Followership in an adversarial context: Uncovering resistance dynamics

Maria Teresa Silva Féria de Almeida

PhD in Management, specialization of Human Resources and Organizational Behavior

Supervisors:

PhD Nelson Campos Ramalho, Associate Professor,
Iscte - Instituto Universitário de Lisboa

PhD Francisco Esteves, Full Professor,
Mid Sweden University
Universidade Católica Portuguesa

December, 2021

iscte

BUSINESS SCHOOL

Department of Human Resources and Organizational Behavior

Followership in an adversarial context: Uncovering resistance dynamics

Maria Teresa Silva Féria de Almeida

PhD in Management, specialization of Human Resources and Organizational Behavior

Jury:

PhD Sílvia Silva, Full Professor, Iscte - Instituto Universitário de Lisboa (President)
PhD Dennis Tourish, Full Professor, University of Sussex Business School
PhD Pedro Neves, Full Professor, Nova School of Business and Economics
PhD Aristides Ferreira, Associate Professor, Iscte - Instituto Universitário de Lisboa
PhD Nelson Campos Ramalho, Associate Professor, Iscte - Instituto Universitário de Lisboa

December, 2021

Acknowledgments

This document is but a collection of characters assembled into words that are sewn into sentences enclosed into paragraphs compiled in chapters meant to make sense. The spaces between comprehend everything else.

To Professor Nelson, a *thank you* that will not be forgotten; without any sense of duty, but with the admiration that legitimizes respect.

To Professor Francisco, I thank the generosity for participating in this project and every discussion we had.

To Iscte's Human Resources and Organizational Behavior department Professors. To Professor Sara Ramos for enabling some data collection. To Professor José Neves, the availability and learning opportunity in my first teaching experience. To Professor Silvia Silva, for the chance to share our research with students. To Professor Ana Passos for returning feedback each year on the evolution of the project.

To those who are mine - family, and friends. To my Mother, Father, and Sister, the bond and freedom. To my Grandmothers, each lesson. To Adriana, Joana, Filipa, Helena, Mariana, Clara, and Mariana, for being there.

To Catarina, Inês, Rúben, and Feng, all of whom I met as colleagues and are now friends.

Finally, to all who ask me "what's that!?" every time I talk about followership. Each of these moments are pieces of evidence of the need to study followership.

Thank you.

FCT Fundação para a Ciência e a Tecnologia   **bru** _ iscte
Business Research Unit

Resumo

O fenómeno de *followership* foi negligenciado devido à premissa de que os *followers* são produto da liderança. Com o tempo, a sua posição na equação mudou e passam a ser vistos como agentes explicativos do processo. Esta tese inscreve-se nesta abordagem, defendendo que a compreensão da natureza do *followership* emerge da análise da resistência/passividade a um líder destrutivo. O primeiro estudo analisa a natureza conceptual de *followership* e propõe uma classificação de tipos de *followers*. Este estudo permite compreender a amplitude do comportamento dos *followers*, propondo a resistência como expressão de *followership*. O segundo estudo analisa o impacto das crenças dos *followers* sobre o seu papel nos processos de resistência a diferentes tipos de líderes destrutivos. Os resultados sugerem que a forma como os *followers* concebem o seu papel é importante na explicação da resistência. O terceiro estudo examina a natureza dos comportamentos de resistência e propõe uma reinterpretação de uma escala existente, em que um dos fatores divide-se em dois. A pertinência da proposta foi testada num modelo que sugere diferenciação nomológica. O último estudo aponta para os novos desafios aos processos de liderança/*followership*, em que os *followers* são chamados a interagir com chefias sintéticas (i.e., inteligência artificial). É analisado um processo explicativo de comportamentos de obediência e os resultados sugerem que a teoria processual existente não se aplica. Globalmente, conclui-se que estudar *followership* num contexto adverso amplia o conhecimento sobre a natureza do fenómeno e sugere que os *followers* são agentes protetores nestes contextos adversariais.

Palavras-chave: follower(ship), lider(ança) destrutivo(a), coprodução, resistência, chefia sintética

Classificação JEL: D23 Comportamento Organizacional, M12 Gestão de Pessoas

Abstract

The phenomenon of followership has been neglected due to the premise that followers are a product of leadership. However, over time, followers' position in the equation has changed, and they are now conceived as explanatory agents of the process. This thesis falls within this approach, arguing that the understanding of followership's nature emerges from the analysis of resistance/passivity to a destructive leader. The first study analyzes the conceptual nature of followership and proposes a classification of follower types. This study contributes to expand the range of followers' behaviors by proposing resistance as an expression of followership. The second study analyzes the impact of followers' role beliefs in resistance processes to different destructive leaders. The results suggest that the way followers conceive their role is important in explaining resistance. The third study examines the nature of resistance behaviors and proposes a reinterpretation of an existing scale, in which one of the factors splits into two. The proposal's relevance was tested in a model that suggests nomological differentiation. The last study points out new challenges to leadership/followership processes, where followers are called upon to interact with synthetic (i.e., artificial intelligence based) supervisors. A process explaining obedience behaviors is analyzed, and the results suggest that extant process theory does not apply. Overall, studying followership in an adversarial context broadens knowledge about the nature of the phenomenon and suggests followers are protective agents in such adversarial settings.

Keywords: follower(ship), destructive leader(ship), coproduction, resistance, synthetic supervisor

JEL Classification: D23 Organizational Behavior; M12 Personnel Management

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction[1]

The best way to introduce this topic is to recall Ronald Riggio's (2020) question "Why followership?". An immediate and possible answer is that there are no leaders without followers (Hollander, 1993). However, followership's relevance goes far beyond that. As it will be developed throughout this chapter, followership is a natural social phenomenon that has always been there; it helps to balance the expression of leadership; and it is required once we conceive leadership as a process.

1.1.   Followership has always been there

Broadly taken, followership refers to the impact of followers in the leadership process (Uhl-Bien et al., 2014). This simple claim brings out two major ideas: a) leadership is not a person; it is a process; and b) followers take part in that process. Although the conceptual debate on followership, its nature, and impact has been gaining attention ever since Robert Kelley's seminal work in 1988, it can be considered a recent topic compared to the long history of leadership research (Baker, 2007; Carsten et al., 2014). It is surprising then that something required for a process to occur has been overlooked for so long. In Uhl-Bien and Carsten (2018, p. 211) words "(…) that does not mean followership is not there, or not foundational. It just means that we are not using a lens that sees it.".

And, in fact, followership has always been around us. This is clearly supported if we look beyond any conceptual debate and see followership as a social phenomenon. According to this idea, a group, under certain conditions, accepts the influence of another agent - i.e., a leader - (Bastardoz & Van Vugt, 2019), that is considered the best strategy to pursue collective goals. So, accepting to follow a leader has emerged as an evolutionary adaptive strategy (Kellerman, 2012; Van Vugt, 2006) which is supported by studies showing these dynamics across species (Bastardoz & Van Vugt, 2019; Smith, 2017).

---

[1] Part of this chapter is submitted for publication as a chapter in an international Handbook of Leadership.

1.2.   Follower(ship) balances leader(ship)

As an adaptive process, followership carries a critical purpose: regulating the leader's behavior to ensure that he or she works to accomplish the group's goals (Bastardoz & Van Vugt, 2019). An extreme example of the group working together to ensure they only keep the leader as long as it is the right agent occurs in a beehive. According to Vollet-Neto et al. (2017) when the queen bee no longer serves the comb's interests because the reproduction became defective, she is killed by the group and replaced by another queen.

We should then acknowledge that followers can develop tactics or strategies to influence those at formal power positions. This means that the asymmetric power relation assumed in leadership (Collinson, 2005) can be rebalanced once followers engage in counterpower behaviors (Yukl & Gardner, 2020).

In contrast to other contexts, such as those with animals where natural leadership processes emerge, human leadership does not necessarily stem from the group processes or sanctioning. In human leadership, even when the leaders are indicated by the group, it is possible that such leaders, at a given time, no longer act in line with the group's interests (Burke, 2006). Still, the group might not be able, unlike bees, to dethrone such a leader.

So, it is essential to develop mechanisms that allow to control and express discontentment when the leader does not perform well. This idea is not new, and many examples of the importance of curbing the leaders' behaviors have been fully described, such as when Winston Churchill's wife rescued him when he started to derail by confronting him with truths that others would not dare, or the role of a trusted aid such as the one held by Eric Schmidt in relation to former Google's CEO, Larry Page (Hyde, 2018). This idea restores a very old, universal, and institutional mechanism aimed to prevent destructive behaviors by those in power: the jester (Otto, 2007). Entertaining the court was but a means to fulfill a much more important purpose: to mirror the truth, in the form of satire, about the behavior of those in power and their surroundings. These are but some examples that show the relevance of not leading alone. When we think about these dynamics at work, shouldn´t followership be in charge of this?

1.3.   Followership and/or leadership?

Neither leadership is leader-exclusive, nor followership is follower-limited. This is conceptually supported by a constructionist approach, and it can be observed when followers engage in some leading behaviors (Riggio, 2014). To engage in leading behaviors means that

at some point, the imbalanced power relationship between leaders and followers (Collinson, 2005) is reversed, and followers start to influence leaders.

So, from a power-based process view, although there is an imbalanced relationship, the apparently powerless individual always has the chance to decide about his or her own behavior, create counterpower, and steer the system using uncertainty zones (Crozier & Friedberg, 1980). Accordingly, although the formal role position may be already defined, there is room for both downward and upward influence (Oc & Bashshur, 2013; Yukl & Tracey, 1992). Leadership is thus a play where both leaders and followers can act. However, acting is not a sufficient condition to state that followers are trying to build leadership. Creating leadership is not merely a behavior; it needs to have a specific reason: help reaching a common purpose (that is expected to align with the organizational good). This idea roots back to early 20th Century with Mary Parker Follett's (1949) idea of following the invisible leader - i.e., the common purpose. Conversely, if one tries to balance the leader-follower relationship based on ambition, it should not be considered as a good followership practice. The same happens when someone engages in behaviors such as retaliation, but still, it remains followership in its essence.

Thus, if motivations are necessary to judge on followership, one should first focus on the underlying cognitive mechanisms that help understand the followers' role orientation. Co-production of leadership beliefs help to predict how followers behave in the leadership process as they show how followers view their role (Carsten & Uhl-Bien, 2012, p. 211): "we define followers' co-production beliefs as the extent to which individuals believe the follower role involves partnering with leaders to advance the mission and achieve optimal levels of productivity". Then, it is also important to look at behaviors that can make a difference (i.e., constructive resistance).

From what has been described, it is clear that we must acknowledge followers as active agents for followership to occur. So, leaders and followers are expected to work together for the collective good (Riggio, 2014). As it will be described later in this work, although not the pioneer (Baker, 2007), it was Kelley (1988) who gave the stage to followers as active agents, and research on followership has established itself as a main perspective on its own, and is gaining momentum. However, it has not always been like this, due to two main and related reasons that might have helped to slow the evolution of this field: 1) the word follower holds a negative connotation (Kelley, 2008; Riggio, 2020); and 2) followership emerges from leadership studies, that have been traditionally leader-centric (Baker et al., 2014; Oc & Bashshur, 2013). This means that, for a long time, mainstream leadership studies treated followers as mere recipients (Shamir, 2007; Thoroughgood et al., 2018), which, in turn, helped

to increase the negative connotation of the word follower. According to this idea, to follow is meaningless, powerless, and no one aims to fulfill such a role (Hoption et al., 2012).

This is in line with leader-centric approaches, where leader's characteristics are overvalued, and they are seen as the major influencers of the followers' attitudes, behaviors, and performance (Shamir, 2007). This idea is well-illustrated by trait-related approaches (Jago, 1982; Kirkpatick & Locke, 1991) and it is not difficult to understand its origin if we consider that for many years the survival of the species may have been associated with the idea of physical traits (Riggio & Riggio, 2010). As mentioned, followers were mere recipients, even when some approaches started to consider that the leader's behavior should fit the type of follower (e.g., maturity level), like in contingency models (Hersey et al., 1979).

Although trait-based theories still have some impact and relevance (House & Howell, 1992; Zaccaro et al., 2018), as needs changed, it was possible to realize that survival/success was not only linked to traits. Moreover, it then became essential to understand what was valued and how followers viewed leaders. Thus, implicit theories of leadership emerged. Within such theories, followers are still not the center, but they are considered relevant in the process for the first time. The most remarkable example of this approach is the Romance of Leadership (Meindl, 1995; Meindl et al., 1985), which paved the way for the follower-centric perspectives. If this helps us explain, for example, the attribution of charisma (Bligh & Schyns, 2007), that is related to the mobilization of masses (Grabo et al., 2017), it does not allow us to analyze the active role of followers in the leadership process. According to Uhl-Bien et al. (2014) we can then describe the relational approach, according to which followers are part of the mutual-relational leadership process, but their role is not emphasized. The Leader-Member-Exchange Theory (Dansereau et al., 1975; Gerstner & Day, 1997) illustrates this approach. The definitive focus on followers comes with followership, which recognizes (emphasizes) the importance of these elements in the leadership process. It can be seen according to two lenses: the position one holds (role) and/or as a social process (Uhl-Bien et al., 2014).

So, followership emerged as research advanced, and academics started to see followers as an heterogenous group where both active and passive followers can co-exist (e.g., Collinson, 2006; Kelley, 1988). Some of these classifications were developed under the role-based approach and started to "reverse the lens" (Shamir, 2007) in the leadership studies. This approach draws our attention to the importance of followers as differentiated agents that can be moved by different purposes (e.g., their implicit theories, level of engagement with the leader). This idea of behavioral drivers in followership has been discussed by critical leadership studies addressing the nature of followership (Collinson, 2017; Learmonth & Morell, 2017; Blom &

Lundgren, 2020) and is linked to another critical issue on the topic: follower's agency or volition. Once this volition is linked to the purpose of taking a leading behavior (as an expression of the volition to follow the common good), a constructionist approach to leadership needs to be considered. In such an approach, leadership emerges from the interaction between the leader and followers (Uhl-Bien et al., 2014). It is then possible to argue that followership theory is necessary to better understand leadership. More than that, these "are two sides of the same construct" (Uhl-Bien & Carsten, 2018, p. 197). In other words, we started from a leader-centric approach to a perspective where both leaders and followers are central players in the construction of the leadership process. So, leadership or followership? Isn't it just a matter of perspective?

1.4. Why (do we need) followership?

As described, given its adaptive nature, followership (through followers' behaviors) should curb leader's behaviors when they are not aligned with or when they threaten the collective good. Those behaviors can be described as destructive once the leader fails to behave ethically or in an effective manner (Kellerman, 2012). There are many classifications of destructive behaviors, such as abusive supervision (Tepper, 2000), petty tyranny (Ashforth, 1997), toxic leadership (Lipman-Blumen, 2005), exploitative leadership (Schmid et al., 2019), destructive leadership (Einarsen et al., 2007), among others. According to Schyns and Schilling (2013) some of the differences between the approaches relate to intentionality (does the leader act on purpose?), the kind of behavior (e.g., physical or verbal), and the target (followers or organization).

Two main reasons make it critical to analyze this phenomenon: its consequences and prevalence. First, destructive leaders have a great impact due to the negative consequences of their behaviors on followers, the organization, and themselves (Mackey et al., 2021; Mullen et al., 2018; Schyns & Schilling, 2013; Tepper et al., 2006; Webster et al., 2016). Secondly, although studies based on the perception of abusive supervision tend to present low mean values (e.g., Ju et al., 2019; Schyns et al., 2018a; Tepper, 2000; Tepper et al., 2008), other studies show that there is an important prevalence of destructive behaviors by leaders in organizations. According to Aasland et al. (2010), 33.5% to 61% of individuals have experienced this exposure to destructive leaders. It is then reasonable to ask why followership is not acting according to its adaptive nature designed to curb destructiveness?

A possible answer is that leaders at organizations take a formal position whose continuance

is not affected by followers (i.e., organizations create an artificial context where leaders do not emerge naturally), and there are no natural (sometimes neither formal) mechanisms to stop the leader's behaviors. Although adverse, this can be considered a critical context to examine followership. In such an adverse context, in which practices to balance leaders' behaviors are not easy to implement, those who resist can be seen as the ones who want to take part in a change process (Collinson, 2006). However, within this specific context, studies have been focusing on followers that enable the leader's behaviors (Lipman-Blumen, 2005; Schyns et al., 2018b; Thoroughgood et al., 2012a). This aligns with the toxic triangle model (Padilla et al., 2007), where destructive leadership is a process that depends on the leaders, followers, and the context to occur. In line, Thoroughgood et al. (2018, p. 627) highlight the dialectical nature of destructive leadership, conceptualizing it as a "dynamic, co-creational process between leaders, followers, and environments, the product of which contributes to group and organizational outcomes". So, as other classifications, Padilla et al. (2007) explains how destructive leadership is developed but does not foresee a way to stop the process. Moreover, as it will be described later in this thesis, we can question whether some of these followers are followers (i.e., have a volitive behavior), and who they are following (i.e., their leader or the invisible one?). By extending followership research and considering that followers behave within a complex and complete behavioral spectrum, with different motivations, followership can be better understood and considered a powerful tool for organizations.

It is not meaningless to notice that "why followership?" should now be taken as "why do we need followership?". Followership is not just a matter of relevance due to its inherent and inevitable existence; it can now be conceived as an organizational asset. And this becomes even more relevant in contexts that represent a severe threat to leadership.

1.5. Aim and overview of the thesis

This thesis is designed to contribute both to a conceptual clarification on the topic and to advance knowledge on the followership field. At the same time, as previous researchers have highlighted, research on followership should also inform practitioners. Especially, it is time to change how followers are looked at (Riggio, 2020). To achieve this, it is essential to gain knowledge on followers' behaviors and their underlying mechanisms. Accordingly, the main goals of this work are to identify followers' behavioral profiles (e.g., confrontational resistance – based on Tepper et al., 2001) and dispositional facilitating pathways (e.g., coproduction beliefs, Carsten & Uhl-Bien, 2012).

The first study (Chapter 2) embraces the critical leadership studies discussion on the nature of followership (Blom & Lundgren, 2020; Collinson, 2017; Learmonth & Morell, 2017). While providing a literature review on the topic and offering an integrated list of followership typologies and classifications, a qualitative data collection was designed to analyze a full range of behaviors when followers face destructive leaders. A mixed-based data analysis uncovered six followership profiles. These findings extended previous research by providing an integrative approach to followers' behaviors and contributing to followership's conceptual clarification. Accordingly, results support the idea that leadership is a coproduced process, and a follower may decide not to follow a leader in order to follow the organizational good and his or her idea of leadership.

The second study (Chapter 3) builds on the idea of followers' beliefs in the coproduction of leadership as an important antecedent of resistance behaviors (Carsten & Uhl-Bien, 2012). Accordingly, this study analyzed the role of coproduction in activating constructive resistance mechanisms and decreasing dysfunctional ones. With an experimental scenario-based approach, this study extends previous research, as the mentioned relationships were tested across four conditions of destructive behaviors by a leader (abusive supervision, exploitative leadership, organization directed behaviors, and laissez-faire) while controlling alternative explanations based on a relevant theoretical framework (model of proactive motivation, Parker et al., 2010). The results from this study corroborate the role of coproduction beliefs. They show such beliefs are an individual resource that protects organizations from destructive leaders by activating constructive resistance behaviors.

The third study (Chapter 4) was designed to explore the differences between resistance behaviors, especially those that Tepper et al. (2001) named as dysfunctional resistance. According to the authors, dysfunctional resistance comprehends both confrontational and passive (subversive) behaviors. However, from the integrative typology developed in Study 1, it became clear that subversive resistance must be a dimension in itself. To check this proposition, a study was designed to test and compare the psychometric quality of a bi-dimensional vs. a three-dimensional solution of a resistance scale. Having found support that the dysfunctional resistance splits in two, another study (Study 2) was conducted to test its nomological validity. Findings showed that a sequential mediation occurs where abusive supervision increases negative affect, which in turn increases defensive silence, which ultimately leads to confrontational resistance. This model also tested boundary conditions as the mediation was moderated by sense of power, which showed that subversion emerges when followers feel no power resources (in their relationship with coworkers). All in all, findings

support the three-dimensional solution that comprises constructive, confrontational, and subversive resistance.

Transposing this line of research about resistance behaviors to current challenges in leadership (namely, those highlighted by Seeber et al. 2020 concerning AI-based agents and leaders), it is reasonable to ask to which extent the cognitive processes that operate when facing a destructive leader (and that can prevent destructive leadership) apply to the emerging figure of synthetic leaders, i.e., when the leader is not a human but an algorithm. This study (Chapter 5) reflects the fact that we live in a context where artificial intelligence (AI) is gaining more and more ground, and organizations are not an exception. In effect, some decisions are already AI-based (Larson & DeChurch, 2020). A lengthy debate on the topic draws attention to the possibility of wrong decisions, some holding an unethical nature (Kellogg et al., 2020). To address this, a two-wave scenario-based experimental study was designed to compare how moral-based individual processes explain the compliance or disobedience to an unethical request when an algorithm vs. a human makes the decision. The findings show no differences in the level of compliance with an unethical request; however, the process that explains this behavior when the requester is a human does not explain the behavior when an algorithm is the one that makes the decision. This study contributes to the literature on business ethics and accountability for unethical behaviors while calling for the need to develop research on the topic, as it suggests different mechanisms are used when facing non-human leader-like agents. This can fundamentally change the fundamental assumptions that have been the cornerstone of followership research. A brief summary of each study is provided in Table 1.1.

