ullah, T. (2019). The power of ‘both/and’: Simultaneous trust and distrust in inter-firm cooperative alliances. *BAM2019 Conference Proceedings*. <https://www.diva-portal.org/smash/get/diva2:1355489/FULLTEXT01.pdf>
- Regorz, A. (2022). *Latent interaction in SEM using lavaan*. https://www.regorz-statistik.de/en/lavaan_sem_latent_interaction.html
- Rehbein, K., Logsdon, J. M., & Van Buren, H. J. (2013). Corporate responsiveness to shareholder activists: Considering the dialogue alternative. *Journal of Business Ethics*, 112, 137–154. <https://doi.org/10.1007/s10551-012-1240-z>
- Reimann, M., Schilke, O., & Cook, K. S. (2017). Trust is heritable, whereas distrust is not. *Proceedings of the National Academy of Sciences of the United States of America*, 114(27), 7007–7012. <https://doi.org/10.1073/pnas.1617132114>
- Risser, H. J., Morford, A. E., Fernandez, F., Moskowitz, K., Doheny, M., Yang, Y., Hersch, E., Murphy, A. N., Pinkerton, L. M., Law, C., & Lattie, E. (2024). Best practices for online and virtual data collection methods to ensure data integrity. *Translational Issues in Psychological Science*, 10(2), 150–161. <https://doi.org/10.1037/tps0000410>
- Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal*, 38(2), 555–572. <https://doi.org/10.2307/256693>
- Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6), 934–960. <https://doi.org/10.1037/0033-2909.128.6.934>
- Robinson, S. L., & Greenberg, J. (1998). Employees behaving badly: Dimensions, determinants and dilemmas in the study of workplace deviance. In C. L. Cooper & D. M. Rousseau (Eds.), *Trends in organizational behavior* (Vol. 5, pp. 1–30). John Wiley & Sons.
- Roseman, I. J. (1984). Cognitive determinants of emotions: A structural theory. In P. Shaver (Ed.), *Review of Personality and Social Psychology* (Vol. 5) (pp. 11–36). Sage.
- Rosen, C. C., Koopman, J., Gabriel, A. S., & Johnson, R. E. (2016). Who strives and who gives up? A daily investigation of curvilinear effects of conscientiousness on work engagement and goal attainment. *Journal of Applied Psychology*, 101(9), 1400–1419. <https://doi.org/10.1037/apl0000134>
- Rosenberg, M. (1979). *Conceiving the Self*. Basic Books.

- Rosenfeld, P. R., Giacalone, R. A., & Riordan, C. A. (1995). *Impression management in organizations: Theory, measurement, and practice*. Routledge.
- Rotter, J. B. (1980). Interpersonal trust, trustworthiness, and gullibility. *American Psychologist*, 35(1), 1–7. <https://doi.org/10.1037/0003-066X.35.1.1>
- Rozin, P., Lowery, L., Haidt, J., & Imada, S. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76(4), 574–586. <https://doi.org/10.1037/0022-3514.76.4.574>
- Rudorf, S., Schmelz, K., Baumgartner, T., Wiest, R., Fischbacher, U., & Knoch, D. (2018). Neural mechanisms underlying individual differences in control-averse behavior. *Journal of Neuroscience*, 38(22), 5196–5208. <https://doi.org/10.1523/JNEUROSCI.0047-18.2018>
- Rupp, D. E., & Cropanzano, R. (2002). The mediating effects of social exchange relationships in predicting workplace outcomes from multifoci organizational justice. *Organizational Behavior and Human Decision Processes*, 89(1), 925–946. [https://doi.org/10.1016/S0749-5978\(02\)00036-5](https://doi.org/10.1016/S0749-5978(02)00036-5)
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
- Sackett, P. R., & DeVore, C. J. (2001). Counterproductive behaviors at work. In N. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work and organizational psychology* (Vol. 1, pp. 145–164). Sage.
- Sagie, A., Birati, A., & Tziner, A. (2002). Assessing the costs of behavioral and psychological withdrawal: A new model and an empirical illustration. *Applied Psychology*, 51(1), 67–89. <https://doi.org/10.1111/1464-0597.00079>
- Salamon, S. D., & Robinson, S. L. (2008). Trust that binds: The impact of collective felt trust on organizational performance. *Journal of Applied Psychology*, 93(3), 593–601. <https://doi.org/10.1037/0021-9010.93.3.593>
- Sato, T. (2005). The Eysenck Personality Questionnaire brief version: Factor structure and reliability. *Journal of Psychology*, 139(6), 545–552. <https://doi.org/10.3200/JRLP.139.6.545-552>
- Saunders, M., Dietz, G., & Thornhill, A. (2014). Trust and distrust: Polar opposites, or independent but co-existing? *Human Relations*, 67(6), 639–665. <https://doi.org/10.1177/0018726713500831>
- Saunders, M., & Thornhill, A. (2004). Trust and mistrust in organizations: An exploration using an organizational justice framework. *European Journal of Work and Organizational Psychology*, 13(4), 493–515. <https://www.tandfonline.com/doi/pdf/10.1080/13594320444000182>