Finally, in Chapter 6, the implications of the thesis - both to theory and practice - are discussed, the main limitations are described, and some directions for future research are presented.

Table 1.1 *Studies' summary*.

Chapters	Research questions	Method and data analysis	Outputs	Outcomes
Chapter 2 [Study 1]	• How do people behave when facing destructive leaders? • What underlies different followers' behaviors? • How do different behaviors and motivations explain different profiles?	• Exploratory questionnaire (open ended questions) • Content analysis (MAXQDA) • Latent class analysis (R software)	• Six followers' behavioral classes • Two resistance classes: active and passive distinguished by strength • Three obedience classes: passives, conflict avoiders, and supporters distinguished by features such as fragility, fear, hierarchical expectations, or critical thinking • One mixed-behavior class	• Conceptual contribution: conceptual clarification on the leader-follower relationship, supporting the idea of resistance as an expression of followership • Practical implications: leadership, as a coproduced process implies co-responsibility
Chapter 3 [Study 2]	• Do coproduction beliefs increase constructive resistance to a destructive leader regardless of the type of destructive behavior? • Do coproduction beliefs decrease dysfunctional resistance to a destructive leader regardless the type of destructive behavior?	• Experimental design • Scenario based questionnaire • Multigroup data analysis using partial least squares (PLS) in Smart-PLS	• Coproduction beliefs are positively associated to constructive resistance regardless of the leader's destructive behavior • Coproduction beliefs are negatively associated to dysfunctional resistance but not for all destructive behaviors	• Conceptual contribution: coproduction beliefs as an individual-based resource • Practical implications: coproduction beliefs can be used to establish governance mechanisms.
Chapter 4 [Study 3]	• Is there a third resistance style? • Are confrontational and subversive resistance (previously merged and named dysfunctional) different behaviors? • How and when is abusive supervision associated to subversive and confrontational resistance, via negative affect and defensive silence?	• Study 1: confirmatory factor analysis (CFA) in Analysis of Moment Structures (AMOS) • Study 2: structural equation modelling (SEM) in AMOS.	• Study 1: There are three-resistance dimensions (constructive, subversive, and confrontational) • Study 2: Negative affect and defensive silence partially mediate the negative relationship between abusive supervision and confrontational resistance; conditional indirect effect between abusive supervision and subversive resistance due to sense of power	• Conceptual contribution: resistance holds three dimensions: constructive, confrontational, and subversive • Practical implications: silence does not necessarily mean "no behavior", and when no power resources are available, subversive strategies emerge

| Chapter 5 [Study 4] | • Does displacement of responsibility mediate the relationship between moral identity and compliance with fraud?
 • Does the supervisor nature (human vs. algorithm) moderate this mediation? | • Experimental design
 • Scenario based questionnaire (two-wave design)
 • Two-stage least squares (2SLS) regression in SPSS
 • Multigroup data analysis using SEM computed in AMOS | • Displacement of responsibility mediates the relationship between moral identity and compliance with fraud when the supervisor is a human but not when the supervision is an algorithm | • Conceptual contribution: extant knowledge on human-human interaction is not directly transferable to AI-based working settings
 • Practical implications: although using a yet unknown mechanism, when implementing AI-based systems, organizations should be aware that when asked to perform unethically, people obey to algorithms in the same way they do with humans |

CHAPTER 2

Can you be a follower even when you do not follow the leader? Yes, you can.[2]

Abstract

In the ongoing debate in the area of critical leadership studies, the nature of leader-follower relationships is a thorny issue. The nature of followership has been questioned, especially whether followers can display resistance behaviours while maintaining their follower position. Addressing this issue requires a dialectical approach in which followers and leaders alike are primary elements in leadership co-production. Followers who face destructive leaders are of special interest when leadership is studied as a co-creational process. This context favours the emergence of a full-range of behavioural profiles in which passives and colluders will illustrate the destructive leadership co-production process, and those who resist demonstrate that followers may not follow the leader and still keep a followership purpose. A two-step data analysis procedure was conducted based on the behaviour descriptions of 123 followers having a destructive leader. A qualitative analysis (i.e., content analysis) showed a set of behaviours and their antecedents that suggest three main groups of followers: resisters, obedient, and mixed behaviour. Treating these data quantitatively (i.e., latent class analysis), six followers' profiles emerged: active resistance, passive resistance, passive obedience, conflict avoidance, support, and mixed. Our findings provide evidence that followers who resist may do it for the sake of the organisation. We discuss our findings in light of followership theory, whereby joining role-based and constructionist approaches allows us to argue that followers may still be followers even when they do not invariably follow their leader.

Keywords: leader(ship), follower(ship), destructive leader(ship), followers' profiles, followers' resistance, CLS

2.1. Introduction

"In the context of over a century of studying leadership, we are still at the beginning of
followership research. There is a great deal of work to be done" (Riggio, 2020)

Recent research highlights conceptual issues pertaining to the studies of leadership calling for
conceptual and empirical clarifications on the nature of the leadership-followership relationship
(Blom & Lundgren, 2020; Collinson, 2017). This call arises in a context in which these and
other authors (e.g., Learmonth & Morell, 2017) advance different positions regarding the
voluntary nature of followership.

The debate mirrors the demanding evolution of the followership approach, in which several
concepts such as leader, leadership, follower, and followership coexist with distinct meanings
and assumptions attached to them. Accordingly, some reviews (Crossman & Crossman, 2011;
Shamir, 2004; Uhl-Bien et al., 2014) demonstrate that this evolution has departed from the
traditional leadership studies (where the focus was placed on leaders) to a real followership
approach.

Followership involves the study of "the nature and impact of followers and following in
the leadership process", and whether this can be achieved by a role-based view (rank-based) or
by a constructionist approach (social influence process) (Uhl-Bien et al., 2014, p. 89). The
constructionist approach conforms to the debate in Critical Leadership Studies (CLS) on the
nature of leadership-followership (e.g., Collinson & Tourish, 2015). Accordingly, and
following a post-structuralist perspective, Collinson (2006, 2008a) argues that leadership holds
a power-based and knowledge nature that allows a dynamic relationship(s) amongst followers
and leaders to be co-produced, and in which followers can develop and deploy different
identities when playing their role in the leadership co-production: conformist selves may hide
different motivations but perform according to the leader's requests; resistance may emerge as
an oppositional behaviour to control practises and can be designed as a way to build an
alternative. However, resistance may entail harmful consequences (e.g., Latan et al., 2019) that
compel followers to disguise their behaviours. Underlying this identity is a dramaturgical self
(Collinson, 2006) in which a manipulative nature allows followers to alternate between
conformist and oppositional forms.

From our standpoint, clarifying these conceptual issues is critical, as labelling is more than
lettering. Langer and Benevento (1978) show some classic experiments in which participants
change their performance levels by being labelled as workers or bosses. This shows how

inducing power differences through labelling produces considerable behavioural and psychological effects (Hoption et al., 2012; Vanderslice, 1988). Conceptual clarification thus has a practical impact. Additionally, once we accept leadership is a co-produced process, we are holding both leader and followers accountable. A third motive to uphold the clarification concerns its possible conceptual redundancy, because if a) followership is ultimately the role followers play in the leadership process (e.g., Uhl-Bien et al., 2014), and b) leadership co-production includes both downward and upward tactics (e.g., Collinson, 2006), then c) followership is intrinsic to leadership (Rost, 1993, 2008). In line with this, followership is considered of such critical importance that it is not possible to conceive of leadership without it (Uhl-Bien & Carsten, 2018).

This study departs from previous research defining leadership as a dynamic and bidirectional-influential process in which the leader and followers both play their role(s) based on asymmetrical power conditions (e.g., Collinson, 2006; Collinson & Tourish, 2015). In such a complex and dynamic process, a follower may decide not to follow the leader. This occurs, for example, when the follower's upwards influence is intended to display resistance (e.g., Tepper et al., 2001). Learmonth and Morrell (2017, p. 265) asked, "In what sense can a person intelligibly remain a follower while simultaneously displaying dissent and resistance?" – stressing the apparent contradiction in terms. In other words, does not following the leader imply not being a follower?

To answer this question, one should depart from "why people follow?". Recent work (Bastardoz & Van Vugt, 2019) shows followership as an adaptive group process, where followers wilfully decide to follow an individual they acknowledge as leader. Conceiving the follower role as an individual decision stresses its voluntary nature (e.g., Alvesson & Blom, 2015). However, the organizational formal hierarchical structure may challenge this free-will assumption (Blom & Lundgren, 2020). This is not to say that voluntariness is not important when studying followership in organisations; it means that it needs to be carefully analysed.

Accordingly, empirical research has been called to contribute to the debate on the followership nature (Blom & Lundgren, 2020; Collinson, 2017). We agree that when a leader displays destructive or unethical behaviours (e.g., Pelletier et al., 2019; Thoroughgood et al., 2018) it is especially important to answer this call. We conceive this as a critical context to analyse these issues because it threatens the basic assumption that leaders and followers work together for the organisational good. Knowing that followership comprehends behaviours displayed in interaction with a leader with the ultimate purpose of meeting organizational goals (Carsten et al., 2014), can someone decide not to follow a destructive leader so he or she can

follow the common organizational goals? Departing from Uhl-Bien et al. (2014) differentiation between the followership role-based and the constructionist views, we will answer this question throughout this work. From a role-based view, subordinates displaying resistance cease their follower role and engage in a manager-subordinate relation. However, the constructionist view allows these subordinates to display leading behaviours (e.g., resistance) and remain followers of the organizational good.

Enabling followers to take part in the leadership process gives them both the opportunity and the responsibility to think critically (Kelley, 1988) and behave courageously (Chaleff, 1995) to check the leader. Although it is well-accepted that destructive leadership is a dynamic process that depends not only on the leader's behaviours but also on followers and context (e.g., Kellerman, 2016; Padilla et al., 2007), followership typologies do not provide profiles of resistant followers (Lipman-Blumen, 2005; Schyns et al., 2018b; Thoroughgood et al., 2012a). Starting with Collinson's (2006, 2008) followership identities, and adopting an exploratory approach, this study contributes to the conceptual clarification involving leadership and followership by empirically identifying a full range of followers' behavioural profiles when facing leaders that display destructive behaviours.

In the first section we present the literature review to explore the nature of the leader-follower relationship. Destructive leadership is discussed as a co-produced and dynamic process. With the premise that followers are active agents in any leadership process, our focus moves to followers themselves and how they have been classified in the literature. This section finishes with followership typologies that have been conceptualised for the specific context of destructive leadership.

In the method section we provide information about the study's participants, procedure, and data analysis strategy, detailing the mixed nature (qualitative and quantitative) of this research. The results section thereafter describes the main findings of this study, which are then examined in the discussion section. Before presenting the main conclusions, we acknowledge some limitations and offer suggestions for future research.

2.2. Literature Review

2.2.1. The nature of the leader-follower relationship

Early leadership research overvalued the leader's role (Uhl-Bien et al., 2014). This is evident in leader-centric approaches in which followers are considered mere recipients (Uhl-Bien &

Carsten, 2018), and both success and failure are of the sole responsibility of leaders (Oc & Bashshur, 2013). Later, follower-centric perspectives on leadership brought followers into the equation. However, these approaches mainly focus on leaders, and how followers' perceptions help to construct leaders and to identify effective or ineffective leaders (Crossman & Crossman, 2011). Other perspectives, such as the relational and the role-based views, emphasise the wilful and active role of followers in the relationship with their leaders. The relational view emphasises that leaders and followers together achieve leadership goals (Uhl-Bien et al., 2014). The role-based perspective is well illustrated with behavioural typologies (e.g., Chaleff, 1995; Kelley, 1988). The evolution of these perspectives paved the way for the study of followership as "a relational role in which followers have the ability to influence leaders and contribute to the improvement and attainment of group and organisational objectives. It is primarily a hierarchically upwards influence" (Crossman & Crossman, 2011, p. 484). However, it was not until the emergence of the constructionist followership perspective that leaders and followers came to be considered as co-producers of leadership (for an historical review on the topic see Uhl-Bien et al., 2014). From this approach, the object of study is the interaction between leaders' behaviours and followers' reactions in producing the leadership process.

The co-production of leadership exposes the critical role followers may play in a partnership (with the leader) aimed to benefit the organisation (e.g., Carsten et al., 2018). Although working together, leaders and followers take different positions in this power imbalanced process (e.g., Collinson, 2008a, 2018). Accordingly, to advance knowledge on followership/leadership, one should first understand why people follow, as evidenced by Lapierre and Carsten (2014) book's title (*Followership: What is it and why do people follow?*). This question is of primary interest since being a follower is a decision, i.e., a voluntary action (Lapierre, 2014), where there is no room neither for fear-based nor for blind obedience (Blom & Lundgren, 2020).

Why people follow is rooted in its evolutionary and adaptive nature that aims to guarantee the group's protection: "humans have a flexible followership psychology that enables them to select and follow the right kind of leaders under the right conditions, determine an appropriate engagement level, and switch from being a follower to a leader whenever appropriate." (Bastardoz & Van Vugt, 2019, p. 81). Therefore, followership is not confounded with a rigid hierarchy or formal position (Alvesson & Blom, 2015), as followers are those who decide who is the leader, based on the recognition of the ability to achieve the common good. Followers monitor the leader's behaviours continuously. Together, these assertions highlight the attributional nature of leadership, where leaders are not just appointed but rather acknowledged

as leaders (Hinrichs & Hinrichs, 2014).

All in all, the voluntary decision to follow the leader departs from the assumption that both the leader, and the group, follow an idea of a common good. Up to this point, we can compare the leadership process to a followership chain where all the actors (leader and followers) are aligned to follow organisational goals. In order to make the process efficient, different roles emerge: the followers follow the leader while the leader is following the common goals. In such a situation those "followers" who decide to resist to a leaders' legitimate influence are "non-followers", and, consequently, may harm the collective good (Carsten et al., 2014).

It is then timely to ask what happens when the leader is not perceived as legitimate, for example, because he or she no longer strives for common goals? When this happens, a fiduciary contract is breached, and previous assumptions are necessarily challenged. This is of interest within the organisational context where followers, although recognised as active and protective agents (Carsten & Uhl-Bien, 2013) cannot replace a bad leader (Bastardoz & Van Vugt, 2019). For all these reasons, followership needs to be explored when basic assumptions are disrupted, answering the call of Blom and Lundgren (2020) to extend knowledge on the voluntary nature of followership.

As mentioned, when followership is conceived as a chain, the ultimate purpose is to achieve the organisational good. This is in line with Carsten et al. (2014) claim that followership is all about how followers relate with leaders with the aim of meeting organisational goals. Thus, when perceiving that the leader is blocking that flow, followers can voluntarily decide to challenge the leader or to resist (Carsten et al., 2010; Carsten & Uhl-Bien, 2013). When they make such a decision, they keep their formal role (i.e., subordinate) but are no longer *that* leader's followers. At this point, three things may happen: a) the follower conforms (Thoroughgood et al., 2012a); b) the follower decides to leave (Schmid et al., 2018); c) the follower resists and challenges the leader's attempts (Tepper et al., 2001). There is no voluntary will in following the leader in any of these situations, ceasing the subordinate's follower role (Blom & Lundgren, 2020). However, minding the chain idea, those subordinates who voluntarily decide to challenge the leader to seek for an alternative (Collinson, 2006) will remain followers, that is organisational followers. Thus, studying followership within this specific context allows us to propose two layers of analysis: (1) a dyadic level (leader/manager-follower/subordinate relationship) related to a role-based approach; (2) a systemic level (organisational goals and follower/subordinate relationship) that recalls a constructive view. By dissecting these levels of analysis, we can conceive resistance as not following the leader (i.e., being a subordinate in a manager-subordinate relation) while following the purpose of a

common good (i.e., being an organisational follower) and reiterating their role in leadership co-production (Carsten et al., 2010).

From what has been described, it becomes clear that the explanation of the conceptual nature of the leadership-followership nature is of critical importance not only for theory but also for practical purposes, for example attributing responsibility when a leader displays destructive or unethical behaviours (e.g., Pelletier et al., 2019; Thoroughgood et al., 2018). In power-imbalanced situations, implicitly assumed in the leader-follower relationship (Blom & Lundgren, 2020; Magee et al., 2005), those who are powerless may displace their responsibility, and commit crimes of obedience – "I did what I have been told to do" (Beu & Buckley, 2004; Carsten & Uhl-Bien, 2013); or, to gain an advantage, may use strategies such as ingratiation (De Clercq et al., 2019; May et al., 2014). Total responsibility for destructive leadership can be wrongfully placed on the leader from a leader-centric perspective. However, when one takes a co-production perspective it becomes clear that leadership is not an exclusive leader-dependent-process (Collinson, 2005), but one in which followers take part and can resist and/or challenge the leader. Co-production of leadership implies a co-responsibility (Uhl-Bien et al., 2014).

2.2.2. Destructive leader(ship)

Echoing the co-production thesis, some researchers (e.g., Kellerman, 2016; Thoroughgood et al., 2018) approach destructive leadership as a dynamic process that depends not only on the leader's behaviours but also on followers and context. Thus, the presence of a leader who displays destructive behaviours is not the same as destructive leadership. Destructive leadership emerges when susceptible followers enable those destructive behaviours (Padilla et al., 2007). As Kellerman (2016) clearly states, "Leadership – It is a system, Not a Person" as it requires a leader, followers, and a context in which to occur. Accordingly, she presents a typology for bad leadership and not for bad leaders (Kellerman, 2005).

Given the exploratory nature of this study, we decided to follow a well-established and comprehensive definition of destructive leaders' behaviours. Moreover, as we place our focus on how followers deal with perceived leaders' destructive behaviours, the definition we choose emphasises this behavioural approach: "The systematic and repeated behaviour by a leader, supervisor or manager that violates the legitimate interest of the organisation by undermining and/or sabotaging the organisation's goals, tasks, resources, and effectiveness and/or the motivation, well-being or job satisfaction of subordinates" (Einarsen et al., 2007, p. 208).

This definition is compatible with the idea that a destructive leader is not the same as a

destructive leadership. The latter holds a dialectical nature that is expressed as a "dynamic, co-creational process between leaders, followers, and environments, the product of which contributes to group and organisational outcomes" (Thoroughgood et al., 2018, p. 627). This dynamic approach draws our attention to the role that followers play within this process (e.g., Kellerman, 2004; Lipman-Blumen, 2005).

2.2.3. Followers as active agents in the leadership process

The CLS dialectical perspective on leadership-followership leverages followers' potential on leadership co-production, arguing that they represent both knowledgeable and proactive agents (Collinson, 2005). An important consequence of this theoretical proposal comes from Collinson's studies (e.g., 2006) from which three followers' identities emerged: conformist (display disciplined behaviours), resistant (show their opposition and try to construct alternatives), or dramaturgical (manipulate their behaviour based on self-consciousness and knowledge).

Based mostly on a role-based view, there is a long history of authors pointing to the relevance of the followers' role (e.g., Pigors, 1934). However, it is only with Kelley (1988, 2008) that a typology comprising independent/critical thinking and activity level as two structuring dimensions gains salience. Table 2.1 summarises the main followers' classifications.

Table 2.1 *Followers' classifications.*

Author(s)	Dimensions	Types
Pigors (1934)	Followers' role in leadership effectiveness depending on their understanding or interest in the job	Constructive Routine Impulsive Subversive
Zaleznik (1965)	Dominance/Submission and Activity/Passivity	Withdrawn Masochistic Compulsive Impulsive
Steger et al. (1982)	Self-enhancement and Self-protection	The game player The achiever The kamikaze The bureaucrat The artist The apathetic The donkey The deviant The super follower
Kelley (1988)	Critical Thinking/Dependent Thinking and Passivity/Activity	Effective Sheep Yes people Alienated Survivor
Chaleff (1995)	Passivity/Activity and Courage to Challenge	Partner Implementer Individualist Resource
Potter and Rosenbach (2006)	Performance Initiative and Relationship Initiative	Politician Partner Subordinate Contributor
Collinson (2006)	Followers' identity in the workplace	Conformist selves Resistant selves Dramaturgical selves
Kellerman (2008)	Level of engagement with the leader	Isolates Bystanders Participants Activists Diehards
Howell and Méndez (2008)	Interaction with the leader	Interactive role Independent role Shifting role
Carsten et al. (2010)	How followers see followership	Passive Active Proactive
Sy (2010)	How leaders see followership	Followership prototype *Industry* *Enthusiasm* *Good citizen* Followership anti-prototype *Incompetence* *Conformity* *Insubordination*

Differently from the others, Collinson's (2006) proposal is founded on a constructionist view. This is meaningful as from a post-structuralist perspective, context plays a critical role (Uhl-Bien et al., 2014). Accordingly, processes are not stagnant but fluid, and in permanent construction, depending on workplace dynamics (Collinson, 2008a). For this reason, it is important to analyse how (and which) followers' profiles emerge under the context of a leader who displays destructive behaviours.