- Schaubroeck, J., Lam, S. S. K., & Peng, A. C. (2011). Cognition-based and affect-based trust as mediators of leader behavior influences on team performance. *Journal of Applied Psychology*, 96(4), 863–871. <https://doi.org/10.1037/a0022625>
- Schelble, B. G., Flathmann, C., Scalia, M., Zhou, S., Myers, C., Mcneese, N. J., Gorman, J., & Freeman, G. (2022). Addressing the spread of trust and distrust in distributed human-AI teaming constellations. *Workshop on Trust and Reliance in AI-Assisted Tasks, at CHI 2023*. <https://chi-trait.github.io/#/>
- Schilke, O., Reimann, M., & Cook, K. S. (2015). Power decreases trust in social exchange. *Proceedings of the National Academy of Sciences of the United States of America*, 112(42), 12950–12955. <https://doi.org/10.1073/pnas.1517057112>
- Schmader, T., & Lickel, B. (2006). The approach and avoidance function of guilt and shame emotions: Comparing reactions to self-caused and other-caused wrongdoing. *Motivation and Emotion*, 30(1), 42–55. <https://doi.org/10.1007/S11031-006-9006-0>
- Schermerhorn, J. R., Hunt, J. G., & Osborn, R. N. (2005). *Organizational behavior* (9th ed.). Wiley.
- Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative model of organizational trust: Past, present, and future. *Academy of Management Review*, 32(2), 344–354. <https://doi.org/10.5465/amr.2007.24348410>
- Schul, Y., Mayo, R., & Burnstein, E. (2008). The value of distrust. *Journal of Experimental Social Psychology*, 44(5), 1293–1302. <https://doi.org/10.1016/J.JESP.2008.05.003>
- Schwartz, S. J., Luyckx, K., & Vignoles, V. L. (Eds.). (2011). *Handbook of identity theory and research*. Springer.
- Schwarz, N., & Clore, G. L. (1996). Feelings and phenomenal experiences. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 433– 465). Guilford Press.
- Schweitzer, M. E., Ho, T. H., & Zhang, X. (2018). How monitoring influences trust: A tale of two faces. *Management Science*, 64(1), 253–270. <https://doi.org/10.1287/mnsc.2016.2586>
- Searle, R. H., Nienaber, A. I., & Sitkin, S. B. (Eds.). (2018). *The Routledge companion to trust*. Taylor & Francis Group.
- Sedikides, C., Meek, R., Alicke, M. D., & Taylor, S. (2014). Behind bars but above the bar: Prisoners consider themselves more prosocial than non-prisoners. *British Journal of Social Psychology*, 53(2), 396–403. <https://doi.org/10.1111/BJSO.12060>
- Selenko, E., Bankins, S., Shoss, M., Warburton, J., & Restubog, S. L. D. (2022). Artificial intelligence and the future of work: A functional-identity perspective. *Current Directions in Psychological Science*, 31(3), 272–279.

- Selenko, E., Berkers, H., Carter, A., Woods, S. A., Otto, K., Urbach, T., & De Witte, H. (2018). On the dynamics of work identity in atypical employment: Setting out a research agenda. *European Journal of Work and Organizational Psychology*, 27(2), 324–334. <https://doi.org/10.1080/1359432X.2018.1444605>
- Sennett, R., & Cobb, J. (1973). *The hidden injuries of class*. Random House.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (K. Prancan, Ed.). Wadsworth Cengage Learning.
- Shamir, B. (1991). Meaning, self and motivation in organizations. *Organization Studies*, 12, 405–424.
- Sharp, E. B., & Johnson, P. E. (2009). Accounting for variation in distrust of local police. *Justice Quarterly*, 26(1), 157–182. <https://doi.org/10.1080/07418820802290496>
- Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52(6), 1061–1086.
- Schneider, T. R. (2004). The role of neuroticism on psychological and physiological stress responses. *Journal of Experimental Social Psychology*, 40(6), 795–804. <https://doi.org/10.1016/J.JESP.2004.04.005>
- Strauger, J. S., & Schoeneman, T. J. (1979). Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, 86(3), 549–573.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7(4), 422–445. <https://doi.org/10.1037/1082-989X.7.4.422>
- Siebert, S., & Czarniawska, B. (2020). Distrust: Not only in secret service organizations. *Journal of Management Inquiry*, 29(3), 286–298. <https://doi.org/https://doi.org/10.1177/1056492618798939>
- Siegel, J. T., & Navarro, M. (2019). A conceptual replication examining the risk of overtly listing eligibility criteria on Amazon's Mechanical Turk. *Journal of Applied Social Psychology*, 49(4), 239–248. <https://doi.org/10.1111/jasp.12580>
- Siegel, J. T., Navarro, M. A., & Thomson, A. L. (2015). The impact of overtly listing eligibility requirements on MTurk: An investigation involving organ donation, recruitment scripts, and feelings of elevation. *Social Science and Medicine*, 142, 256–260. <https://doi.org/10.1016/j.socscimed.2015.08.020>
- Siegrist, M. (2021). Trust and risk perception: A critical review of the literature. *Risk Analysis*, 41(3), 480–490. <https://doi.org/10.1111/risa.13311>
- Sitkin, S. B., & Bijlsma-Frankema, K. M. (2018). Distrust. In R. H. Searle, A. M. I. Nienaber, & S. B. Sitkin (Eds.), *The Routledge companion to trust* (pp. 50–61).

- Sitkin, S. B., & Roth, N. L. (1993). Explaining the limited effectiveness of legalistic “remedies” for trust/distrust. *Organization Science*, 4(3), 367–392. <https://doi.org/10.1287/orsc.4.3.367>
- Sitkin, S. B., & Stickel, D. (1996). The road to hell: The dynamics of distrust in an era of quality. In R. M. Kramer & T. R. Tyler (Eds.), *Trust in organizations* (pp. 196–215). Sage Publications.
- Six, F. E., & Latusek, D. (2023). Distrust: A critical review exploring a universal distrust sequence. *Journal of Trust Research*, 13(1), 1–23. <https://doi.org/10.1080/21515581.2023.2184376>
- Skarlicki, D. P., & Folger, R. (1997). Retaliation in the workplace: The roles of distributive, procedural, and interactional justice. *Journal of Applied Psychology*, 82(3), 434–443. <https://doi.org/10.1037/0021-9010.82.3.434>
- Skarlicki, D. P., Folger, R., & Tesluk, P. (1999). Personality as a moderator in the relationship between fairness and retaliation. *Academy of Management Journal*, 42(1), 100–108. <https://doi.org/10.5465/256874>
- Skiba, T., & Wildman, J. L. (2019). Uncertainty reducer, exchange deepener, or self-determination enhancer? Feeling trust versus feeling trusted in supervisor-subordinate relationships. *Journal of Business and Psychology*, 34(2), 219–235. <https://doi.org/10.1007/s10869-018-9537-x>
- Sloan, M. M. (2012). Unfair treatment in the workplace and worker well-being: The role of coworker support in a service work environment. *Work and Occupations*, 39(1), 3–34. <https://doi.org/10.1177/0730888411406555>
- Smith, R. H., Parrott, W. G., Diener, E. F., Hoyle, R. H., & Kim, S. H. (1999). Dispositional envy. *Personality and Social Psychology Bulletin*, 25(8), 1007–1020.
- Smith, R. H., Webster, J. M., Parrot, W. G., & Eyre, H. L. (2002). The role of public exposure in moral and nonmoral shame and guilt. *Journal of Personality and Social Psychology*, 83, 138–159. <https://doi.org/10.1037/0022-3514.83.1.138>
- Spector, P. E., & Fox, S. (2005). Concluding thoughts: Where do we go from here? In S. Fox & P. E. Spector (Eds.) *Counterproductive work behavior: Investigations of actors and targets* (pp. 297–305). American Psychological Association. <https://doi.org/10.1037/10893-012>
- Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior*, 68(3), 446–460. <https://doi.org/10.1016/j.jvb.2005.10.005>
- Spielberger, C. D. (1988). *Professional manual for the State-Trait Anger Expression Inventory (STAXI) (research ed.)*. Psychological Assessment Resources.
- Spielberger, C. D., Jacobs, G., Russell, S., & Crane, R. S. (1983). Assessment of anger: The State-Trait Anger Scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment (Vol. 2)* (p. 209). Taylor & Francis Group.

- Spielberger, C. D., Krasner, S. S., & Solomon, E. P. (1988). The experience, expression and control of anger. In M. P. Janisse (Ed.), *Health psychology: Individual differences and stress* (pp. 89–108). Springer-Verlag.
- Staw, B. M., Sutton, R. I., & Pelled, L. H. (1994). Employee positive emotion and favorable outcomes at the workplace. *Organization Science*, 5(1), 51–71.
<https://doi.org/10.1287/orsc.5.1.51>
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental psychology (Vol. 21)* (pp. 261–302). Academic Press.
- Stone, J., Cooper, J., Wiegand, A. W., & Aronson, E. (1997). When exemplification fails: Hypocrisy and the motive for self-integrity. *Journal of Personality and Social Psychology*, 72(1), 54–65. <https://doi.org/10.1037/0022-3514.72.1.54>
- Stoner, J. A. F., Gilbert, D. R., & Freeman, R. E. (2013). *Management* (6th ed.). Dorling Kindersley (India) PVT limited.
- Stryker, S. (1980). *Symbolic interactionism: A social structural version*. Benjamin/Cummings
- Stryker, S. (1987). Identity theory: Developments and extensions. In *Self and identity: Psychosocial perspectives* (pp. 89–103). John Wiley & Sons.
- Stryker, S., & Burke, P. J. (2000). The past, present, and future of an identity theory. *Social Psychology Quarterly*, 63(4), 284. <https://doi.org/10.2307/2695840>
- Stryker, S., & Serpe, R. T. (1994). Identity salience and psychological centrality: Equivalent, overlapping, or complementary Concepts? *Social Psychology Quarterly*, 57(1), 16.
<https://doi.org/10.2307/2786972>
- Suh, A., Cheung, C. M., Ahuja, M., & Wagner, C. (2017). Gamification in the workplace: The central role of the aesthetic experience. *Journal of Management Information Systems*, 34(1), 268–305. <https://doi.org/10.1080/07421222.2017.1297175>
- Suls, J., Martin, R., & David, J. P. (1998). Person-environment fit and its limits: Agreeableness, neuroticism, and emotional reactivity to interpersonal conflict. *Personality and Social Psychology Bulletin*, 24(1), 88–98.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33(1), 1–39. <https://doi.org/10.1146/annurev.ps.33.020182.000245>
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7–24). Nelson-Hall.