To our knowledge, none of the presented followers' classifications were developed through the specific lens of destructive leadership. However, some of their authors draw attention to the key role that followers may (should) play: "Instead of viewing followers as the "good soldiers" who carry out commands dutifully, we need to consider followers as the primary defenders against toxic leaders or dysfunctional organisations" (Kelley, 2008, p. 14). In line, Chaleff (1995) argues that followers should show the courage to challenge the leader. Carsten et al. (2010) show that followers view themselves as being able to challenge the leader. This behaviour, in opposition to an obedience role, requires that followers possess a belief in co-production (Carsten & Uhl-Bien, 2013; Carsten et al., 2018). Collinson (2006) highlights the importance of followers' resistance identity when leaders fail to perform in the best interest of the organisation, contrasting it with the conformity identity's negative impact. Kellerman (2004, 2016) stresses this negative impact, arguing that passive followers play an active role in bad leadership by allowing it.

2.2.3.1. Destructive leader(ship) followers

The literature is centred on three theoretically based classifications of followers that co-produce destructive leadership by enabling destructive leaders' behaviours (Table 2.2). Lipman-Blumen (2005) proposes types of followers that can favour toxic leaders (benign followers and the leader's entourage) or overthrown nontoxic ones (malevolent followers). According to the author, there are two kinds of benign followers: anxious and pragmatic. Anxious followers seek safety, and pragmatic followers tolerate destructive behaviours because they are guided by an instrumental orientation (e.g., fear of losing the job). The leader's entourage represents a type of follower that acts as the leader's alter ego. Malevolent followers are those who intentionally act to undermine the leader.

Thoroughgood et al. (2012a) argue that followers can be conformers or colluders. Conformers tolerate destructive behaviours less actively and can be classified into three types: lost souls (emotionally dependent), bystanders (driven by pragmatism), and authoritarians (act

under hierarchical expectations). Colluders have a more active role, and it is possible to find followers that are acolytes (share the leader's values) and those who are opportunists and aim to be the next leader. In a similar vein, Schyns et al. (2018b) propose five types of followers with different kinds of involvement in destructive leadership. Victims strengthen the leader's destructiveness due to their fear. Bystanders allow destructiveness as they simply observe what is happening. Conformers follow the leader's orders to avoid punishment. Colluders align with the leader for personal gains, and they may share the leader's values or be opportunistic. Lastly, puppet masters are those who push leaders to destructive behaviours.

Table 2.2 *Followers' classifications within the destructive leadership context.*

Author(s)	Dimensions	Types
Lipman-Blumen (2005)	Level of participation in destructive leadership	Benign (anxious) Benign (pragmatic) The leader's entourage Malevolent followers
Thoroughgood et al. (2012a)	Level of participation in destructive leadership	Conformers (lost souls) Conformers (bystanders) Conformers (authoritarians) Colluders (acolytes) Colluders (opportunists)
Schyns et al. (2018b)	Level of participation in destructive leadership	Victims Bystanders Conformers Colluders Puppet master

Typologies partially overlap (anxious followers, lost souls, and victims; pragmatic, bystanders, and conformers; the leader's entourage, acolytes, and colluders; malevolent and puppet master) and comprise only followers that in one way or another enable destructive behaviours. As these typologies pertain to the identification of the destructive leadership process (Pelletier et al., 2019), they do not comprehend followers that could help to stop the destructive leadership phenomenon. However, it is known that even under oppressive conditions followers resist (e.g., Thomas, 2003). We contend that this is an important research gap on which we support the goal to provide an integrative and empirically based typology.

Even though followers' classifications overlook a full-range behaviour model, some empirical studies have been analysing followers' specific behaviours. Studies focussing on followers' passivity toward destructive leaders mostly concern crimes of obedience (e.g., Beu & Buckley, 2004; Carsten & Uhl-Bien, 2013; Gibson et al., 2018), and silence (e.g., Kiewitz et al., 2016; Lam & Xu, 2019; Xu et al., 2015). Studies focusing on followers' challenging

behaviours include topics such as insubordination (e.g., Mackey et al., 2021), resistance (e.g., Carsten & Uhl-Bien, 2013; Greenbaum et al., 2013; Tepper et al., 2006), and whistle-blowing (e.g., Latan et al., 2019; Near & Miceli, 1985; Reckers-Sauciuc & Lowe, 2010). More broadly, some studies have drawn attention to followers' coping strategies (May et al., 2014; Nandkeolyar et al., 2014; Yagil et al., 2011) within the context of destructive leaders' behaviours. These behaviours range from passivity to active challenge. Although in different ways, all these behaviours express followers' involvement in the leadership process (Kellerman, 2016). However, sometimes, when facing leaders with destructive behaviours followers decide to withdraw and resign their formal and constructionist participation in the leadership process, by quitting the organisation, or intending to do so (Tepper, 2000). When followers are not aligned with the leader and opt to stay, a possible reason is that they will do it for the sake of the organisation, or they still believe things can change. Accordingly, research shows that turnover intention and organisational commitment are strongly associated (Haque et al., 2019; Vandenberghe & Tremblay, 2008). Intention to stay in the organisation can be taken as an indication of organisational pride, but people can also stay due to a lack of alternatives (e.g., Hom et al., 2012). Reluctant stayers were recently found to show more dysfunctional behaviours such as workplace deviance (Singh, 2019). Overall, these findings suggest that followers who simultaneously display resistant behaviours and experience low levels of turnover intentions, probably match "resistant identity" and thus are striving to build an alternative (Collinson, 2006).

Summarising, we reason that followers' resistance towards destructive leaders' behaviours can be examined in light of a followership approach (Uhl-Bien et al., 2014). This can be done by joining both role-based and constructionist views. In a role-based approach, the relationship is founded in formal hierarchical positions, and subordinates may not follow the manager (Blom & Lundgren, 2020). A constructionist approach comprehends a dynamic nature where power position shifts are allowed and a follower may take the lead (Bastardoz & Van Vugt, 2019). Accordingly, opposing behaviours might be taken as a temporary leading position where a subordinate and a follower role can co-occur at different layers: when one decides not to follow a destructive leader they will remain a role-based follower (i.e., a subordinate) that, in a constructionist approach, is taking a leading position to pursue organisational goals (i.e., a follower). For this reason, we contend that yes, one is still a follower even when one does not follow the destructive leader.

2.3. Method

2.3.1. Participants and procedure

Participants were recruited through snowball sampling using Qualtrics online survey software distribution link. Snowball sampling is an approach that is especially useful for exploratory purposes and reaching hidden populations (Babbie, 2014) and has been chosen in empirical studies targeting destructive leadership (e.g., Schmid et al., 2018; Shaw et al., 2011). Given the specific requirements of the sample, we opted to use this technique.

With an extensive online survey and assuming an exploratory approach (open-ended questions), we provided an adaptation of Einarsen et al.'s (2007) definition of destructive leader(ship), and asked participants whether they have had or are currently dealing with such a leader. Of 145 respondents, 96 (66.2%) declared having had or currently having a destructive leader. Among those with such experience, 51 participants provided enough data to conduct content analysis. As we were interested in broad and full-range behaviours, we asked participants to describe not only their personal experience but also to offer the example of behaviours shown by up to two co-workers when dealing with a destructive leader. We also asked respondents to provide their behavioural antecedents (e.g., motivations) and some demographic variables.

Given that not all participants answered the three conditions, we were able to gather only 123 descriptions of how followers deal/dealt with the destructive leader, their behavioural antecedents, and some demographic variables. We were also able to get descriptions of leaders' behaviours. The mean age of the followers is 38.1 years old ($SD = 11.4$), with 70.7% of women (87 cases) and 29.3% of men (36 cases).

2.3.2. Data Analysis Strategy

To identify how followers behave when dealing with destructive leaders, we opted for a two-step data analysis: qualitative content analysis followed by a latent class analysis (LCA). Content analysis enabled us to categorise followers' behaviours (and antecedents). Then, we defined a matrix where all behaviours (and antecedents) were classified as 2 (when present) and 1 (when absent). After analysing this matrix, we isolated three main groups (those who perform any type of obedient behaviour; those who try to challenge the leader; and followers that perform both). At that point, we computed LCA in order to explore the presence of "hidden" profiles within each group. Figure 2.1 summarises the overall process.

Phase	Procedure (method)	Expected Outputs (results)
Qualitative	**Exploratory survey**	
	Qualtrics online survey software Snowball approach Questions: 1) Can you please give some examples of that manager's actions/behaviors? 2) Two coworker's behavior / actions towards the manager 3) Coworker's characteristics that may help to explain his or her behaviors The same questions regarding the participant	• Descriptions of followers' behaviors • Followers demographic characteristics (gender and age)
	Content analysis	
	Mixed-nature analysis (deductive and inductive) 1) Leaders' behaviors a priori categories based on target and severity (e.g., abusive supervision; laissez-faire) 2) Followers' behaviors and antecedents' a priori categories based on conceptual matrix integrating existing typologies (e.g., to challenge; support; exit/turnover; fear; ambition) 3) Followers behaviors and antecedents' a posteriori categories when content do not match a priori ones 4) MAXQDA 2018 software to organize data and code segments	• Destructive leader's behaviors categories • Followers' behaviors categories • Behavioral antecedents' categories
Quantitative	**Uncover behavioral profiles**	
	Latent class analysis 1) R software PoLCA package 2) Basic latent class model 3) latent class regression modeling with covariates	• Latent Classes with covariates

Figure 2.1. *Data analysis strategy.*

2.3.3. Data analysis strategy: Qualitative content analysis

The data analysis process had a mixed nature combining both deductive and inductive approaches (Patton, 2015). We started our analysis classifying leaders' destructive behaviours according to previous literature (e.g., abusive, exploitative, and laissez-faire). For this purpose, we considered two main ideas: 1) Barbara Kellerman's (2005) continuum of destructive leader's behaviours ranging from ineffective to unethical behaviours, and 2) Einarsen et al.'s (2007) model of displaying destructive behaviours against the followers and/or against the organisation. Crossing levels of severity (Kellerman, 2005) with two possible targets (Einarsen et al., 2007), we labelled categories according to well-established classifications that match each of the described behaviours. For example, when someone described a leader displaying humiliating behaviours (severe behaviour) against followers (target), we classified it as abusive;

a leader showing ineffective behaviours (less severe) due to lack of involvement (organisation-directed) was labelled as laissez-faire.

Regarding followers' behaviours and antecedents' categories, we developed a conceptual matrix crossing existing followers' classifications (e.g., Chaleff, 1995; Collinson, 2006; Kelley, 1988) looking for behavioural patterns. We extracted *a priori* categories such as challenge, support, critical thinking, and hierarchical expectations. While developing the conceptual matrix we were faced with equivalent concepts, that is conceptual overlapping. The nature of those concepts may have different roots, but in the end, they translate into similar behaviours. For example, Zaleznik (1965) observes that some followers show dominance as an expression of the wish to control authority. In the specific context of facing a destructive leader, all these behavioural profiles express a challenging attitude. In a similar vein, when not aligned with the leader, Chaleff's (1995) "partner followers", Kellerman's (2008) "activists" and "diehards", and Collinson's (2006) "resistance selves" will display challenging behaviours. Each time encountered, we considered these as overlapping dimensions and merged them into a single category.

To shed light on possible motivations for resistance behaviour, and to enrich the debate on why resistant followers opt to remain or withdraw, we defined *a priori* a category pertaining to intention to quit. Whenever we identified content that did not fall into one of our *a priori* categories we created a new one. We used MAXQDA 2018 (Kuckartz & Rädiker, 2019) to organise data and code sentence segments.

The content analysis procedure followed Bauer's (2000) recommendations regarding the categorisation procedure (each *a priori* category has a fixed number of codes; codes are independent), coherence (*a priori* codes were derived from a conceptual matrix), transparency (presentation of the codes list, its frequency and illustration quote) and reliability measured through inter-rater agreement (a second independent coder coded 20% of the qualitative data, and a Fleiss' Kappa of 0.936 was obtained). The coding units were words, sentences, or paragraphs, each variable was coded only once, and our unit of analysis was each follower's description (Boyatzis, 1998).

2.3.4. Data analysis strategy: LCA

After coding followers' behaviours and antecedents, LCA was applied to find followers' profiles according to behaviour patterns and their antecedents. LCA is a widely used technique in a number of research fields (e.g., Birmingham et al., 2020; Liu & Vivolo-Kantor, 2020; Parr

et al., 2016) that enables multivariate categorical data testing and computes mixed models based on a probabilistic rationale (Hagenaars & McCutcheon, 2002), estimating the extent of each class in the population, and the probability of someone to belong to a specific class, which is useful to assign individuals in a categorisation system. LCA is an exploratory technique that is more flexible than other clustering procedures and enables the creation of latent variables (classes) form observed data (Oberski, 2016). Thus, this procedure does not use predetermined cut points and provides statistical fit indices that help to retain the best model. To increase data analysis sensitivity, we opted to analyse at the subcategory level, when existing. Moreover, only categories with a frequency of greater than three were considered for convergence's sake.

We used the R software's poLCA package and followed Linzer and Lewis (2011) recommendations to compute LCA. Firstly, we tested a basic latent class model using the behaviour variables. Then, we tested latent class regression modelling with covariates (behaviour antecedents or attributes) that we were expecting to operate as predictors of class membership. The number of classes (i.e., followers' profiles) was judged based on Bayesian information criterion (BIC) and Akaike information criterion (AIC) model comparison fit indices, where the lowest values show the best model. Whenever contradictory information was found, we have considered two recommended decision criteria: BIC is recommended for larger samples (Dziak et al., 2020), and model selection should also be supported on its interpretability (Finch & French, 2015; Oberski, 2016; Parr et al., 2016).

2.4. Results

Following the frame presented in Figure 1, this section starts by showing content analysis findings comprising the leader's destructive behaviours; followers' three major behavioural groups; and their attributes. Then, LCA results are showed to identify the number of final classes and which attributes (behaviour's antecedents) are linked to which profile.

2.4.1. Content analysis

Our results show that followers perceive leaders' unethical and ineffective actions (Table 2.3), in line with Kellerman (2005). Some participants report more than one type of leader behaviour, which means that the occurrences in frequencies are not mutually exclusive.

Table 2.3 *Leaders' behaviours content analysis results.*

Main and subcategories	Number of followers			
	(Resistance) (n=48)	(Mixed behaviours) (n=22)	(Obedience) (n=53)	(Total) (n=123)
Unethical	41	20	49	110
Follower-directed	41	19	49	109
Abusive	34	16	41	91
Exploitative	17	6	17	40
Autocratic	1	1	5	7
Organization-directed	6	3	6	15
Ineffective	19	10	16	45
Laissez-faire	9	3	3	15
Incompetence	16	8	13	37

When conducting followers' behaviours content analysis, we found three main groups: resisters, obedient, and mixed-behaviour followers (who display both opposing and obedience behaviours, or none of these). Table 2.4 presents the categories and an illustrative quote we found for followers' reactions when dealing with destructive leaders, and its frequency (number of followers displaying the behaviour). Resistance behaviours were classified as either active (when the leader is directly faced) or passive (when followers assume an opposing and uncovered behaviour). From active resistance, the most frequent behaviour is *to challenge* (behaviours aimed to stop the leader's action). Passive resistance's strongest subcategories are *to ignore* (behaviours that express the decision to avoid contact with the leader/keeping a formal relationship) and *task-oriented behaviours* (focus on job/tasks/demands).

Obedience behaviours followed the same classification system and are presented as active (those who decide to conform or act in conformity) and passive (those who enable by doing nothing). From active obedience, the most frequent categories are *support* (behaviours that back the leader's actions) and *subservience* (behaviours that manifest active obedience). The passive obedience dimension comprehends *passivity* (being passive to the leader's actions) and *to avoid conflicts* (behaviours aimed to address consensus and avoid conflicts).

The other groups of behaviours comprise horizontal behaviours (to protect or to oppose co-workers); actions that show levels of inconsistency such as sharing disagreement when talking with others but behaving passively with the leader, and pretending different behaviours with different people; and demonstrating the will to leave the organisation.

Table 2.4 *Followers' behaviours content analysis results.*

Main and subcategories	(resistance) (n=48)	(mixed behaviours) (n=22)	(obedience) (n=53)	(total) (n=123)
		Number of followers		
Resistance behaviours	48	9	-	57
Active resistance	35	4	-	39
To challenge *"I had a few confrontations with my manager when I would not agree with a way he wanted me to act and expressed my opinion"*	34*	3	-	37
Verbal aggressiveness *"there is constantly mutual disrespect. The lack of control is also very present and triggers the constant disrespect"*	3*	1	-	4
Passive resistance	18	6	-	24
Task oriented behaviours *"used to study/analyse job issues in detail"*	11*	4	-	15
To ignore *"talking with him only when necessary"*	12*	2	-	14
Obedience	-	10	53	63
Active obedience	-	2	22	24
Support *"In general, the behaviour was of serious support and, in some cases, more serious/evident, inaction and excuse of the leader's actions"*	-	1	10*	11
Subservience *"never questioned leader's decisions"*	-	-	10*	10
Ingratiation *"The x co-worker used to be very sweet and he made compliments to the manager in order to keep his job"*	-	1	6*	7
Passive obedience	-	8	33	41
Passivity / Inaction *"inaction"*	-	4	26*	30
To avoid conflicts *"used to try to find agreements and solutions"*	-	5	8*	13
Co-worker directed behaviours	3	4	1	8
Protective *"was friendly with other employees, she refused to sabotage or report on other employees and attempted to support other team members by assisting with duties"*	-	3	1	4

Against co-workers *"Lack of support and inflexibility of direct leadership generates and increases daily stress with consequences for direct colleagues"*	3*	1	-	4
Inconsistent behaviours	2	6	3	11
To pretend *"Saying one thing to the leader but saying the opposite out of her earshot"*	-	3	3*	6
Passive disagreement *"My co-worker took a passive role. She did not act or contest the decisions. However, when speaking with other employees she used to show displeasure with his decisions"*	2	3	-	5
Exit / Turnover Intentions *"Started looking for other job opportunities to get out of the company"*	4*	3	1	8

Note. Only frequencies identified with a star (*) were considered in the latent class analysis. These cases fulfil the requirement of having been mentioned three or more times and refer to the lowest categories' level.

Although they are not very frequently cited, we were able to find nine types of attributes reported three or more times. In the resistance group ($N = 48$), the most frequent categories were *strength* (feeling confident and empowered; "(…) strong personality, confident and fearless"), *high critical thinking* (being able to question and think clearly; "(…) didn't recognize leader's competency to manage the tasks the leader was supposed to do"), and *high competence* (being perceived as competent; "(…) perform better than the leader").

In the obedience group ($N = 53$) the most frequent categories were *fragility* (low self-esteem and self-confidence; "(…) more fragile/sensitive and with low self-esteem"), *fear* (displaying fear motivated reactions; "(…) because of feeling afraid"), *ambition* (ambition, greed, and opportunism; "(…) most likely because of the willingness to stay within a company and not losing the opportunity for a career growth"), and *low critical thinking*. In the mixed behaviours group, the most prevailing categories were *strength*, *ambition*, and *pragmatism* (being guided by behaviours' consequences; "(…) the need for monthly remuneration and being the only person in his household who contributed for the family's income". So, explanations of the observed behaviours are mostly of a dispositional, attitudinal, and emotional nature.

2.4.2. LCA

As detailed in the data analysis strategy section, we computed LCA with each of the three groups uncovered by the qualitative analysis (i.e., resistance, obedience, and mixed-behaviours). Table 2.5 shows the fit indices for models for each of the three groups. Regarding the resistance group, both AIC and BIC indicate there are two classes, which means there are two different resistance profiles. For the obedience group, the joint analysis of AIC and BIC, considering both the sample size and data interpretability (Dziak et al., 2020; Finch & French, 2015; Oberski, 2016), indicates three classes (i.e., three different profiles of obedience behaviours). For the mixed- behaviour's group, a single class model has the best fit to the data.

Table 2.5 *Fit indices for behaviour model decision.*

Group	Nr. of Classes	AIC	BIC	df
Resistance (n=48)	1	248.031	259.259	42
	2	**230.027**	**254.353**	**35**
	3	235.762	273.186	28
	4	244.417	294.939	21
	5	254.365	317.985	14
	6	264.950	341.670	7
	7	278.950	368.768	0
Obedience (n=53)	1	293.600	305.422	47
	2	268.871	294.485	40
	3	**263.684**	**303.090**	**33**
	4	258.814	312.012	26
	5	265.358	332.348	19
	6	272.708	353.490	12
	7	286.708	381.282	5
Mixed-behaviour (n=22)	**1**	**189.206**	**200.116**	**12**
	2	190.794	213.706	1

Note. Boldface type shows the chosen model.

As the latent classes were identified, we continued by conducting class regression modelling with covariates[3] (behaviour's antecedents). Table 2.6 shows the fit indices for different models. For the obedience group, AIC and BIC suggest, based on interpretability and sample size criteria (Dziak et al., 2020; Finch & French, 2015; Oberski, 2016), that the resistance group comprehended two classes while the obedience group comprehended three classes.