- Tangney, J. P. (1989). *A quantitative assessment of phenomenological differences between shame and guilt. Poster presented at the 97th Annual Convention of the American Psychological Association.*
- Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology*, 61(4), 598–607.
- Tangney, J. P., & Dearing, R. L. (2002). *Shame and guilt*. Guilford Press.
- Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, 70(6), 1256–1269. <https://doi.org/10.1037/0022-3514.70.6.1256>
- Tangney, J. P., Youman, K., & Stuewig, J. (2009). Proneness to shame and proneness to guilt. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 192–209). Guilford Press.
- Tedeschi, J., & Felson, R. (1994). *Violence, aggression, and coercive action*. American Psychological Association.
- Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of Management Journal*, 43(2), 178–190. <https://doi.org/10.2307/1556375>
- Tepper, B. J., Duffy, M. K., Henle, C. A., & Lambert, L. S. (2006). Procedural injustice, victim precipitation, and abusive supervision. *Personnel Psychology*, 59(1), 101–123. <https://doi.org/10.1111/j.1744-6570.2006.00725.x>
- Tepper, B. J., Simon, L., & Park, H. M. (2017). Abusive supervision. *Annual Review of Organizational Psychology and Organizational Behavior*, 4, 123–152. <https://doi.org/10.1146/annurev-orgpsych-041015-062539>
- Thau, S., Crossley, C., Bennett, R. J., & Sczesny, S. (2007). The relationship between trust, attachment, and antisocial work behaviors. In R. Liden & R. Klimoski (Eds.), *Research in Social Issues in Management* (Vol. 6, pp. 59–84). Information Age Publishing.
- Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, 7(2), 233–256. <https://doi.org/10.1080/741944069>
- Tong, E. M. W., Bishop, G. D., Enkelmann, H. C., Why, Y. P., Diong, S. M., Ang, J., & Khader, M. (2006). The role of the Big Five in appraisals. *Personality and Individual Differences*, 41(3), 513–523. <https://doi.org/10.1016/J.PAID.2006.01.018>
- Tost, L. P., Hardin, A. E., Roberson, J. W., & Gino, F. (2022). Different roots, different fruits: Gender-based differences in cultural narratives about perceived discrimination produce divergent psychological consequences. *Academy of Management Journal*, 65(6), 1804–1834. <https://doi.org/10.5465/AMJ.2020.0339>

- Tracy, J. L., & Robins, R. W. (2004). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry*, 15(2), 103–125.
https://doi.org/10.1207/S15327965PLI1502_01
- Tracy, J. L., Robins, R. W., & Tangney, J. P. (2007). *The self-conscious emotions: Theory and research*. Guilford Press.
- Trevino, L. K., & Nelson, K. A. (2010). *Managing business ethics: Straight talk about how to do it right* (5th ed.). Wiley.
- Tripp, T. M., Bies, R. J., & Aquino, K. (2007). A vigilante model of justice: Revenge, reconciliation, forgiveness, and avoidance. *Social Justice Research*, 20(1), 10–34.
<https://doi.org/10.1007/s11211-007-0030-3>
- Tripp, T. M., & Bies, R. J. (2010). Righteous anger and revenge in the workplace: The fantasies, the feuds, the forgiveness. In M. Potegal, G. Stemmler, & C. Spielberger (Eds.), *International handbook of anger* (pp. 413–431). Springer. https://doi.org/10.1007/978-0-387-89676-2_23
- Tristram, H., Marriott, J., & Wellens, J. (2012a). Dealing with ethical issues in online research. In *What is online research?: Using the Internet for social science research*. (pp. 25–38). Bloomsbury Academic.
- Tristram, H., Marriott, J., & Wellens, J. (2012b). Online experiments. In *What is online research?: Using the Internet for social science research* (pp. 91–106). Bloomsbury Academic.
- Tristram, H., Marriott, J., & Wellens, J. (2012c). Online surveys. In *What is online research?: Using the Internet for social science research*. (pp. 39–52). Bloomsbury Academic.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.
- Turnley, W. H., & Bolino, M. C. (2001). Achieving desired images while avoiding undesired images: Exploring the role of self-monitoring in impression management. *Journal of Applied Psychology*, 86(2), 351–360. <https://doi.org/10.1037/0021-9010.86.2.351>
- Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207–232.
- Tyler, T. R. (2001). The psychology of public dissatisfaction with government. In J. R. Hibbing & E. Theiss-Morse (Eds.), *What is it about government that Americans dislike?* (pp. 227–242). Cambridge University Press.
- Tyler, T. R. (2006). *Why people obey the law* (2nd ed.). Princeton University Press.
- Tyler, T. R., & Lind, E. A. (1992). A relational model of authority in groups. *Advances in Experimental Social Psychology*, 25, 115–191.