Table 2.6 *Fit indices for regression model decision.*

Group	Nr. of Classes	AIC	BIC	df
Resistance (n=48)	**2**	**230.679**	**260.618**	**32**
	3	241.441	290.093	22
	4	253.588	320.951	12
	5	272.478	358.553	2
Obedience (n=53)	2	268.615	304.081	35
	3	**263.156**	**322.264**	**23**
	4	282.728	365.480	11

Note. As these are regression models to identify predictors for different classes, we have not computed models comprising a single class. Boldface type shows the chosen model.

---

[3] As our findings showed that the mixed-behaviour group is a single-class group, it was not possible to test the model with covariates (i.e., different antecedents cannot predict belongingness to any subgroup as there is only one profile for the mixed-behaviours group).

Resistance classes were named active resistance (class 1; 55.6%) and passive resistance (class 2; 44.4%). *Active resistance* class is characterised by challenging behaviours such as "(…) dialogue-based confrontation" or "(…) opposed to decisions and impositions by the supervisor". We named the second resistance class as *passive resistance* because it combines more covert behaviours, such as ignoring the leader "(…) performing everything in detail, speaking to the person concerned [supervisor] only when necessary" and focussing on their own tasks "(…) trying to continue the way of doing things". Some indication of behavioural complexity can be found in direct confrontation, although there is a much lower than the active resistance class. Table 2.7 shows the prevalence of the behaviours in each of the classes. There is a 100% chance of a follower in class 1 to challenge the leader; 15% exhibiting a desire to leave the organisation; 7.1% demonstrating verbal aggressiveness. Regarding class 2, there is a 56.3% chance of a follower to ignore the leader; 51.7% to stay focused on their task; 34.3% to challenge the leader; 14.1% to exhibit behaviours against their co-workers, and 5.1% to manifest verbal aggressiveness.

Table 2.7 *Manifest variables probabilities for the two-class resistance group model.*

	Latent Classes	
	1 (55.6%)	2 (44.4%)
Manifest variables		
To ignore	.000	**.563**
Task oriented	.000	**.517**
To challenge	**1.000**	.343
Verbal aggressiveness	.071	.051
Against co-workers	.000	.141
Exit / Turnover intentions	.150	.000

Note. Boldface type shows the highest probabilities in each class.

Table 2.8 presents the results for covariates' influence. Only strength is a predictor of class differentiation. Those who have this characteristic are more likely to exhibit active resistance when facing destructive leaders.

Table 2.8 *Latent class regression results (Resistance group).*

	Coefficient	Std. error	*t* value	Pr(>|t|)
Class 1: Active opposition		Reference Group		
Class 2: Passive opposition				
(Intercept)	-93.561	.780	119.909	.000
Strength	-93.956	.780	**-120.414**	**.000**
High Competence	.914	1.181	.774	.444
High Critical thinking	.581	.935	.622	.539

Obedience classes were named *passives* (class 1; 46.6%), *conflict avoiders* (class 2; 36.4%), and *supporters* (class 3; 17%). The *passive class* is characterised by passivity behaviours such as "inaction behaviours". The *conflict avoiders* class shows an important chance of subservience "I used to have an attitude of submission and acceptance. Followed orders, even when did not agree" and appeaser behaviours "(…) was an affable and friendly person, who got along with everyone and avoided confrontation". The last class was named *supporters* as the behaviour that stands out is to support the leader "(…) tended to support the supervisor's actions". Table 2.9 shows the prevalence of the behaviours in each of the classes. There is a 100% chance a follower in the passives class shows passivity and inaction and a 4.1% chance of supporting the leader. Regarding the *conflict avoiders* class, there is a 46.7% chance of a follower to be subservient; 41.5% to avoid conflicts; 15.6% to be hypocritical; 15.6% to ingratiate themselves with the leader, and 6.7% to be passive and inactive. Within the *supporters class*, there is a 100% chance a follower supports the leader's behaviours and actions; 33.3% to ingratiate, and 11.1% to be subservient.

Table 2.9 *Manifest variables probabilities for the three-class obedience group model.*

	Latent Classes		
	1 (46.6%)	2 (36.4%)	3 (17%)
Manifest variables			
Passivity	**1.000**	.067	.000
To avoid conflicts	.000	**.415**	.000
Ingratiation	.000	.156	**.333**
Subservience	.000	**.467**	.111
Support	.041	.000	**1.000**
To pretend	.000	.156	.000

Note. Boldface type shows the highest probabilities in each class.

To compare covariates' influence between all classes (Table 2.10), we ran three paired comparisons (pair 1 = passives vs. conflict avoiders; pair 2 = passives vs. supporters, and pair 3 = conflict avoiders vs. supporters).

Table 2.10 *Latent class regression results (Obedience group).*

	Coefficient	Std. error	t value	Pr(>|t|)
Model 1				
Class 1: Passivity		Reference Group		
Class 2: Conflict avoidance				
(Intercept)	-2.295	.664	3.457	.002
Fragility	.515	.881	.584	.565
Fear	.118	1.108	.106	.916
Ambition	20.596	.664	**31.03**	**.000**
Low Critical thinking	-1.099	1.482	.742	.466
Hierarchical expectations	-16.526	.664	**-24.9**	**.000**
Class 3: Support				
(Intercept)	16.223	.099	163.259	.000
Fragility	-14.136	.099	**-142.252**	**.000**
Fear	-13.243	.099	**-133.265**	**.000**
Ambition	40.937	.099	**411.961**	**.000**
Low Critical thinking	-13.593	.099	**-136.797**	**.000**
Hierarchical expectations	-17.159	.099	**-172.681**	**.000**
Model 2				
Class 1: Passivity				
Class 2: Conflict avoidance		Reference Group		
Class 3: Support				
(Intercept)	-35.829	.099	-361.775	.000
Fragility	-13.128	.099	**-132.561**	**.000**
Fear	-12.506	.099	**-126.279**	**.000**
Ambition	65.590	.099	**662.284**	**.000**
Low Critical thinking	-11.883	.099	**-119.989**	**.000**
Hierarchical expectations	6.742	.099	**68.074**	**.000**

Characteristics such as fragility, fear, and low levels of critical thinking do not predict whether or not a follower belongs to conflict avoiders or passives groups. Fragility is least likely to be witnessed in the supporters group, which is also the least fearful. Ambitious followers are more likely to be found within supporters, closely followed by avoiders, and lastly by passives. The same is observed regarding hierarchical expectations, which are more likely to be observed in followers belonging to the passives, and least likely in those belonging to the conflict avoiders' group. Low critical thinking is a feature found equally in passives and avoiders but not in the supporters' group, which has the lowest likelihood of including such followers.

Correspondingly, the passive group most likely includes fragile, fearful, and unambitious followers who have the lowest critical thinking and the highest levels of hierarchical

expectations of all groups. Conflict avoiders stand out for being the group with the lowest hierarchical expectations, resembling passives in their fragility, fearfulness, low levels of critical thinking, and moderate levels of ambition. Supporters have the highest number of ambitious followers, co-occurring with the lowest fragility, fear, low critical thinking, and having intermediate levels of hierarchical expectations.

2.5. Discussion

The main purpose of this study is to contribute to the debate on the nature of followership. More specifically, we focus on the question "are followers who do not follow still followers?" This issue emerges from a conceptual debate in which authors call for empirical contributions (Blom & Lundgren, 2020; Collinson, 2017; Learmonth & Morell, 2017). Departing from a destructive leader-based context that challenges basic and evolutionary assumptions (Bastardoz & Van Vugt, 2019), we propose a two-layer approach based on Uhl-Bien et al.'s (2014) role-based and constructionist approaches. We argue that an integrative interpretation of these perspectives allows us to reason that, when attempting to resist to a destructive leader, subordinates (as defined by their position) may display leading-position behaviours (from a transitory power shift) aimed to achieve the organisational good, and thus remain organisational followers while not following the leader.

Leadership studies have long treated followers as mere recipients in a leader-follower relationship, placing attention on leaders' characteristics and styles (e.g., trait and behaviour leadership approaches) (Oc & Bashshur, 2013). This resulted in negative associations with the role of followers, who were considered passive agents (Agho, 2009; Chaleff, 2008). Overlooking followers' agency may have led to a conceptual confusion between the terms of leader and leadership. This theoretical misperception gained orderliness through the emergence of the followership field of studies, which has been gradually growing and mobilising researchers' interest in followers as important players in the leadership process (e.g., Carsten et al., 2010; Chaleff, 1995; Kellerman, 2008).

We believe that studying followers in the specific context of facing a dysfunctional leader will help to clarify the followership nature. This specific context allows us to analyse both destructive leadership co-production and resistance mechanisms, enabling the development of a full-range behavioural typology. Moreover, as we will detail throughout this section, the emergence of resistance profiles provides evidence that "not following" may be compatible with followership.

Accordingly, we looked for participants who have experienced a destructive leader(ship) situation and asked them about their leader's behaviours. Findings revealed two main categories in line with Kellerman's (2005) unethical and ineffective dimensions. Within the unethical dimension, followers' descriptions emphasised follower- and organisation-directed behaviours. The former includes abusive (Tepper, 2000; Tepper et al., 2017), and exploitative behaviours (Schmid et al., 2019). The latter was scarcely mentioned, which may reflect its less harmful effects on followers (Schmid et al., 2018) or possible gains for followers (Einarsen et al., 2007). Concerning ineffective behaviours, the emergence of both incompetence and laissez-faire categories provides evidence of its destructive nature (Aasland et al., 2010; Schilling, 2009; Skogstad et al., 2007). Ultimately, the main features of participants' descriptions match those proposed in the literature, which offers encouragement that our findings pertain to these types of leader.

The literature review suggests that when facing a destructive leader, followers may display diverse behaviours (e.g., obedience, ingratiation, and resistance). However, as mentioned, an integrative model is lacking. Our results mirror this complex behavioural portfolio showing individuals are complex entities where an all-or-nothing approach is illusory. Accordingly, we found different profiles within both the obedience and resistance groups. Within each of the subgroups (e.g. passive resisters, conflict avoiders etc.), behavioural consistency is expected (Sherman et al., 2010), meaning that we can anticipate followers to behave consistently when interacting with a destructive leader. However, followers' agency and contextual constraints interact and allow a shifting nature in followers' profiles (Collinson, 2006) which is evidenced by the mixed-group eclectic profile.

Essentially, findings regarding followers' behavioural profiles allow a more comprehensive interpretation of leadership's/followership's dynamic nature, in which one needs first to distinguish and then to integrate the role-based and constructionist perspectives. In light of this dialectic, we discuss our findings for the three main behavioural groups: obedience, resistance and mixed.

2.5.1. Obedience: When followers co-produce destructiveness

Through the naked eye, the obedience group recalls a traditional conformity role, resulting from a corrective nature of power (Collinson, 2018). Milgram's (1963) experiment is a classic illustration of this, in which obedience has a hierarchical nature. However, one must notice that compliance may disguise other motives. According to Fromm (1977), and his "Fear of

Freedom," obedience may derive from the security and comfort of simply leaving decisions to someone else. Shamir (2004) argues that conformity behaviours may camouflage calculative motivations. Regardless of the motivation, obedience behaviours reinforce the *status quo* (Kellerman, 2012).

Our findings shed light on how destructive leadership is co-produced in this group. The three-class model covers three different obedience profiles (passive followers, conflict avoiders, and supporters) expressing distinct processes that enable destructive leadership. Passive followers' behaviours are mainly characterised by showing total inaction regarding the leader's behaviours. Their most distinctive characteristic is the presence of hierarchical expectations that seem to emerge as justification for the follower's behaviour based on the idea of authority (e.g., Gibson et al., 2018; Milgram, 1963). These followers share some characteristics with individuals classified as conflict avoiders: fragility, fear, and low levels of critical thinking. These features, especially fear and fragility, have been highlighted by Lipman-Blumen (2005) and Thoroughgood et al. (2012a) in describing how followers enable leaders' destructive behaviour. Low levels of critical thinking are also related to less effective followership styles (Kelley, 1988).

In line with theoretical predictions on the topic, findings show a group of followers that support the destructive leader's behaviour (Lipman-Blumen, 2005; Schyns et al., 2018b; Thoroughgood et al., 2012a). Alongside the supportive behaviours, ingratiation and subservience behaviours are also observed in this class. Accordingly, Yagil et al. (2011) showed that ingratiation behaviours are a strategy that increases in the presence of a destructive leader. This group's most distinctive characteristic is ambition, which has been discussed within followership literature as a characteristic of followers that enables destructive leadership, also related to greed and opportunism (Lipman-Blumen, 2005). The existence of different profiles of obedient followers suggests that obedience is a multi-meaning concept that ranges from pure conformism to calculative behaviours, as proposed by Collinson (2008a) and Shamir (2004). Despite its diverse manifestations, we reason that obedience can be taken as consent in which individuals enable the destructive process. In other words, these actors are taking part and co-producing the situation, and as a consequence, are co-responsible for its destructiveness. Although all three obedience classes are co-responsible for the destructive leadership process, the motives underlying the profiles suggest different levels of alignment. This idea was proposed by Blom and Lundgren (2020) who discussed the involuntary nature of followership, extending previous research on the idea that followership (in opposition to the superior-subordinate relationship) implies a "voluntary reduction of influence" (Alvesson & Blom, 2015,

p. 272). Accordingly, it is reasonable to question whether the passive and conflict avoiders can be conceived as voluntary followers. Once their behaviour is rooted in motivations such as hierarchical expectation, and fear, it seems fair to consider them as subordinates, who fail the follower role at both the dyadic and the systemic layer. Conversely, those who support the leader's actions are leader's followers. However, they fail to meet the systemic follower role when we consider the organisational good. That being said, we also argue, in line with Collinson (2008a), that regardless of being leader-follower or superior-subordinate, the critical nature of these relations resides in power dynamics. Moreover, from our standpoint, the pivotal issue regarding followers' willingness to follow (Blom & Lundgren, 2020) is linked to followers' motives, which deserve further attention and research. We reason that this more in-depth inspection of followers' motivations to obey helps organisations to engage in practises and develop policies that supply or feed followers' needs and/or concerns (e.g., fear of losing job). If an organisation adopts policies whereby employees' participation is valued (e.g., providing feedback systems), we can expect that obedience elicited by hierarchical expectations will be reduced.

Regardless of the voluntariness of their behaviours, obedient followers display compliance behaviours and follow their leader. This willingness to follow shows how obedient followers and destructive leaders co-produce destructive leadership.

2.5.2. Resistance: When followers decide not to follow the leader

As mentioned, when facing destructive leaders, followers may display resistance behaviours (e.g., Tepper et al., 2001). When followers resist, they are by definition not following the leader. Not following the leader means not taking part in the destructive leadership process. With this idea in mind, it is appropriate to ask whether these followers are, indeed, followers. Recalling our two-layer approach in articulation with Uhl-Bien et al.'s (2014) views of followership - role-based and constructionist – we can interpret resistance as a manifestation of followership. Accordingly, these individuals are role-based subordinates in a dyadic level where they resist their manager's attempts. At the same time, from the systemic layer, the constructionist perspective allows individuals to display resistance as a power-based behaviour where followers do follow the organisational good. Moreover, the idea of scrutinising the leader, and act to benefit the collective, restores the evolutionary and adaptive role of followership (Bastardoz & Van Vugt, 2019), and is in line with works that highlight the active and courageous nature of good followers (e.g., Chaleff, 1995; Kelley, 1988).

There are both conceptual arguments in the literature as well as empirical grounds reported in this study that support the contention that followers may remain followers when they decide not to follow a destructive leader. When followers are not aligned and do not want to comply, they can opt to resist or leave. When leaving the organisation is not possible (Hom et al., 2012; Sheridan et al., 2019) followers are known to display deviant behaviours (Singh, 2019). There are two main forms or workplace deviance: organisational and interpersonal (Bennett & Robinson, 2000). Interpersonal deviance has been reconceptualised and it is now well established that it involves behaviours targeted against co-workers and supervisors (Hershcovis et al., 2007; Mitchell & Ambrose, 2007). Our findings show vertical deviance through verbal aggressiveness (*"(…) talk back to the supervisor"*) and horizontal deviance when followers describe behaviours against co-workers ("(…) spreading negativity through other co-workers"). The existence of resisters who simultaneously show low intentions to leave and low levels of deviant behaviours may suggest their pro-organisational motivation to resist a destructive leader. Under these circumstances, resisters may not follow the destructive leader precisely in order to not co-produce destructive leadership and thus remain faithful to and follow their ideal of "good" leadership. We agree that resistance may guide the goal of re-establishing the desired condition or building an alternative (Collinson, 2006), sending a message of disapproval (Tepper et al., 2001).

Our findings support this contention. Accordingly, it is informative that amongst the emerging categories there is a low frequency of intentions/behaviours of leaving the organisation. The probability of finding intentions to abandon the organisation was nil in the passive resistance class and rose to only 15% in those that fall into the active resistance class. Likewise, the probability of finding individuals who express verbal aggressiveness is only 5% in the passive resistance class and 7% in the active resistance class. Regarding behaviours against co-workers, the probability within passive resisters is only 14% and nil in the active resistance class. These probabilities make deviant behaviour very unlikely in the sample, which is also concomitant with a low intention to leave.

As mentioned, our findings showed two resistance profiles: those who play an active-challenging role and those that opt to take a more covert resistance strategy. This partially goes in line with Tepper et al.'s (2001) proposal of constructive resistance and passive dysfunctional resistance behaviours. The active resistance group shows a defiant profile in which followers dare to challenge the leader (e.g., Chaleff, 1995). With respect to the passive resistance class, the two main behaviours are ignoring the leader, and focussing on the task. Although it is possible to find followers who actively display resistance, this is not the core behavioural profile

of this group. We take passive resistance as operating at an underground level, which is in line with previous proposals of camouflage dynamics and cover behaviours linked to resistance (Collinson, 2005; Rodrigues & Collinson, 1995). We contend that passive resistance may be an alternative path of defiance preferred by those who do not possess enough strength to engage in a straightforward confrontation. Findings from the content analysis showed that strength is related to feelings of confidence and empowerment, which is linked to a perceived sense of power (Anderson et al., 2012). This perception is known to relate to voicing (Kim et al., 2019) and silence (Morrison et al., 2015) behaviours. Moreover, the exhibition of behaviours against the co-workers, although with low occurrence, is aligned with the idea that followers may displace aggression when they cannot carry it out against the transgressor (Mitchell & Ambrose, 2012), suggesting the lack of power to oppose the leader.

2.5.3. The mixed group: When followers play both sides

This group is the most ambiguous and needs further research. However, it may echo Collinson's (2006, p. 185) dramaturgical selves who can turn into "skilled manipulators of self and information" when facing more authoritarian contexts. Thus, these followers will become tacticians and typically display both resistance and conformist behaviours. Our findings detect followers who exhibit this mixed profile, supported by antecedents such as ambition and pragmatism. The complex nature of these individuals may justify its relative scarcer occurrence in the sample. Most of the descriptions within this profile combine passive obedience behaviours (e.g., inaction and avoiding conflicts) with active resistance (e.g., to challenge) or passive resistance (e.g., task orientation and ignoring the leader). It is unclear if followers deliberately engage in different behaviours at different moments, in a manipulative and calculated manner, or if these behaviours simply co-occur. However, this mixed nature can be an expression of disguised resistance (Collinson, 2018).

2.5.4. Limitations and future research

This study presents some limitations that may guide future research. This work is based on the well-established assumption that followership assumes working towards organisational goals (Carsten et al., 2014). However, this may not match all possibilities. Accordingly, future research may challenge this assumption and ascertain its limits, namely, how strong is this motivation when contrasted with selfish motivations when voluntarily following or resisting.

A larger sample may have been preferable, but one should note that data were collected from open-ended questions, and previous or present experience with destructive leaders was required. This may explain the relatively low-frequency levels for attributes. Despite this, results show some influential features such as strength or ambition.

Working simultaneously with self-reports and observer ratings can be a limitation since the cognitive processes may bias data. However, there is an indication in the literature that observer ratings can be as precise as self-reports in such latent constructs as personality traits (Oh et al., 2011), and conducting our analyses with and without self-report data showed the same number of classes for the main behavioural groups.

This study reports specific types of destructive leaders' behaviours. However, we did not segment the followers' analysis accordingly, due to insufficient frequency. Considering Schmid et al.'s (2018) findings, future research may consider targeting each of the subsamples to test whether our results depend on a specific leader's action(s). Moreover, as destructive leadership is a dynamic process (e.g., Kellerman, 2005; Padilla et al., 2007), it will be important to study how contextual factors influence followers' behaviours. This issue seems particularly important in order to gain knowledge of followers belonging to the mixed behaviour group.

Our results show that resistant followers scarcely mention any intention to quit, displaying challenging strategies. It will be important for future research on this topic to collect data regarding followers' organisational commitment (Allen & Meyer, 1990) to assess whether followers' low rates of turnover intentions are related to organisational loyalty (normative or affective commitment) or lack of alternatives (continuance commitment). However, one should notice that although commitment is negatively related to turnover intentions (Meyer et al., 2002), in the presence of abusive supervision, normative commitment is expected to be linked to voicing behaviours, and continuance commitment to be related to low levels of voice (Gabler & Hill, 2015). It is, therefore, reasonable to expect that the resistance group shows low intentions to leave due to ethical and moral values, unlike passive followers (from the obedience group) who simply have no alternatives.