- Umphress, E. E., Bingham, J. B., & Mitchell, M. S. (2010). Unethical behavior in the name of the company: The moderating effect of organizational identification and positive reciprocity beliefs on unethical pro-organizational behavior. *Journal of Applied Psychology*, 95(4), 769–780. <https://doi.org/10.1037/a0019214>
- Umphress, E. E., & Bingham, J. B. (2011). When employees do bad things for good reasons: Examining unethical pro-organizational behaviors. *Organization Science*, 22(3), 621–640. <https://doi.org/10.1287/orsc.1100.0559>
- van der Ploeg, H. M., Defares, P. B., & Spielberger, C. D. (1982). *Manual of the Dutch State-Trait Anger Scale Self-Analysis Questionnaire*. Swets & Zeitlinger.
- Van De Walle, S., & Six, F. (2014). Trust and distrust as distinct concepts: Why studying distrust in institutions is important. *Journal of Comparative Policy Analysis: Research and Practice*, 16(2), 158–174. <https://doi.org/10.1080/13876988.2013.785146>
- van Dijke, M., De Cremer, D., Brebels, L., & van Quaquebeke, N. (2015). Willingness to justify past organizational misdeeds: The role of Machiavellianism and organizational identification. *Human Relations*, 68(5), 765–789. <https://doi.org/10.1177/0018726714543734>
- van Dijke, M., Wildschut, T., Leunissen, J. M., & Sedikides, C. (2015). Nostalgia buffers the negative impact of low procedural justice on cooperation. *Organizational Behavior and Human Decision Processes*, 127, 15–25. <https://doi.org/10.1016/j.obhdp.2014.11.005>
- Van Dijk, R., & van Dick, R. (2009). Navigating organizational change: Change leaders, employee resistance and work-based identities. *Journal of Change Management*, 9(2), 143–163.
- Van Dyck, C., Frese, M., Baer, M., & Sonnentag, S. (2005). Organizational error management culture and its impact on performance: A two-study replication. *Journal of Applied Psychology*, 90(6), 1228–1240. <https://doi.org/10.1037/0021-9010.90.6.1228>
- van Prooijen, J. W., & Ellemers, N. (2015). Does it take intuition to behave fairly? Fairness judgments and the role of affect and intuitions. *Social Justice Research*, 28, 173–193. <https://doi.org/10.1007/s11211-015-0233-0>
- Vecchio, R. P., Justin, J. E., & Pearce, C. L. (2010). Empowering leadership: An examination of mediating mechanisms within a hierarchical structure. *The Leadership Quarterly*, 21(3), 530–542. <https://doi.org/10.1016/j.leaqua.2010.03.014>
- Verbeke, W., Dietz, B., & Verwaal, E. (2011). Drivers of sales performance: A contemporary meta-analysis. Have salespeople become knowledge brokers? *Journal of the Academy of Marketing Science*, 39(3), 407–428. <https://doi.org/10.1007/s11747-010-0211-8>
- Vidourek, R. A., & Burbage, M. L. (2019). Workplace bullying among U.S. adults: Causes, controls and prevention. *Journal of Social Sciences Research*, 5(2), 338–343. <https://doi.org/10.32861/jssr.52.338.343>

- Wang, Q., Bowling, N. A., & Tian, Q. (2013). How does task conflict promote counterproductive work behavior? The moderating role of job autonomy. *Journal of Business and Psychology*, 28, 1–13. <https://doi.org/10.1007/s10869-012-9265-6>
- Wang, X. H. F., Kim, T. Y., & Lee, D. R. (2016). Cognitive diversity and team creativity: Effects of team intrinsic motivation and transformational leadership. *Journal of Business Research*, 69(9), 3231–3239. <https://doi.org/10.1016/j.jbusres.2016.02.026>
- Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the positive and negative affect schedule-expanded form*. University of Iowa.
- Webb Hooper, M., Mitchell, C., Marshall, V. J., Cheatham, C., Austin, K., Sanders, K., & Grafton, L. L. (2022). Responding to healthcare distrust among underserved communities: Phase II. In *Psycho-Oncology* (Vol. 31, Issue 1, pp. 3–8). John Wiley and Sons Ltd. <https://doi.org/10.1002/pon.5841>
- Weber, M. (1947). *The theory of social and economic organization* (A. M. Henderson & T. Parsons, Trans.). Free Press.
- Weinberg, J. (2023). Feelings of trust, distrust and risky decision-making in political office. An experimental study with national politicians in three democracies. *Comparative Political Studies*, 56(7), 935–967. <https://doi.org/10.1177/00104140221139376>
- Weiss, H. M., & Cropanzano, R. (1996). Affective events theory. *Research in Organizational Behavior*, 18(1), 1–74.
- Weiss, J. W. (2021). *Business ethics: A stakeholder and issues management approach* (6th ed.). Berrett-Koehler Publishers.
- Wellman, N., Mayer, D. M., Ong, M., & DeRue, D. S. (2016). When are do-gooders treated badly? Legitimate power, role expectations, and reactions to moral objection in organizations. *Journal of Applied Psychology*, 101(6), 793–814. <https://doi.org/10.1037/APL0000094>
- Wellman, N., Tröster, C., Grimes, M., Roberson, Q., Rink, F., & Gruber, M. (2023). Publishing multimethod research in AMJ: A review and best-practice recommendations. *Academy of Management Journal*, 66(4), 1007–1015. <https://doi.org/10.5465/amj.2023.4004>
- Welten, S. C. M. (2011). *Concerning shame*. Unpublished Doctoral Thesis. Tilburg University, the Netherlands.
- Wessling, K. S., Huber, J., & Netzer, O. (2017). MTurk character misrepresentation: Assessment and solutions. *Journal of Consumer Research*, 44(1), 211–230. <https://doi.org/10.1093/jcr/ucx053>
- Whitener, E. M., Brodt, S. E., Korsgaard, M. A., & Werner, J. M. (1998). Managers as initiators of trust: An exchange relationship framework for understanding managerial trustworthy behavior. *Academy of Management Review*, 23(3), 513–530. <https://doi.org/10.5465/amr.1998.926624>

- Wicker, F. W., Payne, G. C., & Morgan, R. D. (1983). Participant descriptions of guilt and shame. *Motivation and Emotion*, 7(1), 25–39. <https://doi.org/10.1007/BF00992963>
- Williams, K. D. (2007). Ostracism. *Annual Review of Psychology*, 58, 425–452. <https://doi.org/10.1146/annurev.psych.58.110405.085641>
- Williams, K. D., & Nida, S. A. (2011). Ostracism: Consequences and coping. *Current Directions in Psychological Science*, 20(2), 71–75. <https://doi.org/10.1177/0963721411402480>
- Williamson, O. E. (1996). Economic organization: The case for candor. In *Source: The Academy of Management Review* (Vol. 21, Issue 1). <https://about.jstor.org/terms>
- Wilson, N. L. (2019). Employees' behavioural responses to workplace mistreatment and person and workplace factors influencing their responses [*Doctoral dissertation, University of Alberta*].
- Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, 14(3), 131–134. <https://doi.org/10.1111/j.0963-7214.2005.00355.x>
- Wiltermuth, S. S., & Flynn, F. J. (2013). Power, moral clarity, and punishment in the workplace. *Academy of Management Journal*, 56(4), 1002–1023. <https://doi.org/10.546>
- Woodyatt, L., & Wenzel, M. (2014). A needs-based perspective on self-forgiveness: Addressing threat to moral identity as a means of encouraging interpersonal and intrapersonal restoration. *Journal of Experimental Social Psychology*, 50(1), 125–135. <https://doi.org/10.1016/J.JESP.2013.09.012>
- Yan, T. Ta., Tangirala, S., Vadera, A. K., & Ekkirala, S. (2022). How employees learn to speak up from their leaders: Gender congruity effects in the development of voice self-efficacy. *Journal of Applied Psychology*, 107(4), 650–667. <https://doi.org/10.1037/apl0000892>
- Yukl, G., Mahsud, R., Hassan, S., & Prussia, G. E. (2013). An improved measure of ethical leadership. *Journal of Leadership & Organizational Studies*, 20(1), 38–48. <https://doi.org/10.1177/1548051811429352>
- Zapf, D. (2002). Emotion work and psychological well-being: A review of the literature and some conceptual considerations. *Human Resource Management Review*, 12(2), 237–268. [https://doi.org/10.1016/S1053-4822\(02\)00048-7](https://doi.org/10.1016/S1053-4822(02)00048-7)
- Zeelenberg, M., Nelissen, R.M.A. Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Judgment and Decision Making*, 3, 18–27.
- Zeelenberg, M., & Pieters, R. (2006). Feeling is for doing: A pragmatic approach to the study of emotions in economic behavior. In D. De Cremer, M. Zeelenberg, & K. Murnighan (Eds.), *Social psychology and economics* (pp. 117–137). Lawrence Erlbaum Associates.

- Zhao, H. A., Xia, Q., He, P., Sheard, G., & Wan, P. (2019). Workplace ostracism and discretionary work behavior: A self-determination perspective. *Journal of Business Ethics*, 158, 1083–1104. <https://doi.org/10.1007/s10551-017-3721-y>
- Zhang, R., Liu, Y., & Wang, X. (2024). A trusted employee who distrusts: Exploring the joint effect of being trusted and distrust on employees' performance. *Current Psychology*, 43(17), 14960–14974. <https://doi.org/10.1007/s12144-022-03622-w>
- Zheng, X., Hall, R. J., & Schyns, B. (2019). Investigating follower felt trust from a social cognitive perspective. *European Journal of Work and Organizational Psychology*, 28(6), 873–885. <https://doi.org/10.1080/1359432X.2019.1678588>
- Zinner, L. R., Brodish, A. B., Devine, P. G., & Harmon-Jones, E. (2008). Anger and asymmetrical frontal cortical activity: Evidence for an anger withdrawal relationship. *Cognition and Emotion*, 22(6), 1081–1093. <https://doi.org/10.1080/02699930701622961>
- Zuckerman, E. (2021). *Mistrust: Why losing faith in institutions provides the tools to transform them* (1st ed.). W. W. Norton & Company, Inc.
- Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods*, 6(1), 21–29. <https://doi.org/10.22237/jmasm/1177992180>
- Zumbo, B. D., & Kroc, E. (2019). A measurement is a choice and Stevens' scales of measurement do not help make it: A response to Chalmers. *Educational and Psychological Measurement*, 79(6), 1184–1197. <https://doi.org/10.1177/0013164419844305>

Appendices

Appendix A: Study 1 Experimental Vignettes

INSTRUCTIONS (the same for both scenarios):

As you read the scenario, imagine yourself as the employee in this situation. After you read the scenario, you will be asked several questions about how this situation makes you feel.

NEUTRAL CONDITION

You are an employee at company X and you enjoy working there. Part of your job responsibilities includes working on a project that is very important to your supervisor. To complete some of the tasks, your team needs access to highly sensitive information. Your supervisor explains to everyone working on the project:

“Password-protected access to sensitive information is a standard policy for all employees. It guarantees secure and responsible treatment of high-stake company data. You all have been pre-approved for access to the system since you’ve been assigned to this project. The system is set up so that the password changes daily as a function of the software. So, whenever you need access to the data, please check in with the IT specialist to get a password.”