A more detailed and in-depth analysis of followership and the shifting nature of followers' roles (resulting from the interaction between agency and contextual constraints) require an extension of our work, specifically targeting the mixed group. Such a group's emergence provides evidence that followers are not static agents and may engage in different and sometimes contradictory behaviours. As mentioned, the nature of the study design precludes a deep understanding of this group and calls for qualitative and longitudinal research specifically designed to comprehend all its complexity. Although behavioural consistency is likely (e.g.,

Sherman et al., 2010) and favours group intra-homogeneity, we expect that longitudinal designs and non-structured qualitative methods (e.g., interview) will provide empirical robustness to this issue. To close this section, we must highlight that despite the need for further research, the mere emergence of a mixed group in our study allows us to gain evidence on the non-static nature of being a follower.

2.6. Conclusion

The nature of leadership-followership remains immersed in an intense debate (Blom & Lundgren, 2020), especially in the area of critical leadership studies (Collinson, 2014; Collinson, 2017; Learmonth & Morrell, 2017; Tomkins et al., 2020). In line with a dialectical approach to leadership-followership dynamics, in which workplace conditions and power imbalances play an essential role (Collinson, 2008a, 2018), we studied followers' profiles when they face destructive leaders.

This study provides two major contributions. Firstly, our results contribute to a conceptual clarification regarding the nature of followership, by showing why and how resistance should be considered in a comprehensive followers' behaviour model. Moreover, this study helps to consolidate the post-structuralist perspective, in which followers (although or because (Foucault, 1977, 1979) of being powerless participants) play an active and crucial role in leadership dynamics. Secondly, to achieve our main purpose, we developed an integrative empirical-based model of six behavioural profiles that emerge in the presence of a destructive leader: active resistant followers, passive resistant followers, passives, conflict avoiders, supporters, and mixed behaviour followers. Although followers volition and agency have been widely considered (e.g., Chaleff, 1995; Kellerman, 2008; Kelley, 2008) as the most important factors in uncovering followers' types, within the destructive leadership context, they are still theoretically-based and favour followers that enable destructive leadership (Lipman-Blumen, 2005; Schyns et al., 2018b; Thoroughgood et al., 2012a). Our results largely confirm such profiles but highlight a group of individuals who opt not to follow the leader, and display resistance seemingly for pro-social reasons. Accordingly, the low levels of intention to quit within the resistance group suggest that these followers are active producers of the leadership process.

Moreover, joining role-based and constructionist perspectives in a two-layer approach allows us to uphold the idea that followers may still be followers even if they do not follow the destructive leader. This claim is not inconsequential, as we have detailed in the introduction section, because labelling is not lettering. Accordingly, it is important to close this work with

two critical assertions: (1) regardless of the voluntary nature of followers' (subordinates') obedience, following a destructive leader (supervisor) helps to develop destructive leadership (organisational destructiveness); (2) followers who decide not to follow the leader are nevertheless organisational followers, and play a critical role in the non-destructive leadership process.

Coproducing leadership: A reason to resist destructive leaders[4]

Abstract

Follower's individual differences have been receiving increased attention in studying destructive leadership because followers may enable or disable it. One of these yet under-researched features is the role of followers' leadership coproduction beliefs (a role construal) in explaining their resistance to destructive leaders. Departing from the proactive motivation theory, this paper explores the robustness of coproduction beliefs by testing its ability to predict followers' resistance to destructive leaders across four situations – abusive supervision, exploitative leadership, organization directed behaviors, and laissez-faire. With a sample of 359 participants that answered a scenario-based survey, findings support the idea that individual differences in coproduction beliefs contribute to explain resistance behaviors. Constructive resistance is always favored by coproduction beliefs independently of the leader's type of destructive behavior. Dysfunctional resistance, however, is sensitive to the leader's type of destructive behavior. These findings are discussed while detailing some practical implications and offering insights for future research.

Keywords: followership, coproduction, role construal, resistance, destructive leadership

---

[4] This chapter is under review in an international indexed journal.

3.1. Introduction

Research on the dark side of leadership has been gaining momentum (Zhu et al., 2019), in part because of its nefarious consequences (Mackey et al., 2021; Schyns & Schilling, 2013). For a long time, the focus of destructive leadership was single placed on leaders (Thoroughgood et al., 2018). However, followers, are now considered an intrinsic part of the leadership process (for a review, see Uhl-Bien et al., 2014) and play an active role in trying to curb destructiveness (Almeida et al., 2021; Wee et al., 2017). There are many ways for followers to express disapproval about the leader's behavior (e.g., resistance, voice, retaliation), and there are differences between them (Brett et al., 2016). This study focuses on resistance behaviors (Tepper et al., 2001), and places special attention on constructive resistance as an expression of ethical followership (Carsten & Uhl-Bien, 2013).

Research has uncovered some resistance roots (Brett et al., 2016; Greenbaum et al., 2013; Tepper et al., 2001). Among these, followers' beliefs on the coproduction of leadership (Carsten & Uhl-Bien, 2012) are attracting researchers' attention. Although the relationship between coproduction and constructive resistance has been established (Carsten & Uhl-Bien, 2013), this study goes a step further as it examines the role of coproduction in light of Parker et al. (2010) model of proactive motivation. According to this model, efforts to produce work-related changes are driven by three motivational states: "can do", "energized to", and "reason to". Whereas the first and second are more volatile and context-dependent, the latter may have a more stable nature as it is linked to the individual's sense of personal responsibility (Curcuruto et al., 2019). Accordingly, we propose coproduction beliefs as a "reason to" motivation and study its role in explaining both constructive and dysfunctional resistance while controlling for affect ("energized to") and personal sense of power ("can do"). We contend these beliefs gain relevance when analyzed within a context where a leader displays destructive behaviors, because holding a sense of participation in the leadership process can be conceived as an important individual-level organizational resource.

Destructive leadership studies are known both to show a lack of clarity in operationalizing destructiveness (Schyns & Schilling, 2013; Tepper, 2007) and to remain focused on a single sort of profile (e.g., abusive supervision). However, previous research shows that different destructive leader's behaviors lead to different followers' reactions (Schmid et al., 2018). Accordingly, we designed an experimental study to test whether

coproduction beliefs explain resistance across different destructive profiles (abusive, exploitative, organization directed, and laissez-faire).

3.2. Literature Review

3.2.1. Followers' resistance

Destructive leaders' behaviors lead to many consequences related to the job, the organization, the individual, and the leader (Schyns & Schilling, 2013). Among these consequences, followers' actions that help to balance power in the leader-follower relationship can be considered especially relevant as they may contribute to change (Wee et al., 2017).

Tepper et al. (2001) introduced the study of followers' resistance mechanisms. The authors defined two resistance strategies: constructive resistance has a negotiation-based nature, while dysfunctional resistance comprises passive-aggressive responses. Constructive resistance expresses nonconformity and dysfunctional resistance is retaliatory, aiming to re-establish a sense of justice. The latter contributes to a conflict spiral that can be stopped using constructive resistance behaviors.

Both strategies lead to important consequences to the individual as well as to the job, and the leader. Accordingly, dysfunctional resistance is negatively associated with performance and job satisfaction (Vecchio et al., 2010) and positively related to the leader's emotional exhaustion, while constructive resistance results in lower levels of the leader's emotional exhaustion (Brett et al., 2016). However, these relationships are not context-free and Tepper et al. (2006) show that the leader-follower relationship plays a critical role in how leaders perceive resistance. Another study reveals that showing dysfunctional resistance to a leader that displays downward hostility, lessens negative consequences (Tepper et al., 2015).

Supervisors' abusive behaviors are an important antecedent of resistance behaviors (Tepper et al., 2001), and it is enough to witness abuse to activate resistance (Greenbaum et al., 2013). There are, additionally, follower-related features explaining resistance behaviors: goal orientation is positively associated both to constructive and dysfunctional resistance, and learning goal orientation is positively related to constructive resistance and negatively related to dysfunctional resistance (Brett et al., 2016); aggressive humor is associated to dysfunctional resistance (Goswami et al., 2015); low levels of

consciousness and agreeableness strengthen the relationship between abusive supervision and dysfunctional resistance, while high levels of consciousness heighten the relationship between abusive supervision and constructive resistance (Tepper et al., 2001), which also occurs when followers have high levels of moral identity (Greenbaum et al., 2013); coproduction beliefs are positively associated to constructive resistance (Carsten & Uhl-Bien, 2013). Based on what has been described, one can assert that individual differences, most of them ethic-based, play a key role in routing different resistance types.

3.2.2. How do followers' resist? The role of coproduction beliefs

When trying to change something, people need to have a reason to engage and persist in risky behaviors (Urbach et al., 2020) such as resistance (Carsten & Uhl-Bien, 2012). In line with this, coproduction beliefs deserve special attention because they inform about followers' commitment to leadership, which becomes critical when accounting for what makes followers resist or defer a destructive leader (Carsten & Uhl-Bien, 2013).

Shamir (2007) introduced the idea of coproduction, where both the leader and followers are causal agents and active contributors to leadership. This approach marks an important turning point in the study of followership and highlights the importance of followers own role perception (Carsten et al., 2010; Knoll et al., 2017). Uhl-Bien et al. (2014) present a role-based approach where the followers role is influenced by the way they perceive their duties and responsibilities. These coproduction beliefs are neither a state nor a trait (Carsten et al., 2018) but a cognitive mechanism - role construal - representing how individuals build different social roles (Biddle, 1986; Vial et al., 2021). Research has been showing that individual differences in role construal explain how individuals behave in the organizational context (Zellars et al., 2002). Among these studies, Simpson and Laham (2015) showed that individual differences in moral role construal (cognitive moral motives) are linked to moral judgment. This finding provides evidence on the cognitive roots of ethical decision-making, encouraging the study of coproduction beliefs as an individual difference. Accordingly, as predicted by the Theory of Planned Behavior (Ajzen, 1991), these beliefs underlie followership attitudes that will help predict followers' behaviors.

Carsten and Uhl-Bien (2012) proposed a measure of coproduction beliefs as these cognitive construals explain constructive resistance and voice (Carsten et al., 2013, 2018). These can be conceived as problem solving strategies (Mitchell & Ambrose, 2012),

contrasting with retaliatory behaviors (e.g., Skarlicki & Folger, 2004) such as dysfunctional resistance (Tepper et al., 2001, 2015).

Parker et al. (2010) introduced a motivation model whereby self-efforts to change the work environment depend on the activation of three motivational states: "reason to", "can do", and "energized to". Finding a "reason to" is related to the individual's will to engage in certain behaviors. The "can do" driver is linked to the perception of self-efficacy, cost and control. Feeling "energized to" relates to the emotional boost underlying proactive behaviors. Both "can do" and "energized to" can be conceived as more context-dependent as they result from a situation's assessment (i.e., "can I?"/"Is it risky?" and "how do I feel?"). Conversely, "reason to" refers to a stable individual difference as it is related to the individuals' integrated and identified motivation (Deci & Ryan, 2000) and sense of personal responsibility towards achieving a goal related to a constructive change (Curcuruto et al., 2019; Lebel & Patil, 2018; Parker et al., 2010). Moreover, as highlighted by Parker and Wang (2015), individuals may feel they are able to do something but will not do anything unless they have a reason to. This draws our attention to how followers define their roles, and the specific contribution of a moral role construal (Simpson & Laham, 2015), and is in line with previous research showing that beliefs may operate as dispositional variables helping to predict compliance with unethical requests (Blass, 1991). In light of this, we focus our attention on the "reason to" driver (coproduction beliefs), as we aim to understand individual-based differences that make people resist destructive leaders (Carsten & Uhl-Bien, 2013). According to the same authors, followers who hold deferential roles will show compliance with unethical requests, while those who report higher levels of coproduction beliefs engage in constructive resistance. However, it is yet unclear how independent are coproduction beliefs effects from its context.

3.2.3. Destructive leader's behaviors: the role of context

Although it is well-established that destructive leaders have an important impact on followers' behaviors (Schyns & Schilling, 2013), there is a lack of knowledge regarding how the different leader's destructive behaviors impact on followers (for an exception see Schmid et al., 2018). Leaders' destructive behaviors can be classified regarding their intensity and target (Einarsen et al., 2007). According to this classification, and in line with previous research (Schmid et al., 2018), there are three main leaders' destructive profiles: abusive (when leaders humiliate subordinates, Tepper, 2000), exploitative (when

leaders take advantage to overwork subordinates, Schmid et al., 2019), and organization directed (when leaders' actions deviate from organizational interest, Thoroughgood et al., 2012b). Schmid et al. (2018) have not considered the controversial laissez-faire behavior, and there is no consensual agreement on whether laissez-faire leadership is destructive or not (Schyns & Schilling, 2013). However, its negative effects (Judge & Piccol, 2004; Skogstad et al., 2007) suggest that destructive behaviors can be displayed in a more active or passive way and a destructive leader can be both related to unethical and ineffective behaviors (Kellerman, 2012).

The universality of coproduction effects has been under scrutiny. On the one hand, leaders may condition how followers express their own characteristics (Zhang et al., 2020), and distinct destructive profiles do have different consequences on followers (Schmid et al., 2018). Conversely, coproduction beliefs - as role construals - are taken as being stable and central in decision making (e.g., Vial et al., 2021) which can be encouraged for the case of coproduction beliefs as Carsten and Uhl-Bien (2012) found no moderation effect of context (consideration leadership and relational quality) in the relationship between coproduction and constructive resistance. This strengthens the claim that coproduction beliefs, as a relatively stable individual difference, operate independently of its context. Thus, the present study evolves on the idea that coproduction is an organizational resource powered by followers, who use these beliefs as a motivational driver to choose between resistance behaviors.

Constructive resistance is consistent with coproduction beliefs as they foster the willingness to participate. Conversely, dysfunctional resistance expresses the will to undermine the leader (Tepper et al., 2006) and is therefore inconsistent with believing in coproduction. We thus hypothesize that:

H1. There is a positive relationship between coproduction beliefs in leadership and constructive resistance regardless of the type of leader's destructive behavior.

H2. There is a negative relationship between coproduction beliefs in leadership and dysfunctional resistance regardless of the type of leader's destructive behavior.

The conceptual model is depicted in Figure 3.1.
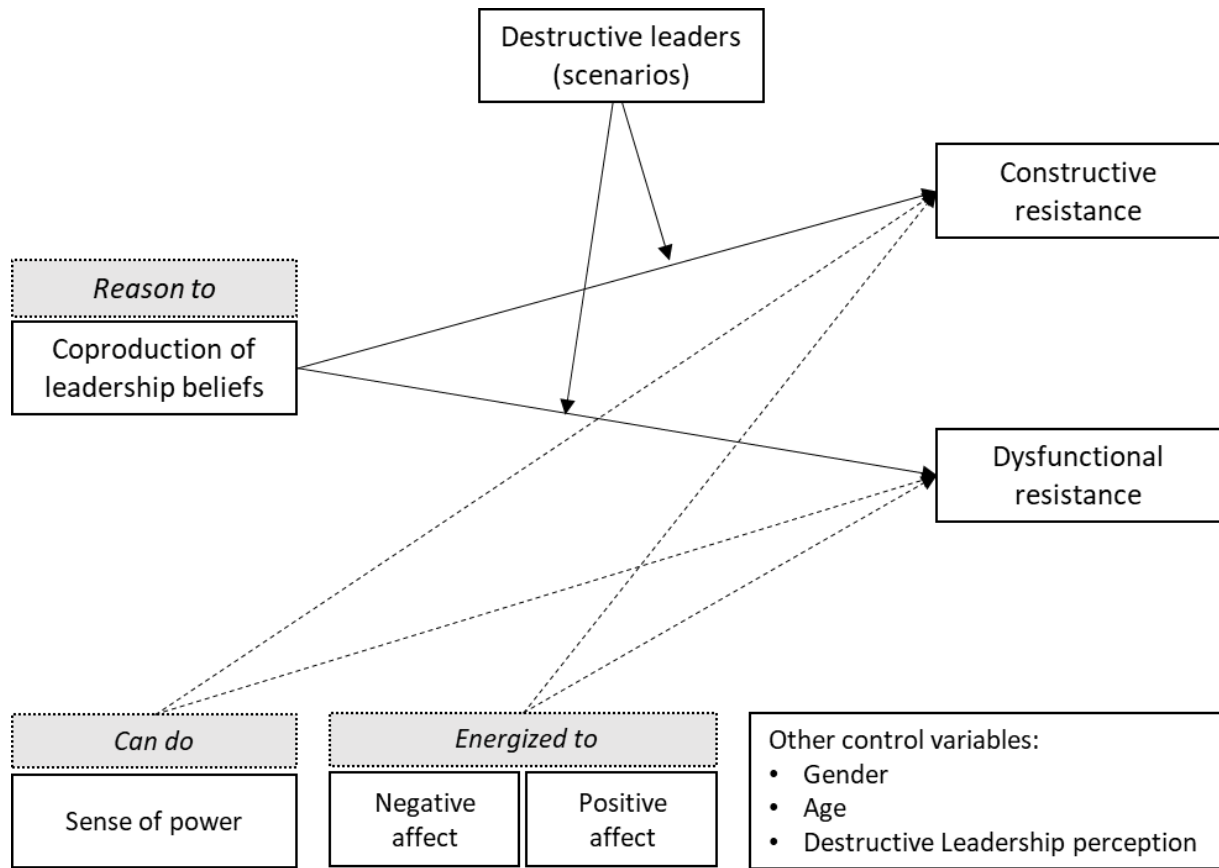
Figure 3.1 *Research model.*

3.3. Method

3.3.1. Procedure and Participants

To test the proposed hypotheses, 359 participants (M_{age}=39.95, SD_{age}=13.68; 68.2% female) took part in an online experimental study (using Qualtrics online survey software). First participants responded to the coproduction scale and demographic variables. Then, individuals were randomly assigned to one adapted scenario: abusive supervision, exploitative leadership, organization-directed, and laissez-faire (Christie et al., 2011; Hughes & Harris, 2017; Kelloway et al., 2003; Schmid et al., 2018). Then, participants answered the remaining scales.

3.3.2. Controlling for Alternative Explanations

As detailed in the literature review section, three important motivational drivers can help explaining followers' resistance. This study aims to analyze the "reason to" motivation

(i.e., coproduction), and to make sure we are able to test its impact on followers' behavior, we controlled the other drivers ("can do" and "energized to"). The "can do" comprehends the perception that one is able to attain goals, balancing its risks/feasibility (Urbach et al. 2020). So, individuals must own a sense of self-efficacy that allows them to pursue goals (Parker & Wang, 2015). The leader-follower interaction can be conceived as an interdependent and asymmetric power-based relationship (Collinson, 2005). To balance this relationship, the power gap should be reduced (Wee et al., 2017), and the follower needs to believe he or she can influence the leader. Moreover, it is known that covert behaviors (present in dysfunctional resistance) may occur due to power imbalance situations (O'Reilly & Aquino, 2011). The personal sense of power is the individual's perception of his or her ability to influence others (Anderson et al., 2012), and, therefore, we asked participants their perceived power to influence the leader. Feeling "energized to" is mostly linked to positive affect (Parker et al., 2010), especially activating positive affect (Bindl & Parker, 2010). However, negative affect is also informative as it is known to make people display defensive behaviors (Fredrickson, 2001), which may be critical to understand dysfunctional resistance. Accordingly, deactivating negative affect may inhibit risky behaviors (Urbach et al., 2020). For this reason, we asked participants how they would feel at work, having the described supervisor.

3.2.3. Measures

Destructive leadership. Abusive supervision was measured using three items from Tepper (2000) scale (α = .86; e.g., *Puts me down in front of others)*. Exploitative leadership was assessed with three items from Schmid et al. (2019; α = .84; e.g., *Sees employees as a means to reach his or her personal goals*). Organization directed behaviors were measured with three items from Thoroughgood et al. (2012b) anti-organizational behaviors subscale (α=.89; *Violates company policy/rules*) and laissez-faire leadership was assessed with Avolio et al. (1999) subscale (α = .87; e.g. *Avoids deciding*).

Coproduction of leadership was assessed with four items from Carsten and Uhl-Bien (2012) scale (α = .81; *Followers should communicate their opinions, even when they know leaders may disagree*).

The personal sense of power in the relationship with the leader was measured with four items from Anderson et al. (2012) scale (α = .82; *Even if I voiced them, my views would have little sway*).

Activated positive affect was measured with two items (enthusiastic and inspired) from the positive affect subscale of PANAS (Galinha et al., 2014; Spearman-Brown coefficient = .85). *Deactivated negative affect* was measured with two items (frightened and tormented) from the negative affect subscale of PANAS (Galinha et al., 2014; Spearman-Brown coefficient = .75).

Constructive and dysfunctional resistance were measured with four items each from Tepper et al. (2001) scales ($\alpha_{constructive}$ = .92; *I would explain that it should be done in a different way*; $\alpha_{dysfunctional}$ = .83; *I would disregard what this supervisor says*)".

All items were rated on a six-point Likert scale (1 = strongly disagree to 6 = strongly agree) with the exception of the resistance scale (1 = never to 6 = always).

Control variables. Following previous research, we controlled for age and gender (e.g., Brett et al., 2016; Haggard & Park, 2018). We also controlled "can do" (sense of power) and "energized to" (affect) motives. As the perception of leaders' behaviors is related to followers' resistance (Vecchio et al., 2010), we controlled destructive leadership for each scenario.

Whenever the scales were not available in Portuguese, we followed Brislin (1970) translation / back-translation procedure.

3.2.4. Manipulation check

One-way ANOVA and Tukey HSD post-hoc comparisons were conducted to test the manipulation check. There was a significant effect of abusive supervision manipulation [$F_{(3, 353)}$ = 61.72, $p < .001$] and, abusive supervision rates (M = 5.42, SD = 0.95) were higher than organization-directed (M = 3.61, SD = 1.43), exploitative (M = 2.96, SD = 1.44), and laissez-faire rates (M = 3.63, SD = 1.22). Similarly, the exploitative scenario showed a significant manipulation effect [$F_{(3, 355)}$ = 6.48, $p < .001$] exploitative leadership rates were higher (M = 5.16, SD = 1.04) than abusive supervision (M = 4.43, SD = 1.46), organizational-directed (M = 4.63, SD = 1.42), and laissez-faire rates (M = 4.48, SD = 1.02). There was also a significant effect of the organizational-directed manipulation [$F_{(3, 353)}$ = 72.53, $p < .001$], and organizational-directed rates were higher (M = 5.43, SD = 0.85) than abusive supervision (M = 3.25, SD = 1.44) exploitative (M = 2.89, SD = 1.54) and laissez-faire rates (M = 3.29, SD = 1.29). Finally, there was a significant manipulation effect for the laissez-faire condition [$F_{(3, 351)}$ = 39.76, $p < .001$], laissez-faire leadership

rates were higher ($M = 5.21$, $SD = 1.00$) than abusive supervision ($M = 3.71$, $SD = 1.42$), organizational-directed ($M = 3.29$, $SD = 1.67$), and exploitative ($M = 3.07$, $SD = 1.49$).

3.4. Results

The hypotheses were tested using PLS-SEM via SmartPLS3 software (Ringle et al., 2015). PLS-SEM analysis comprises two steps: the measurement model analysis, and the structural model assessment (Hair et al., 2016).

3.3.1. Measurement model

The scales reliability, convergent, and discriminant validity (Bagozzi & Yi, 1988; Fornell & Larcker, 1981; Hair et al., 2014) were verified. The heterotrait-monotrait ratio (Henseler et al., 2015) corroborated discriminant validity as all values were below .85. Overall, findings show good psychometric quality (Table 3.1). Table 3.2 present the correlations for each condition and for the total sample.

Table 3.1 *Reliability and convergent validity.*

	Mean (*SD*)	Composite Reliability					Average variance extracted				
		Total	Ab	Exp	Org	LF	Total	Ab	Exp	Org	LF
Coproduction	5.26 (0.66)	.88	.86	.89	.88	.86	.64	.60	.68	.65	.61
Constructive resistance	4.33 (1.19)	.94	.94	.94	.94	.95	.79	.79	.68	.78	.81
Dysfunctional resistance	2.45 (1.04)	.89	.86	.91	.85	.90	.66	.60	.71	.60	.70
Sense of power	3.02 (1.16)	.87	.85	.90	.86	.77	.63	.58	.69	.62	.47
Negative affect	2.94 (1.44)	.88	.90	.87	.90	.84	.79	.81	.77	.82	.72
Positive affect	2.17 (1.10)	.93	.93	.97	.91	.88	.87	.87	.93	.83	.78
Destructive leadership	-	-	.91	.96	.85	.90	-	.78	.90	.66	.75

Table 3.2 *Correlations per condition.*

	Abusive					Exploitative					Organization					Laissez-Faire				
	CP	CR	DR	NA	PA	CP	CR	DR	NA	PA	CP	CR	DR	NA	PA	CP	CR	DR	NA	PA
CR	.23*					.26*					.29**					.33**				
DR	-.26*	-.21*				-.15	-.08				-.05	-.06				-.34**	-.06			
NA	-.01	-.26*	-.10			-.11	-.27*	.44***			.02	-.20	.32**			-.07	-.21	-.01		
PA	-.13	.05	.17	-.21*		-.02	.14	-.20	-.21*		.00	.18	-.03	-.23*		.17	-.12	.12	-.26*	
SP	-.05	.33**	-.14	-.42***	.28**	-.12	.08	-.09	-.28**	.28**	.02	.08	-.45***	-.28**	.25*	-.04	-.23*	-.08	-.20	.29**

Note. CP = coproduction; NA = negative affect; PA = positive affect; SP = sense of power; CR = constructive resistance; DR = dysfunctional resistance

*p<.05; **p<.01; ***p<.001

3.3.2. Structural model

Following Hair et al. (2016) recommendations, we started by assessing collinearity. The highest VIF value was 1.46 (total sample and each group), showing no collinearity issues. Then, we obtained the path coefficients to test the hypothesized relationships. After that, the model's predictive power was assessed (coefficient of determination). To evaluate the impact of each exogenous variable, the effect size was calculated. To close the structural model analysis, the model's predictive relevance was tested by examining the Stone-Geisser's Q^2 value. The described results are depicted in Table 3.3.

Table 3.3 *Structural model results.*

	Total		Abusive		Exploitative		Organization		Laissez-Faire	
	β	f^2	β	f^2	β	f^2	β	f^2	β	f^2
Gender → CR	-.07	.01	-.16	.03	-.13	.02	-.04	.02	.00	.00
Age → CR	.03	.00	.03	.00	.00	.00	.02	.00	.00	.00
DL → CR	.13*	.02	-.10	.01	-.01	.00	.28*	.07	.22*	.05
CP → CR	.23***	.05	.25*	.07	.25†	.06	.23*	.06	.24*	.06
SP → CR	.08	.00	.32*	.10	.04	.00	.07	.01	-.17	.03
NA → CR	-.20**	.03	-.06	.00	-.18	.03	-.19	.04	-.24*	.06
PA → CR	.01	.00	-.05	.00	.11	.01	.09	.01	-.13	.02
R^2	.15		.22		.15		.21		.28	
Q^2	.11		.14		.08		.12		.08	
Gender → DR	-.10†	.01	-.07	.01	-.19†	.05	-.11	.01	-.05	.00
Age → DR	.09	.01	.14	.02	.14	.02	-.08	.01	.11	.01
DL → DR	.09	.01	.10	.01	.10	.01	-.01	.00	.07	.01
CP → DR	-.20***	.04	-.29*	.09	-.10	.01	-.03	.00	-.42***	.16
SP → DR	-.12†	.00	-.25	.06	.08	.01	-.42***	.19	-.15	.02
NA → DR	.15*	.02	-.16	.02	.48**	.24	.21†	.05	.04	.00
PA → DR	.03	.00	.17	.03	-.10	.01	.11	.02	.25	.06
R^2	.09		.18		.28		.28		.19	
Q^2	.05		.06		.17		.11		.01	

Note. †$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

3.3.3. Multigroup analysis

To test the moderation effect of the conditions, we conducted a multigroup analysis. Measurement invariance was assessed through the MICOM three-step procedure (Henseler et al., 2016). Configural invariance is assumed as the same factor structure is represented in all groups. Compositional invariance was also attained as indicators have equal weights across groups. Once configural and compositional invariance have been established, we are able to assume partial measurement invariance that allows group comparisons. Although analyzing the pooled data was not our goal, we conducted the last step in invariance testing that assesses equality between mean values and variances. Our models failed this step, so we kept our initial purpose on not interpreting the global results.

After establishing partial invariance, we used permutation testing (Hair et al., 2017) to determine if the model's paths were different between the four groups. Significant differences ($p < .05$) were found only for dysfunctional resistance and between two cases: exploitative versus laissez-faire ($\Delta\beta = .31$, $p < .05$), and organizational versus laissez-faire ($\Delta\beta = .38$, $p < .05$).

3.5. Discussion and Conclusion

This research aims to contribute to the study of coproduction as an individual-based organizational resource. Our results support the robustness of coproduction beliefs in explaining constructive resistance, as this effect occurs regardless of the leader's destructive behavior. Although the magnitude of the effects are mostly modest in size, they are in line with previous research (Carsten & Uhl-Bien, 2012). Moreover, there is no difference in the effects across groups. Unexpectedly, however, these beliefs reduce dysfunctional resistance only in the abusive and laissez-faire groups. We will discuss these findings while highlighting three ways in which this study extends previous research.

First, we studied the impact of followers' coproduction beliefs on resistance behaviors within a conceptual framework that explains proactive behaviors at work (Parker et al., 2010). Even when controlling for alternative explanations ("can do" and "energize to") coproduction does make a difference, especially for constructive resistance. Second, by examining this relationship across four scenarios, we answer the call for more conceptual clarification as regards the types of destructive leader behaviors in line with Schmid et al. (2018) while testing the robustness of these beliefs (i.e., its universal nature). Our findings support the claim that differences in coproduction beliefs may define ethical followership (Carsten & Uhl-Bien,

2013). Finally, extending previous research (Carnevale et al., 2018; Greenbaum et al., 2013) this study analyzed both dysfunctional and constructive resistance. Studying both types is useful as these behaviors hold divergent natures (Tepper et al., 2001) but are not mutually exclusive (Haggard & Park, 2018). Having found no association between constructive and dysfunctional resistance (except for the abusive condition), it seems unlikely one can infer one from the other. Against expectation, coproduction beliefs do not play a role in both exploitative and organization-directed conditions. According to Schmid et al. (2018), these two styles convey low hostility when compared with abusive supervision. Thus, hostility may activate the expression of coproduction beliefs in decreasing dysfunctional resistance. We should highlight that the items assessing dysfunctional resistance have a subversive nature mirroring passive resistance and therefore: when high hostility (abusive supervision) or no leading behavior (laissez-faire) is observed, believing in one's own active role as a follower reduces chances of behaving in a concealed manner. Hence, when destructive behaviors are more adverse, followers may need to go subversive. For the laissez-faire case, a follower who hold coproduction beliefs will not need to resist as they most likely will take the leading role in the relationship.

3.4.1. Limitations and Future Research

Findings must be interpreted at the light of the limitations of this study. First, it is a scenario-based research assessing behavioral intentions. However, this is a widely used approach in organizational studies, especially in ethics-related topics (e.g., Cianci et al., 2014; Tseng, 2019), and it is presented as a valuable alternative research strategy where large samples are not required (Ehrhart & Naumann, 2004).

Another limitation concerns the four scenarios, which do not allow room for mixed contexts as well as the occurrence of opposing behaviors. However, destructive behavior may damage the organization while protecting followers and vice-versa (Einarsen et al., 2007). Therefore, future research could explore resistance behaviors using scenarios that include these variations.

Another limitation concerns the exclusive individual-level focus of this study. Although findings are informative on the relevance of coproduction for constructive resistance regardless of the leader's destructive profile, the role of team as context may be considered in future research, e.g., by testing coproduction beliefs against group conformity.

3.4.2. Implications

Despite limitations, findings from this study offer a novel view and more extensive insight on resistance behaviors and coproduction by bridging conceptual frameworks (role theory and motivation for proactive behavior) and testing its stable nature. We trust our findings also have practical relevance by offering guidance for organizational design. Decision-makers interested in preventing this phenomenon may conclude that some of the strategies inspired in extant literature may not suffice. Namely, uncovering the roots of leaders' destructive behaviors (e.g., Zhang & Bednall, 2016) so that they become aware of such behaviors (Goswami et al., 2015), and acknowledge and apologize (Basford et al., 2014; Haggard & Park, 2018) to rebuild the relationship with followers. This may explain why destructive behaviors are prevalent (Aasland et al., 2010), and organizations do attract destructive profiles to leading positions (Chamorro-Premuzic, 2013). Therefore, a shift in focus, e.g., followership research, may offer novel solutions for this enduring problem.

In line with this, the present study highlights the importance of coproduction beliefs, which decision-makers may enhance to establish governance mechanisms that build on this role construal, thus fostering built-in immunity to prevent or break the destructive leadership cycle.

CHAPTER 4

Subversive resistance: When low sense of power fosters covert action[5].

Abstract

Although research reports that abusive leaders induce followers' resistance, these leaders are also known to have nefarious effects on followers due to the negative emotions they produce that can lead to silence and no visible resistance. Focusing on the idea of resistance reduction through fear and silence, we question whether all resistance behavior is visible and, therefore, likely to be silenced. Departing from Tepper et al. (2001) two-dimension resistance scale – constructive and dysfunctional resistance – we propose that dysfunctional resistance can be split in two (covert and overt behaviors). We designed a first study to reassess the psychometric nature of the scale, which showed a three-factor solution is preferable: constructive, confrontational, and subversive. Then, a second study tested the effect of abusive leaders on confrontational and subversive resistance via negative affect and defensive silence. It was then possible to verify that along with a positive direct effect, there is a negative effect of abusive leaders on confrontational resistance, but not on subversive resistance. An additional interaction effect was found for sense of power with colleagues, which showed subversive (or silent) resistance is activated in the absence of perceived power resources.

Keywords: abusive supervision, subversive resistance, constructive resistance, dysfunctional resistance, defensive silence, sense of power

---

[5] This chapter has been submitted to an international indexed journal.

4.1. Introduction

Destructive leadership is neither a cause nor a consequence: it is a process. For a long time, studies operationalized destructive leadership following a leader-centric approach (Thoroughgood et al., 2018). However, Padilla et al. (2007) through the toxic triangle model drew researcher's attention for the dynamic nature of the destructive process. Accordingly, to became destructive, the leadership process needs to find alignment between the leader's behaviors, the context, and followers (Kellerman, 2016). As regards to followers' consenting, two important remarks should be highlighted: a) followers may align even when they disagree with the leader's behavior (e.g., followers that are driven by instrumental purposes such as keeping their job); b) followers may do nothing, which means they are contributing to keep the *status quo* (Kellerman, 2012).

Taking destructive leadership as a process enables us to think about followers as active agents. In place of considering them as mere recipients of leaders' actions, followers have now gained their place (Crossman & Crossman, 2011; Shamir, 2007; Uhl-Bien et al., 2014). Within this line of research, many followers' types that enable the destructive process were proposed (Lipman-Blumen, 2005; Schyns et al., 2018b; Thoroughgood et al., 2012a), and recent research expanded this model introducing followership profiles that try to challenge destructive leaders in a more active or passive way (Almeida et al., 2021). Even before, some studies surfaced that put effort into uncovering followers' behaviors, such as insubordination (e.g., Mackey et al., 2019) or whistleblowing (e.g., Latan et al., 2019). Followers' resistance, another challenging behavior (e.g., Carsten & Uhl-Bien, 2013; Greenbaum et al., 2013; Tepper et al., 2006), emerges as a critical and still scarcely studied phenomenon. Tepper et al. (2001) conceived resistance as a two-factorial construct as it can hold either a constructive or dysfunctional nature. While not belittling previous efforts, we propound a reinterpretation of dysfunctional resistance; we contend it incorporates two types of resistance: confrontational resistance, and subversive resistance, which, somehow, mirrors previous research on opposing leaders' behaviors (Almeida et al., 2021). In addition to the scales' facial validity, we believe that the justification for this conceptual difference lies in the path that links destructive leaders' behavior to follower's resistance. The path arguably involves psychological variables where emotions and behavioral intention are key in explaining behaviors. Amongst these possible mediators, negative affect may play a role. It is known that destructive leaders favor negative affect (Schmid et al., 2018) and that negative affect is positively associated with defensive silence (Kiewitz et al., 2016). Such silence is motivated by self-protection and can be

automatically generated (Kish-Gephart et al., 2009), but may drive to different behaviors. Drawing on Conservation of Resources Theory (Hobfoll, 1989), this study tests whether defensive silence reduces a high-demanding resource strategy (confrontational resistance) but not a low-demanding one (subversive resistance). While acknowledging the impact of resource depletion, we argue that employees may hold external resources, such as perceiving power to influence others (i.e., their coworkers).

Accordingly, this paper is structured to test whether the dysfunctional resistance subscale by Tepper et al. (2001) comprehends not one but two types of resistance (confrontational and subversive) as well as the nomological validity of this proposal (Campbell, 1960; Netemeyer et al., 2003). It will start by introducing the claim and testing it, with a first empirical study focused on the comparative construct validity of different factorial solutions. It then builds the theoretical basis of the nomological model and reports the findings from its empirical test in a second study.

Study 1 - A third type of resistance

4.2. Literature Review

4.2.1. Dysfunctional resistance(s)

Destructive leader's behaviors encompass nefarious consequences (Einarsen et al., 2007; Mackey et al., 2017; Tepper, 2007; Xu et al., 2012). Abusive supervision, the most studied type (Schyns & Schilling, 2013), is described as "subordinates' perceptions of the extent to which supervisors engage in the sustained display of hostile verbal and nonverbal behaviors, excluding physical contact" (Tepper, 2000, p. 178). It is well documented that abusive supervision leads to many negative outcomes lowering performance (e.g., Walter et al., 2015), job satisfaction (e.g., Caesens et al., 2019) and engagement (Yan et al., 2020) while increasing turnover intentions (e.g., Peng et al., 2019), psychological distress (Park et al., 2018), and negative affect (Schmid et al., 2018).

Followers are mostly conceived as passive or resigned recipients of leaders' behaviors, which explains the extent of the negative outcomes; however, researchers showed that when facing abusive supervisors, followers may actively engage in opposing behaviors, like resistance (Goswami et al., 2015). The main types of resistance have been firstly operationalized by Tepper et al. (2001) with a dimension expressing constructive (dialogue-

based) and another one dysfunctional (passive/aggressive-based) resistance. Nevertheless, these two may not depict the full extent of possibilities as resistance may be displayed in a frontal or covert way (e.g., Collinson, 2006). A finer analysis of the dysfunctional resistance subscale (Tepper et al., 2001) reveals items expressing active opposing behaviors such as "saying no to a supervisor's request" – that we conceive as confrontational resistance - as well as passive ones such as "I ignore my supervisor" and "I act as if I was never asked to do it". The latter can be understood as an expression of subversion. If so, this subscale entails two subtypes of dysfunctional resistance, namely "confrontational resistance" and "subversive resistance".

To explore this possibility, we conducted a preliminary study to test the psychometric quality of Tepper et al. (2001) resistance scale via a confirmatory factor analysis contrasting a two-factor with a three-factor measurement model. From the items' facial validity, we expect dysfunctional resistance to split in two. Moreover, there is some encouraging evidence in published empirical studies that do not use the 9-item solution (e.g., Brett et al., 2016; Vecchio et al., 2010), suggesting that the dysfunctional resistance subscale might benefit from revision as from the original nine items.

4.3. Method

4.3.1. Participants and procedure

Three-hundred and sixty-nine participants[6] (68.5% female, M_{age} = 40.28, SD_{age} = 13.96;) answered an online survey that allowed to assess the factorial structure of Tepper et al. (2001) resistance scale.

4.3.2. Measures

Resistance was measured with Tepper et al. (2001) dysfunctional and constructive scales. A sample item for constructive scale is "I would explain that it should be done in a different way", and a sample item for dysfunctional resistance is "I would disregard what this supervisor says". The items were rated on a six-point scale (1 = never to 6 = always). The scale was translated to Portuguese, and then another researcher performed a back-translation to English (Brislin, 1970).

---

[6] As the purpose of this study was the examination of the factor structure of Tepper et al. (2001) resistance scale we reused the sample of study #2 to which some more valid entries were added afterwards.

4.4. Results

To judge the proposed 3-dimension structure, we assessed the reliability, construct and discriminant validity, and the confirmatory factor analysis for 5 models. Results are depicted in Table 4.1. Model A represents a unifactorial structure of resistance. Model B is the original 2-factor structure of resistance. Model C shows an alternative bifactorial structure judged upon the facial validity of the items. Accordingly, three items describing confrontational resistance were removed ("I would just say no"; "I would refuse to perform the request"; "I would tell this supervisor that I'm not available"). Model D excludes two items due to low loadings ($\lambda = .41$ in both cases). Model E represents our final proposal of a 3-dimension structure, where subversive resistance and confrontational resistance are taken as individual dimensions. This model has the highest fit indices with all factors reliable, with convergent validity (all average variance extracted [AVE] higher than .500) and divergent validity (highest heterotrait-monotrait [HTMT] value is .498, and Fornell & Larcker criterion). The items retained in the final factorial solution are depicted in Table 4.2.