DISTRUST CONDITION

You are an employee at company X and you enjoy working there. Part of your job responsibilities include working on a project that is very important to your supervisor. To complete some of the tasks, your team needs access to highly sensitive information. Your supervisor explains to everyone working on the project:

“Password-protected access to sensitive information is a standard policy for all employees. It guarantees secure and responsible treatment of high-stake company data. You all have been pre-approved for access to the system since you’ve been assigned to this project. The system is set up so that the password changes daily as a function of the software. So, whenever you need access to the data, please check in with the IT specialist to get a password.”

One day, as you make your usual request for a password to access project-related sensitive data, you are asked to provide a written note to your supervisor explaining the intended use of this information. You try to approach the supervisor about it, but they avoid direct contact with you. Moreover, you notice that they look worried and uneasy when you join in informal group conversations. At lunch, a colleague confidentially shares with you that your supervisor asked some personal information about you. You decide to ask your supervisor about their concerns, and they respond:

“Some of your recent decisions have been inconsistent with my vision of how the project should evolve. Your actions were also incongruent with your own words on a few occasions. All of this made me question your values and raised my concerns about your intentions to take advantage of the sensitive information in a way that may hurt the project. I can’t have this project harmed in any way; my reputation depends on it. That’s why I want to have a paper record of how you intend to use the sensitive information every time you access the system.”

Appendix B: Results of Confirmatory Factor Analyses

Table B1. Confirmatory Factor Analyses of the core variables

Core variables and factors	Items	Standardized loadings	CR	AVE	Cronbach's alpha	Omega 3 CR
Interpersonal felt distrust (manipulation check)	Manipch_1	0.950***	0.97	0.9	0.97	0.97
	Manipch_2	0.898***				
	Manipch_3	0.972***				
	Manipch_4	0.974***				
Identity threat	Idthreat_1	0.940***	0.97	0.8	0.97	0.96
	Idthreat_2	0.905***				
	Idthreat_3	0.928***				
	Idthreat_4	0.878***				
	Idthreat_5	0.894***				
	Idthreat_6	0.796***				
	Idthreat_7	0.889***				
	Idthreat_8	0.906***				
Moral identity centrality	MIC_1	0.650***	0.78	0.4	0.76	0.75
	MIC_2	0.765***				
	MIC_3	0.493***				
	MIC_4	0.504***				
	MIC_5	0.803***				
Anger-experience	Angerexp_1	0.897***	0.95	0.77	0.95	0.95
	Angerexp_2	0.844***				
	Angerexp_3	0.890***				
	Angerexp_4	0.907***				
	Angerexp_5	0.855***				
	Angerexp_6	0.876***				
Shame-experience	Shameexp_1	0.895***	0.96	0.81	0.95	0.96
	Shameexp_2	0.930***				
	Shameexp_3	0.840***				
	Shameexp_4	0.912***				
	Shameexp_5	0.918***				

Note: ***p < 0.001

Table B2: Confirmatory Factor Analyses of control variables

Control variables and factors	Items	Standardized loadings	CR	AVE	Cronbach's alpha	Omega 3 CR
Shame-trait	Shametrain_1	0.659***	0.73	0.41	0.73	0.73
	Shametrain_2	0.599***				
	Shametrain_3	0.653***				
	Shametrain_4	0.645***				
Anger-trait	Angertrait_1	0.746***	0.85	0.39	0.86	0.73
	Angertrait_2	0.757***				
	Angertrait_3	0.797***				
	Angertrait_4	0.700***				
	Angertrait_5	0.452***				
	Angertrait_6	0.428***				
	Angertrait_7	0.265***				
	Angertrait_8	0.448***				
	Angertrait_9	0.771***				
	Angertrait_10	0.599***				
Neuroticism	Neuroticism_1	0.673***	0.94	0.39	0.94	0.93
	Neuroticism_2	0.677***				
	Neuroticism_3	0.666***				
	Neuroticism_4	0.705***				
	Neuroticism_5	0.688***				
	Neuroticism_6	0.838***				
	Neuroticism_7	0.824***				
	Neuroticism_8	0.786***				
	Neuroticism_9	0.764***				
	Neuroticism_10	0.848***				
	Neuroticism_11	0.657***				
	Neuroticism_12	0.735***				
Narcissism	Narcissism_1	0.964***	0.89	0.77	0.78	0.86
	Narcissism_2	0.741***				
	Narcissism_3	0.576***				
	Narcissism_4	0.245***				
	Narcissism_5	0.680***				
	Narcissism_6	0.403***				
	Narcissism_7	0.491***				
	Narcissism_8	0.384***				
	Narcissism_9	0.980***				
	Narcissism_10	0.709***				
	Narcissism_11	0.504***				
	Narcissism_12	0.452***				
	Narcissism_13	0.693***				
	Narcissism_14	0.638***				
	Narcissism_15	0.468***				
	Narcissism_16	0.523***				

Note: *** $p < 0.001$. Narcissism is a latent variable with categorical indicators. For constructs with categorical indicators, Ordinal Alpha (Zumbo et al., 2007; Zumbo & Kroc, 2019) is calculated, which treats ordinal variables as numeric. For Narcissism, $\alpha_{ord} = 0.88$.

Likewise, for Narcissism, AVE is calculated from polychoric (polyserial) not Pearson correlations.

Appendix C: Report on Hypotheses Testing

Table C1. List of Hypotheses and Report on Their Testing.

Hypotheses	Study 1 Chapter 4	Study 2 Chapter 5
H1: Felt distrust leads to work-related identity threat.	(Felt Distrust operationalization: Scenario) Supported	(Felt Distrust operationalization: CIT) Supported
H2: Moral identity centrality strengthens the relationship between felt distrust and work-related identity threat.	Supported	Supported
H3: Felt distrust leads to anger, mediated by work-related identity threat.	Supported	Not supported
H4: Felt distrust leads to shame, mediated by work-related identity threat.	Supported	Supported
H5: Anger mediates the positive relationship between work-related identity threat and counterproductive behavior.	N/A	Not supported
H6: Anger mediates the positive relationship between work-related identity threat and psychological withdrawal behavior.	N/A	Not supported
H7: Shame mediates the positive relationship between work-related identity threat and psychological withdrawal behavior.	N/A	Not supported
H8: Shame mediates the positive relationship between work-related identity threat and exemplification behavior.	N/A	Not supported

Appendix D: Study 2 Results of Confirmatory Factor Analyses

Table D1. Study 2. Confirmatory Factor Analyses of the core variables

Core variables and factors	Items	Standardi zed loadings	CR	AVE	Cronbach' s alpha	Omega 3 CR
Interpersonal Felt Distrust	IFDs1	0.485***	0.49	0.55	0.78	0.81
	IFDs2	0.707***				
	IFDs3	0.888***				
	IFDs4	0.659***				
Work-Related Identity Threat	IDT1	0.679***	0.37	0.42	0.85	0.81
	IDT2	0.696***				
	IDT3	0.747***				
	IDT4	0.445***				
	IDT5	0.412***				
	IDT6	0.576***				
	IDT7	0.806***				
	IDT8	0.683***				
Moral Identity Centrality	MIC1	0.795***	0.43	0.42	0.76	0.77
	MIC2	0.689***				
	MIC3	0.379***				
	MIC4	0.628***				
	MIC5	0.704***				
Anger-Experience	AngExp1	0.615***	0.46	0.48	0.83	0.84
	AngExp2	0.630***				
	AngExp3	0.583***				
	AngExp4	0.720***				
	AngExp5	0.719***				
	AngExp6	0.788***				
Shame-Experience	ShamExp1	0.728***	0.60	0.59	0.88	0.88
	ShamExp2	0.876***				
	ShamExp3	0.665***				
	ShamExp4	0.790***				
	ShamExp5	0.806***				
Counterproductive Work Behavior (CWB)	CWB1	0.631***	0.39	0.37	0.82	0.80
	CWB2	0.681***				
	CWB3	0.673***				
	CWB4	0.595***				
	CWB5	0.516***				
	CWB6	0.608***				
	CWB7	0.650***				
	CWB8	0.632***				
Psychological Withdrawal (PWB)	PWB1	0.601***	0.44	0.44	0.86	0.85
	PWB2	0.557***				
	PWB3	0.769***				
	PWB4	0.697***				
	PWB5	0.797***				
	PWB6	0.753***				
	PWB7	0.500***				
	PWB8	0.551***				
Exemplification	Exemp1	0.815***	0.58	0.59	0.83	0.84
	Exemp2	0.434***				
	Exemp3	0.834***				
	Exemp4	0.885***				

Note: ***p < 0.001

Table D2. Study 2. Confirmatory Factor Analyses of the control variables

Control variables and factors	Items	Standardized loadings	CR	AVE	Cronbach's alpha	Omega 3 CR
Shame-Trait	ShamTr1	0.738***	0.48	0.48	0.78	0.78
	ShamTr2	0.702***				
	ShamTr3	0.674***				
	ShamTr4	0.649***				
Anger-Trait	AngTr1	0.778***	0.31	0.35	0.83	0.71
	AngTr2	0.823***				
	AngTr3	0.757***				
	AngTr4	0.736***				
	AngTr5	0.419***				
	AngTr6	0.428***				
	AngTr7	0.231***				
	AngTr8	0.398***				
	AngTr9	0.490***				
	AngTr10	0.487***				
Neuroticism	Neuro1	0.732***	0.52	0.52	0.93	0.93
	Neuro2	0.743***				
	Neuro3	0.673***				
	Neuro4	0.723***				
	Neuro5	0.691***				
	Neuro6	0.762***				
	Neuro7	0.755***				
	Neuro8	0.761***				
	Neuro9	0.684***				
	Neuro10	0.736***				
	Neuro11	0.670***				
	Neuro12	0.683***				
Narcissism	Narci_key1	0.949***	0.37	0.38	0.78	0.83
	Narci_key2	0.709***				
	Narci_key3	0.656***				
	Narci_key4	0.134***				
	Narci_key5	0.638***				
	Narci_key6	0.320***				
	Narci_key7	0.468***				
	Narci_key8	0.396***				
	Narci_key9	0.973***				
	Narci_key10	0.588***				
	Narci_key11	0.553***				
	Narci_key12	0.499***				
	Narci_key13	0.596***				
	Narci_key14	0.641***				
	Narci_key15	0.660***				
	Narci_key16	0.501***				

Note: *** $p < 0.001$. Narcissism is a latent variable with categorical indicators. For constructs with categorical indicators, Ordinal Alpha (Zumbo et al., 2007; Zumbo & Kroc, 2019) is calculated, which treats ordinal variables as numeric. For Narcissism, $\alpha_{\text{ord}} = 0.88$.

Likewise, for Narcissism, AVE is calculated from polychoric (polyserial) not Pearson correlations.