Table 4.1 *Model comparison for the Resistance Scales.*

Model	Factor(s)	AVE	CR	Const.	Conf.	Sub.	χ^2 (df), p value	CFI	TLI	RMSEA CI90, PCLOSE	SRMR	$\Delta \chi^2$	ΔCFI
A *1-factor*	Resistance (14)	.270	.685	-	-	-	χ^2 (77) = 1478.242, χ^2/df = 19.198, p < .001	.496	.404	.222 [.213, .232] .000	.221	-	-
B *2-factor*	Constructive (5)	.687	.916	**.829**	-	-	χ^2 (76) = 697.345, χ^2/df = 9.176, p < .001	.776	.732	.149 [.139, .159] .000	.135	$\Delta\chi^{2a}$ (1) = 780.897, p < .001	.280[a]
	Dysfunctional (9)	.358	.823	*.019*	**.599**	-							
C *2-factor*	Constructive (5)	.687	.916	**.829**	-	-	χ^2 (34) = 168.313, χ^2/df = 3.914, p < .001	.940	.923	.089 [.075, .103] .000	.068	$\Delta\chi^{2b}$ (42) = 529.032, p < .001	.164[b]
	Dysfunctional (6)	.437	.812	*-.073*	**.661**	-							
D *2-factor*	Constructive (5)	.687	.916	**.829**	-	-	χ^2 (26) = 89.957, χ^2/df = 3.268, p < .001	.969	.957	.078 [.060, .097] .006	.043	$\Delta\chi^{2c}$ (50) = 607.388, p < .001	.193[c]
	Dysfunctional (4)	.570	.839	*-.057*	**.755**	-							
E *3-factor*	Constructive (5)	.686	.916	**.819**	-	-	χ^2 (51) = 138.796, χ^2/df = 2.721, p < .001	.966	.956	.068 [.055, .082] .014	.044	$\Delta\chi^{2d}$ (25) = 558.549, p < .001	.190[d]
	Confrontational (3)	.671	.858	*.338*** *	**.756**	-							
	Subversive (4)	.571	.839	*-.056*	*.470*** *	**.828**							

Note. Boldface type elements show the square root of the AVE.
The off-diagonal values in italic show the correlations among constructs.
$\Delta \chi^2$ and ΔCFI present values regarding differences between the following models: [a] B-A; [b] C-B; [c] D-B; [d] E-B
*** p < .001
χ^2 – chi-square; df – degrees of freedom; CFI – comparative fit index; TLI – Tucker–Lewis index; RMSEA – root mean square error of approximation; SRMR – standardized root mean square residual; $\Delta\chi^2$ – chi-square difference; ΔCFI – CFI difference

Table 4.2 *Three-factor resistance scale final solution.*

Dimension	Items I would ...	Loading (λ)
Constructive	… ask for additional clarification and explanation.	.689
	… explain that it should be done in a different way.	.888
	… convince my supervisor to reassess whether or not the task is worthwhile.	.871
	… explain that the task will not yield the expected benefits.	.862
	… present logical reasons for doing the task differently or at a different time.	.816
Subversive	… ignore this supervisor.	.610
	… not pay attention to my supervisor.	.728
	… act as if I was never asked to do it.	.770
	… disregard what my supervisor says.	.888
Confrontational	… just say "no."	.781
	… refuse to perform the request.	.931
	… tell this supervisor that I'm not available.	.731

4.5. Brief conclusion of study 1

As a conclusion, these findings offer support to the claim that the dysfunctional resistance, as measured by Tepper et al. (2001) subscale conflates the confrontational resistance with the subversive resistance. However, the mere existence of better psychometric qualities is not sufficient to claim the advantage of differentiating between both constructs. Hence, only through testing the nomological validity of these constructs would it be reasonable to assert the need to consider both types.

Study 2 - Testing the nomological network of confrontational and subversive resistance

In building the nomological network, we departed from the observation that abusive supervision has a twofold opposing effect on resistance. On the one hand, it is already known that it increases its occurrence (e.g., Goswami et al., 2015), but on the other hand, it can also inhibit resistance (e.g., Haggard & Park, 2018). This should draw researchers' attention as it may mean that inhibitory psychological processes can compromise the frontal expression of resistance while, concomitantly, leaving untouched other non-confrontational expressions of resistance. This is instrumental to the objective of this study because such processes may uncover differential relationships with confrontational as against subversive silence.

Accordingly, a key construct in this process that emerges from the akin literature is defensive silence (Van Dyne et al., 2003).

4.6. Literature Review

4.6.1. Defensive silence and abusive supervision

Silence is not the mere opposite of voice (e.g., Detert & Edmondson, 2011; Kish-Gephart et al., 2009; Pinder & Harlos, 2001; Van Dyne et al., 2003). The absence of voice is often motivated by employees having nothing to say, but this behavior is not equivalent to silence (Morrison, 2014). Silence is not only speechless behavior; it may be a conscient and deliberate choice based on a motivation to withhold information (Van Dyne et al., 2003). These authors introduce three motives to withhold information: resignation (acquiescence silence), self-protection-based fear (defensive silence), and other-oriented (prosocial silence). Amongst the typologies of silence, there is always a fear-based type that can be named as defensive (Brinsfield, 2013) or quiescent (Knoll & Van Dick, 2013).

Abusive supervision is commonly assumed to play an important role in predicting defensive silence (Lam & Xu, 2019; Mannan & Kashif, 2019), which is in line with Knoll and Van Dick (2013) findings that silence occurs when employees are victims or witnesses of moral standards breaches in organizations. This, adding to the emotional nature of defensive silence, makes it more critical to organizations as it may emerge as an automatic process (Detert & Edmondson, 2011; Kish-Gephart et al., 2009) within a fundamentally compromised ethical situation (Wang et al., 2020). Accordingly, it is known that abusive supervisors provoke negative emotions at work (Goswami et al., 2015; Simon et al., 2015), which can be helpful in explaining how abusive supervision fosters defensive silence through negative affect activation like fear (Guo et al., 2018; Kiewitz et al., 2016; Kish-Gephart et al., 2009; Milliken et al., 2003). Accordingly, we hypothesize that:

H1. Negative affect mediates the relationship between abusive supervision and defensive silence.

4.6.2. Defensive Silence and Resistance Mechanisms

Contrary to voice, silence motives are easily misattributed (Van Dyne et al., 2003), as no overt behavior is observed. As detailed, defensive silence is a fear-based behavior, and the source of fear stems from the anticipated negative reaction of challenging authority, which may help to explain its prevalence within the organizational context as reviewed by Kish-Gephart et al. (2009). Although defensive silence may be automatically activated (e.g., Detert & Edmondson,

2011), it may lead to volitional and conscious decisions such as the choosing to engage or not in resistance behaviors. From what has been described, we acknowledge the relevance of analyzing how defensive silence relates to both confrontational and subversive silence.

Such opposition to authority can be either unequivocal and assumed (such as confrontational resistance) or covert and silent (such as subversive resistance). When opting for one of these, the resistor needs to have resources to uphold his or her opposition. Conservation of Resources Theory (Hobfoll, 1989) helps to frame this process, where a succession of events is chained by resource depletion. Accordingly, this theory suggests that people manage their resources in order to avoid their depletion and to retain them, especially under stressful situations, such as an abusive supervisor (e.g., Lee et al., 2018; Wang et al., 2020; Whitman et al., 2014). As suggested by Wang et al. (2020) when facing such conditions, employees will reduce proactive behaviors to save resources. Confrontational resistance can be conceived as resource-demanding behavior, as it requires high levels of exposure and proactivity (i.e., active strategies). Accordingly, individuals will refrain in their decision to engage in this behavior when perceiving resource loss and risk. Conversely, subversive resistance can be conceived as less resource-demanding, as the behavior is not exposed (i.e., quiet resistance). As such, it can be taken as an avoidance-oriented coping strategy. Therefore, we hypothesize:

H2. Defensive silence is negatively associated to confrontational resistance but not negatively associated to subversive resistance

As a consequence of the previous hypotheses, it is reasonable to expect that there is a sequential mediation between abusive supervision and resistance via negative affect and defensive silence so that:

H3. The indirect sequential effect is negative towards confrontational resistance but not negative towards subversive resistance.

This reasoning does not take into account that the resources conservation/depletion are also contingent upon power resources, e.g., how much support one gathers from his or her coworkers (Rasheed et al., 2021).

4.6.3. Sense of power as a moderator

Power can be seen as a resource for organizational members (Munduate & Medina, 2017), with similar status of intangible resources as perspective in the Job Demands-Resources Theory (Bakker & Demerouti, 2014). It is known, from the Approach-Inhibition Theory of Power (Keltner et al., 2003), that perceiving power activates the behavioral approach system (e.g., disinhibition) and lack of power is related to the inhibition system, increasing awareness of

risks. All in all, it is already well established that power shapes people's behavior (Guinote, 2017), and, as suggested by previous research (Almeida et al., 2021), the sense of power may have a say when followers decide which resistance strategy should be used. In line, previous works had already informed that covert behaviors are expected when followers perceive feel powerless (Greenbaum et al., 2013; Lian et al., 2014).

For a long time, power was associated with the possession of resources with insufficient focus being given to power as a psychological state that mirrors the individual's perception of his or her influence capability (Anderson et al., 2012). However, this notion of sense of power was found to be an antecedent of silence (Morrison et al., 2015), and voice (Kim et al., 2019), especially when it is conceived as a consequence of the interaction with a leader (e.g., Lin et al., 2019). Nevertheless, power holds many facets; it is not solely linked to the leader-follower interaction, but it also related to peer relations, as illustrated by lateral political behavior (Farrell & Peterson, 1982). This is in line with Anderson et al. (2012) findings showing sense of power is relationship-specific. Accordingly, sense of power can be a critical process variable but can also be taken as a contextual one, especially when analyzed as a protective agent or resource (Van Loo & Rydel, 2013). Moreover, once the position within a context is similar, the main differences in sense of power are related to dispositional/stable features such as personality (Anderson et al., 2012). The relationship with coworkers has been gaining ground, and research shows, for example, that coworker support is a stronger predictor of turnover than supervisor support (Basford & Offermann, 2012). Also, recent research (Rasheed et al., 2021) found that coworker support buffers the negative impact of abusive supervision.

Accordingly, conceiving sense of power in the relationship with coworkers as a resource, we reason that, although silenced by fear, holding higher or lower levels of sense of power will help to determine the impact of silence on resistance behaviors. Thus, we hypothesize:

H4a. Sense of power interacts with defensive silence such that higher levels of sense of power will buffer the negative relationship between defensive silence and confrontational resistance.

H4b. Sense of power interacts with defensive silence such that higher levels of sense of power will have a positive effect on the relationship between defensive silence and subversive resistance, and lower levels of sense of power will have a negative effect on this relationship.

Considering the above, the overall model can be framed in the light of the Job Demands-Resources Theory (Bakker & Demerouti, 2014). Accordingly, the strain and fear (i.e., negative affect and defensive silence) induced by the supervisors' behaviors (i.e., abusive supervisor) operate as demanding factors that have an impact in followers' behavior (i.e., resistance).

However, the individual resource of power perception may interact with this process in such a way that:

H5a. Sense of power moderates the indirect effect of abusive supervision on confrontational resistance, such that the negative indirect effect is weaker among individuals perceiving high sense of power.

H5b. Sense of power moderates the indirect effect of abusive supervision on subversive resistance, such that the indirect effect is positive among individuals perceiving low sense of power and negative among those perceiving high sense of power.

4.7. Method

4.7.1. Participants and procedure

A new sample was collected with an online survey (using Qualtrics) to test the proposed model. Participants were required to be employed and to report to a direct supervisor. The sample comprised 247 employed participants, 79.4% female, aged from 19 to 63 years old ($M_{age}=$ 34.65, $SD_{age} = 10.86$), working with the same supervisor, on average, for 4.57 years ($SD_{dyadic\ tenure} = 5.78$).

4.7.2. Measures

Abusive supervision was measured with the 11-item scale (e.g., "Ridicules me") from Harris et al. (2007), who adapted the original 15-item scale from Tepper (2000). Participants answered on a frequency scale from 1 (never) to 5 (always).

Negative affect in the workplace was measured with the five items (e.g., scared) from the negative affect subscale of PANAS (Galinha et al., 2014). Participants answered on a frequency scale from 1 (never or rarely) to 5 (frequently or extremely frequent).

Defensive silence was assessed with the Van Dyne et al. (2003) five item-scale (e.g., "I do not speak up and suggest ideas for change, based on fear"). The items were tailored to make them self-report instead of other-report. Participants answered on a Likert scale from 1 (strongly disagree) to 6 (strongly agree).

Confrontational and subversive resistance were measured with the new factorial version of Tepper et al. (2001) dysfunctional resistance scale (e.g., "I refuse to perform the request"; "I ignore my supervisor"), developed in Study 1. Participants answered on a frequency scale from 1 (never) to 5 (always).

The personal sense of power in the relationship with colleagues was measured with five

items from Anderson et al. (2012) scale (e.g." I can get them to listen to what I say"). Participants answered on a Likert scale from 1 (strongly disagree) to 6 (strongly agree).

As in the previous study, whenever there was no Portuguese validated version of a scale, Brislin (1970) translation / back-translation procedure was followed.

4.7.3. Control variables

We controlled for age and dyadic tenure as previous research on the topic (Brett et al., 2016; Carnevale et al., 2018; Haggard & Park, 2018; Tepper et al., 2006). As we measure the personal sense of power in the relationship with coworkers, we controlled for the team size. We also controlled for displacement of responsibility as it is an important predictor of deviant behaviors at work (Christian & Ellis, 2014). It was measured with the 4-item scale adapted from Bandura et al. (1996: e.g., "If people are not properly supervised, they should not be blamed for misbehaving"), answered on a 5-point Likert scale (1-strongly disagree; 5 – strongly disagree).

4.7.4. Measurement Model

A confirmatory factor analysis was conducted using AMOS v. 26. The measurement model presented good fit to data $\chi2(474) = 783.548$, $p <.001$, $\chi2/df = 1.653$, CFI = .932, TLI = .924, RMSEA = .052, CI90 [.045; .058] PClose = .342, SRMR = .062 (Hair et al., 2019). The baseline model was compared with alternatives (Table 4.3), which shows it has the best fit, as indicated by $\Delta\chi^2$ and ΔCFI (Cheung & Rensvold, 2002).

Table 4.3 *Models' comparison.*

Model	χ² (*df*), *p* value	CFI	TLI	RMSEA CI90, PCLOSE	SRMR	Δ χ²	ΔCFI
Baseline model	χ² (474) = 783.548, χ²/*df* = 1.653, *p* < .001	.932	.924	.052 [.045, .058] .342	.062	-	-
Model 1 (Abusive and negative affect together)	χ² (479) = 1163.725, χ²/*df* = 2.429, *p*< .001	.849	.833	.076 [.071, .082] .000	.080	Δχ (5) = 380.177, *p* < .001	.083
Model 2 (Negative affect and defensive silence together)	χ² (479) = 1240.244, χ²/*df* = 2.589, *p*< .001	.832	.814	.080 [.075, .086] .000	.080	Δχ (5) = 456.696, *p* < .001	.100
Model 3 (Defensive silence and confrontational resistance together)	χ² (479) = 1013.410, χ²/*df* = 2.116, *p*< .001	.882	.870	.067 [.062, .073] .000	.089	Δχ (5) = 229,862, *p* < .001	.050
Model 4 (Defensive silence and subversive resistance together)	χ² (479) = 1147.366, χ²/*df* = 2.395, *p*< .001	.852	.837	.075 [.070, .081] .000	.104	Δχ (5) = 363.818, *p* < .001	.080
Model 5 (Confrontational and subversive resistance together)	χ² (479) = 909.094, χ²/*df* = 1.898, *p*< .001	.905	.895	.060 [.054, .066] .003	.067	Δχ (5) = 125.546, *p* < .001	.027
Model 6 (Single factor)	χ² (489) = 2500.306, χ²/*df* = 5.113, *p*< .001	.555	.520	.129 [.124, .134] .000	.134	Δχ (15) = 1716.758, *p* < .001	.377
Model 7 (Independent factors)	χ² (489) = 1012.827, χ²/*df* = 2.071, *p*< .001	.884	.875	.066 [.060, .072] .000	.171	Δχ (15) = 229.279, *p* < .001	.048

Note. χ² – chi-square; df – degrees of freedom; CFI – comparative fit index; TLI – Tucker–Lewis index; RMSEA – root mean square error of approximation; SRMR – standardized root mean square residual; Δχ² – chi-square difference; ΔCFI – CFI difference

4.8. Results

4.8.1. Preliminary analyses

The descriptive statistics and correlations are depicted in Table 4.4.

Table 4.4 *Descriptive statistics and correlations.*

	Mean	SD	1	2	3	4	5	6	7	8	9
1.Abusive supervision	1.42	0.60	(.937)								
2.Negative affect	1.61	0.73	.569**	(.855)							
3.Defensive silence	2.02	0.95	.327**	.426**	(.879)						
4.Confrontational resistance	1.57	0.61	.225*	.087	-.099	(.760)					
5.Subversive resistance	1.30	0.44	.359**	.129	.079	.494**	(.800)				
6.Sense of power (colleagues)	4.50	0.85	-.073	-.200	-.277*	-.067	-.248*	(.808)			
7.Moral Disengagement	2.20	0.75	.116*	.134	.232*	.130*	.128	-.213**	(.805)		
8.Age	34.65	10.86	-016	-.055	-116	.053	.155	-.128	-.090		
9.Dyadic Tenure	4.57	5.78	-.001	-.034	-.052	.113	.026	-.059	.042	.482*	
10. Team size	17	34.56	-.034	.007	.034	-.035	-.036	.043	-.053	.213**	.090*

Note. $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$, Cronbach alpha between parentheses in diagonal.

4.8.2. Structural Model

The hypothesized moderated mediation presented an acceptable fit to data ($\chi2(615) = 984.990$, $p < .001$, $\chi2/df = 1.602$, CFI = .921, TLI = .909, RMSEA = .049, CI90 [.044; .055] PClose = .556, SRMR = .0684). Results show a significant indirect effect from abusive supervision to defensive silence, via negative affect ($\beta = .184$, $p = .004$, 95% CI = .059; .401), and no direct effect ($\beta = .126$, $p = .262$, 95% CI = -.093; .349), suggesting a full mediation and supporting H1. Defensive silence is negatively associated to confrontational resistance ($\beta = -.239$, $p = .005$, 95% CI = -.419; -.081), and no association was found between defensive silence and subversive resistance ($\beta = -.078$, $p = .380$, 95% CI = -.248; .093), supporting H2. Abusive supervision is indirectly and negatively related to confrontational resistance, sequentially mediated through first negative affect and then defensive silence ($\beta = -.047$, $p = .004$, 95% CI = -.156; -.011), and a positive direct effect was found between abusive supervision and confrontational resistance ($\beta = .319$, $p = .014$, 95% CI = .074; .592) suggesting a partial mediation. No sequential mediation was found between abusive supervision and subversive resistance ($\beta = -.011$, $p = .209$, 95% CI = -.051; .007), and a direct effect was found between abusive supervision and

74

subversive resistance (β = .494, p = .001, 95% CI = .243; .732); thus, H3 was partially supported. The sense of power moderates the relationship between defensive silence and subversive resistance (β = -.224, p = .017, 95% CI = -.332; -.049). The relationship is positive when there are low levels of sense of power (β = .255, p = .039, 95% CI = .018; .516) and negative when there are high levels of sense of power (β = -.322, p = .011, 95% CI = -.571; -.081). These findings support H4a. No moderation effect of sense of power was found in the relationship between defensive silence and confrontational resistance (β = -.091, p = .173, 95% CI = -.275; .052), thus no support was found for H4b.

A significant index of moderated mediation (β = -.036, p = .010, 95% CI = -.120; -.006) showed a conditional indirect effect between abusive supervision and subversive resistance for both low (β = .081, p = .022, 95% CI = .009; 270) and high (β = -.103, p = .007, 95% CI = -326; -.022) levels of sense of power, showing a positive effect of abusive supervision on subversive resistance through negative affect and interaction of defensive silence with low sense of power, and a negative effect when high levels of sense of power were presented, thus supporting H5b.

4.9. Discussion

Once we accept leadership as a process where followers play an active role, it becomes clear that followers also have a say in destructive leadership (Padilla et al., 2007; Thoroughgood et al., 2018). Although mainstream research assumed followers' passivity and has been focusing on the consequences of having a destructive leader (e.g., Schyns & Schilling, 2013; Park et al., 2018; Peng et al., 2019), recent research on followership is focusing also on behaviors designed to curb destructiveness (Almeida et al., 2021). Accordingly, the present work was designed to extend previous knowledge on resistance behaviors in the line of this emerging trend.

While it is widely accepted that resistance may emerge due to the abuse by a leader (e.g., Schyns & Schilling, 2013), we argue that there is an important gap when considering dysfunctional resistance, as it blends different behaviors. First, departing from Tepper et al. (2001) scale, we proposed a three-dimension measure, where dysfunctional resistance splits in two: confrontational, and subversive. Then, we tested the nomological of this split proposal. A paradoxical reality drove us: while there is evidence of increasing dysfunctional resistance when followers face destructive leaders (e.g., Goswami et al., 2015), it is reasonable to expect that abusive leaders also frighten followers, thus reducing their capacity to display resistance. Then, it becomes inevitable to ask: is there any mechanism operating in parallel that explains a

decrease in resistance expression? If so, we reasoned that it would be explained by the increased depletion of resources, which would reduce active expressions of resistance (i.e., confrontational resistance). This may not be the case for avoidant strategies (i.e., low resource demanding), like subversive resistance.

Overall, the results from the first study support a three-factor resistance scale where the dysfunctional dimension gives place to confrontational and subversive resistance. Adding to the psychometric evidence, a second study finds a differential path in explaining the new resistance types, corroborating the relevance of that dimension. This section provides the findings' interpretation while highlighting some conceptual and practical contributions.

The first part of the present research acknowledges the relevance and impact of Tepper et al. (2001) proposal to assess resistance behaviors while identifying some potential for conceptual advancement. Thus, the cumulative evidence of facial validity issues, plus empirical studies not using the whole scale (e.g., Vecchio et al., 2010), plus previous suggestions of a covert resistance style (e.g., Almeida et al., 2021; Collinson, 2006; Rodrigues & Collinson, 1995, led us to test a three-factor solution. In Tepper et al. (2001) scale, constructive resistance is already consolidated with all the items describing the same behavioral pattern (i.e., negotiation-based) in the same direction. Our attention thus fell upon dysfunctional resistance which, to our understanding combines confrontational resistance and subversive resistance. Our findings supported this prediction which led us to propose a restructured resistance scale. Accordingly, our first study contributes to research on resistance mechanisms, drawing scholars' attention to a behavioral pattern that goes beyond labeling a behavior as constructive or dysfunctional. In other words, resistance displayed against an abusive leader may be more active (i.e., overt) or passive (i.e., covert). This sort of classification has already been described and empirically supported in recent research (Almeida et al., 2021). These results also seem to echo research on coping strategies (e.g., Lazarus & Folkman, 1984) that build on a framework that classifies stress management behaviors as holding an avoidance or approach nature (May et al., 2014; Webster et al., 2016). In light of this, both constructive and confrontational resistance would be taken as approach strategies, while subversive would be considered as an avoidance-based behavior.

Splitting dysfunctional resistance into two different strategies impelled us to test whether the path explaining them was different. Drawing on both Conservation of Resources (Hobfoll, 1989) and Job Demands-Resources theories (Bakker & Demerouti, 2014), we designed and tested hypotheses from a moderated sequential mediation model. In this chained process, defensive silence (Van Dyne et al., 2003) gains particular relevance in the sense that although

it can be automatically generated (Detert & Edmondson, 2011), it can foster different resistance decisions.

As expected, we found an affective route in explaining how abusive leaders foster defensive silence. This is in line with previous research anchoring this kind of silence in negative emotions activation (Kish-Gephart et al., 2009). Additionally, findings regarding the sequential mediation enable us to highlight the following ideas. First, when explained by defensive silence, abusive supervision lowers confrontational resistance levels. This is in line with the idea that a chained resource depletion process, induced by abusive behaviors, drives followers to protect the remaining resources by avoiding exposure to risky situations or behaviors. At the same time, one should note that although there is a negative indirect effect, the direct path is still significant ($B = .322$, $p = .006$, 95% CI = .125; .621), meaning that, overall, the mediation is only partial. Moreover, due to the opposite directions of the direct and indirect effects, one could expect to have no total effect; however, this does not happen, and the effect is positive ($B = .239$, $p = .017$, 95% CI = .067; .453). It is possible to interpret these results as others did (e.g., Goswami et al., 2015), where confrontational resistance could emerge as a source of retaliation, which is known to be a mechanism intended to restore justice (Liang et al., 2018). Another reason for this to happen can be related to the low levels of abusive supervision in our sample. Although this value is in line with other research in the field (Harvey et al., 2007; Tepper, 2000), we should acknowledge this issue which was already highlighted previously (Tepper et al., 2017).

Secondly, with regards to subversive resistance, our results show no association with defensive silence, i.e., no sequential mediation. This finding is quite interesting as it suggests that silence does not necessarily mean absence of resistance, which is in line with the idea that one can decide and behave silently. However, these results themselves limit their own interpretation, which makes the role of the moderator variable critical. We found a moderated sequential mediation explaining subversive resistance, which means that external factors do make a difference in this process. Specifically, we found that individuals that perceived themselves as having influence power in the relationship with their coworkers experience lesser need to engage in covert strategies. Conversely, when perceiving low levels of peer-directed power, the opposite applies. That is, our findings suggest that when explained through defensive silence, subversive resistance emerges when one has no resources. It will be taken as a last resort behavior. Overall, because sense of power relates to the individual's perception of being able to influence others (Anderson et al., 2012), then high levels of sense of power can leverage the use of collective strategies such as coalitions (Farrell & Peterson, 1982; Yukl & Tracey, 1992). This may explain the decrease in the use of an individual and protective strategy such as

subversive resistance. Conversely, no moderation was found when explaining confrontational resistance leading us to reason that an active and risky strategy such as engaging in confront does not depend on external power resources. In other words, this kind of external resource (ability to influence peers) is not enough to fight the depletion of the process-related resources.

4.9.1. Limitations and future research

Findings must be gauged against the study's limitations. The mediation process was tested with variables collected from the same source and at the same time, which can raise issues related to the common method bias (Podsakoff et al., 2003, 2012). However, there is some evidence that this may not be a concern. Accordingly, we tested a single factor solution, which did not fit the data. Moreover, additional evidence that helps to rule out the common-method bias comes from the interaction effect which is not expected whenever common-method variance occurs (Siemsen et al., 2010). Another possible limitation relates to the sample. As discussed, we need to be cautious when interpreting models that inquire the general working population about abusive supervision. This happens because although some studies present data suggesting an important prevalence (e.g., Aasland et al., 2010), when answering scales of abuse, the means are generally low (e.g., Harvey et al., 2007; Tepper, 2000). Then it would be important to test if this model works in the same way when higher levels of abuse occur. An experimental methodology might help to address this.

Future research should focus on extending this model, or a similar one, to understand the consequences (for the leader, the follower, and the organization) of different resistance strategies. Although there is some research on the impact on the consequences of resistance (Tepper et al., 2006), these studies are scarce. Moreover, it might be more informative to test and compare the three dimensions of resistance or even explore subdimensions. Also, since we found that some individual external resources affect and modify the path that explains resistance, then and according to the Job Demands-Resources Theory (Bakker & Demerouti, 2014), it can be useful to extend research on this kind of resources to employ strengths-based interventions related.

Future research could investigate our proposition that higher levels of lateral sense of power reduce the need to engage in subversive resistance due to the adoption of collective strategies. In this research, those holding power decrease the usage of subversive resistance; however, other sorts of resources may enhance active strategies like confrontational resistance. Although these kinds of resistance were previously classified as dysfunctional (Tepper et al., 2001), we trust that it can be important to make them happen when facing a destructive leader and thus its

"dysfunctional" nature may be questioned.

Finally, the differential findings for both paths throw light to the nature of covert resistance, such as the subversive. As expected, covert behaviors emerge when followers feel they do not hold enough power-related resources (e.g., Greenbaum et al., 2013). Thus, apparently, to behave subversively can be conceived as a protective strategy, making it a last resort rather than a willingly first choice. Although we were expecting these results from a power-based perspective, they shed light on the need to further extend research on covert behaviors. In other words, are covert behaviors always driven by fear? Can they be strategically defined, being related to behaviors such as manipulation? (Fairholm, 2009) Or does the strategy emerge from a perceived lack of alternatives, mirroring the idea that *necessity is the mother of invention*? It might be relevant for future research to address this topic and detail subversive mechanisms and their drivers in a similar way to what researchers have been done to uncover the true nature of organizational silence (Brinsfield, 2013; Knoll & Van Dick, 2013; Pinder & Harlos, 2001; Van Dyne et al., 2003).

4.10. Conclusion

Once we conceive followers as inherent to the leadership process, we should explore and understand resistance strategies, especially when considering destructive leadership. Very recent research suggests that when facing a destructive leader, a real follower may be precisely the one who does not follow the leader (Almeida et al., 2021). This is to say that resisting a destructive leader is critical in curbing the destructive leadership process. Thus, this field of research benefits from scholars' attention to the study of the different resistance strategies. In this way, the present work sets out to detail these behaviors and proposed three instead of two resistance types. Specifically, regarding dysfunctional resistance, this dimension holds two different natures: overt and covert.

The second part of this study draws attention to the need to listen to silence, a pervasive and sometimes automatically generated problem in the organizational setting (Kish-Gephart et al., 2009). So, researchers and practitioners should acknowledge that silence is not necessarily equivalent to doing nothing. Individuals can decide what to do in silence, and although it is reasonable to expect confrontation to reduce, this might not be the case when a subversive intent is present. In that case, it will depend on the resources the individual has. When holding power of influence over others, an individual will not feel the need to act subversively, but if no power resources are available, then subversion emerges as a backup strategy.

Complying with the supervisor's unethical request: The moderation effect of AI vs. human supervisor[7].

Abstract

Subordinates' compliance with unethical requests provided by non-human agents is an underdeveloped topic in the artificial intelligence and business ethics research field. Drawing on social cognitive theory and identity-based motivation theory, we designed a two-wave study to address this issue. Two-hundred and eight participants answered about moral-based variables – moral identity centrality and displacement of responsibility – before being exposed to a set of scenarios to test whether they would align with fraud. Participants were randomly assigned to one of two conditions (1) a human supervisor performs the request or (2) the request is provided by an algorithm. Our findings show that displacement of responsibility mediates the relationship between moral identity centrality and intention to comply with fraud. However, contrary to expectations, the mediation was not found in the algorithm supervisor condition, where displacement of responsibility did not predict unethical behavior. Implications are discussed, after which opportunities for future research are explored.

Keywords: supervisor's unethical request, artificial intelligence, displacement of responsibility, moral identity, accountability

---

[7] This chapter has been submitted to an international indexed journal.

5.1. Introduction

Artificial intelligence (AI) already established itself in organizations (Makarius et al., 2020; von Krogh, 2018), and its continuous development and increasing complexity bring important challenges to business ethics. Artificial intelligence, defined by Kaplan and Haenlein (2019, p. 17) as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation", is not a simple tool anymore. It now takes part in the process and needs to be considered a teammate or even a more complex agent that already makes decisions (Wesche & Sonderegger, 2019). This phenomenon raises many questions and ethical concerns (Munoko et al., 2020; Tambe et al., 2019; Wright & Schultz, 2018). Are organizations prepared to face AI systems mistakes with ethical implications? Is there enough knowledge on how subordinates will manage AI systems requests? Specifically, how will people behave when faced with an unethical request provided by a non-human? The last question is the starting point of this paper.

Organizations face several ethical threats, and there is abounding research on it (Kish-Gephart et al., 2010; Treviño et al., 2014). Minor unethical situations tend to build up to collapse into ethical scandals often associated with fraud, such as Enron and WorldCom (van Driel, 2019). When such scandals happen, the focus is placed on leading positions (Cohen et al., 2010). However, employees are often important players in unethical dynamics (Beu & Buckley, 2004; Scharff, 2005), particularly when they follow unethical requests (Johnson et al., 2019). Consequently, it becomes relevant to uncover which subordinates' characteristics make them more vulnerable to defer to such demands. Accordingly, it is known, and well-established, that individual moral-based differences, such as moral identity and moral disengagement, matter (Detert et al., 2008; Reckers & Samuelson, 2016; Shao et al., 2008).

Artificial intelligence adds complexity to the analysis of ethical risks in organizations. With the expansion of the fourth industrial revolution (Schwab, 2017), it is pivotal to go back to basics and integrate previous knowledge on human behavior in the AI reality. Non-human agents are perceived as more rational and less emotional (Lee, 2018), which would suggest lower levels of unethical decisions. However, there are cases of biased decisions with important ethical consequences, like gender and racial biases (Kellogg et al., 2020; Martin, 2019; Meyer, 2018). Thus, these systems are not value-free and reflect their designing process (Mittelstadt et al., 2016), allowing us to argue that non-human agents can indeed make unethical decisions.

All the research that has been studying employees' alignment with unethical behaviors in organizations departs from the same underlying assumption: employees react to human

supervisors. The technological breakthroughs are changing the possibilities and challenging this assumption. Does previous knowledge on unethical behavior routes (e.g., moral-based differences) help to understand unethical practices, such as following unethical requests from a non-human agent? In human-human interaction, employees' compliance with supervisors' unethical requests are enabled by displacing responsibility upwards (Bandura, 2016). This well-known cognitive mechanism assumes that when someone follows an unethical request in the organizational context, it can be considered a crime of obedience (Carsten & Uhl-Bien, 2013; Kelman & Hamilton, 1989). When this happens, employees with higher levels of propensity to disengage will displace their responsibility into a senior representative (Bandura, 2016). An implicit assumption underlying the link between displacement of responsibility and unethical behavior is that the person can assign responsibility to another agent. Is this likely to occur when the request comes from an AI system? This question links with the debate around AI and accountability (Orr & Davis, 2020; van de Poel, 2020), where we agree that responsibility matters and exists even if it is not in the system itself (Martin, 2019).

This study aims to test how moral disengagement, through displacement of responsibility, predicts intention to comply with fraud when a human supervisor provides the request compared to when it comes from a non-human system (an algorithm). Answering this goal will develop knowledge on human-AI agents' relations, specifically when ethical issues are at stake. Additionally, it will add to the debate on accountability and AI. All in all, our study will also help to inform organizations about AI-based non-ethical decisions threats.

This paper proceeds as follows. First, we introduce literature on the processes that activate unethical behaviors in organizations, focusing on moral disengagement, how it develops, its link to unethical practices, and its mediator role (through the displacement of responsibility mechanism). Then, we propose the supervisor nature (i.e., human or AI-based) as a moderator in the process mentioned above. Next, we explain our research method. This is followed by the results section that presents findings pertaining to the mediational process as well as the moderated mediation and precedes the discussion. After presenting the study's implications, we conclude by putting forward some limitations and future research opportunities.

5.2. Literature Review

5.2.1. Moral disengagement

With the growing number of ethical scandals (Newman et al., 2019) and recognition of subordinates' involvement in such episodes (Beu & Buckley, 2004; Scharff, 2005), it has become critical to ask why people behave unethically in organizations? Research provides multiple answers to this question (e.g., Kish-Gephart et al., 2010; Moore & Gino, 2013). According to Treviño et al. (2014) and De Cremer and Moore (2020), factors that influence (un)ethical behaviors in organizations can be placed into different levels. Organizational factors have been mainly studied through the implication of ethical climates (e.g., Peterson, 2002). There are also interpersonal influences such as the leader (e.g., Beu & Buckley, 2004; Moore, et al., 2019; Valle et al., 2019), and peers (e.g., Moore & Gino, 2013). The last category of reasons is placed at the individual level and comprehends stable individual differences (e.g., Johnson et al., 2019), cognitive (e.g., Detert et al., 2008), and affective (e.g., Kouchaki & Desai, 2015) mechanisms. Within the cognitive processes, moral disengagement has been receiving growing interest (e.g., Newman et al., 2019).

Moral disengagement is a cognitive process introduced by Bandura, et al. (1996), as a derivation of his social cognitive theory (Bandura, 1991), to illustrate how individuals can detach from their moral standards to behave unethically without the resulting guilt or strain. This mechanism is usually activated to self-allow unethical behaviors such as social undermining (Duffy et al., 2012), workplace deviance (Valle et al., 2019), and following unethical requests (Carsten & Uhl-Bien, 2013).

Bandura (2016) proposed eight moral disengagement mechanisms that deactivate the moral self-regulatory process at four different *loci*, namely, behavior, agency, outcome, and victim. Behavior locus allows a cognitive restructuring of the behavior so that it is not assessed as immoral and comprises moral justification, euphemistic labeling, and advantageous comparison. Agency locus enables an agency relocation so that the individual feels less responsible for the immoral conduct and comprehends displacement of responsibility and diffusion of responsibility. Outcome locus changes the way consequences are analyzed so that the action is not perceived as immoral and includes disregard or distortion of consequences. Victim locus pushes the responsibility for the victims themselves and comprises dehumanization and attribution of blame.

Once moral disengagement has been established as an important precursor of unethical behaviors (Moore, 2015; Newman et al., 2019), it becomes critical to understand where it comes from and how it develops. Research has uncovered different routes to moral disengagement, and while some show evidence for contextual influences (Astrove et al., 2015; Huang et al., 2017), others build on the idea of an individual difference (Detert et al., 2008; Knoll et al., 2016; Moore et al., 2012). Conceiving moral disengagement as an individual difference means that some people are more susceptible to deactivated moral standards and engage in unethical behaviors. This assumption opened the way for studies focusing on characteristics that make people more (or less) predisposed to morally disengage. Empirical evidence shows a set of features that influence this cognitive process, namely: empathy, trait cynicism (Detert et al., 2018), honesty-humility (Ogunfowora & Bourdage, 2014), and moral identity (Chowdhury & Fernando, 2014; Detert et al., 2008; Harris & He, 2019). From these predictors, moral identity has been the most frequently reported, consistently showing associations with moral disengagement.

5.2.2. Moral identity

Moral identity refers to how individuals think about themselves regarding moral issues. Accordingly, Aquino and Reed (2002, p. 1424) define it as "as a self-conception organized around a set of moral traits" and position their conceptualization roots on social identity theory (Tajfel & Turner, 1979). Therefore, moral identity works like a *schema*, which makes people behave according to how they view themselves (McFerran et al., 2010). It is conceived as being relatively stable in time (Aquino & Reed, 2002), like other moral standards (Beu & Buckley, 2004), making moral identity a trait-based concept (Reed II et al., 2016). Aquino and Reed (2002) propose that moral identity's self-importance is critical for individuals' self-definitions and is associated with beliefs, attitudes, and behaviors. This means that what determines moral reasoning and behaviors is not the level of moral identity itself but its centrality in the individual self-perception. This is in line with Blasi (1994) proposal of moral identity as an individual difference that informs about the centrality of being moral. Accordingly, those who feel a strong moral identity will behave correspondingly in order to keep self-consistency. Thus, the level of moral identity and its centrality for the sense of self plays a critical role in how moral reasoning determines moral behaviors (Aquino & Freeman, 2009). Research supports this idea as moral identity was found to relate both to prosocial (Aquino & Reed, 2002) and unethical behavior (Shao et al., 2008; Wowra, 2007). Alongside the research of antecedents of moral

disengagement, much research has focused on its consequences.

5.2.3. Moral disengagement and unethical behavior

Moral disengagement, a cognitive mechanism that helps to suppress moral consequences (i.e., self-censure), is an important predictor of unethical behaviors (e.g., Detert et al., 2008). Moral disengagement enables unethical behaviors across different settings (e.g., Chowdhury & Fernando, 2014; Detert et al., 2008; Newman et al., 2019). However, there are some specific for the organizational context, such as turnover intentions (Huang et al., 2017), counterproductive work behaviors (Fida et al., 2015; Hystad et al., 2014; Seriki et al., 2020), social undermining (Duffy et al., 2012), workplace deviance (Valle et al., 2019), employee pilferage (Harris & He, 2019), and following unethical requests (Carsten & Uhl-Bien, 2013; Reckers & Samuelson, 2016).

Ethical scandals have been widely reported (Newman et al., 2019), and employees may play an important role in these scandals (e.g., Johnson et al., 2019). These situations correspond to "crimes of obedience" (Carsten & Uhl-Bien, 2013; Kelman & Hamilton, 1989), where individuals excuse their behavior, claiming they were just following orders. When faced with these requests, subordinates are taking part in the process and need to decide whether or not to obey (Beu & Buckley, 2004). Although they can feel a strong pressure to comply, it is known that some do resist (Carsten & Uhl-Bien, 2013). Moral disengagement plays a critical role in explaining this phenomenon (e.g., Reckers & Samuelson, 2016).

As mentioned, moral disengagement mechanisms take action at different moments, and their manifestation depends on the specificity of the context (Bandura, 2016, 2018). Mechanisms that relocate responsibility are critical when studying unethical behaviors within the organizational context (Seriki et al., 2020), and from those, the most studied is displacement of responsibility. Displacement of responsibility relocates responsibility to an authority figure (Bandura, 2016) and is of particular interest within the organizational context due to its hierarchical structure (Hinrichs et al., 2012; Newman et al., 2019; Reckers & Samuelson, 2016). A neutralization mechanism similar to displacement of responsibility, denial of responsibility (Ashforth & Anand 2003), has also been linked to organizational fraud (Anand et al., 2004; Fleming et al., 2020; Zyglidopoulos et al., 2009). Accordingly, when people deactivate moral regulation and do not feel accountable for behavior (displacing it to another agent), they will be more likely to engage in unethical behavior (Barsky, 2011).

5.2.4. The mediator role of moral disengagement

Moral identity is related to a principled ideology (McFerran et al., 2010; Wowra, 2007), meaning that the conduct is guided by moral principles (Schlenker, 2008). Accordingly, research found that this cognitive *schema* is positively associated with prosocial behaviors and negatively related to unethical actions (Aquino & Reed, 2002; Shao et al., 2008), although it is considered a weak effect (Moore et al., 2012). Hertz and Krettenauer (2016) conducted a meta-analysis on the relation between moral identity and (un)ethical behavior, and their findings suggest only small to moderate effect sizes. The authors conclude that moral identity is an important variable that should be studied integrating other variables (e.g., mediators) in models designed to explain (un)ethical behaviors. Indeed, Detert et al. (2008) have already found a full mediation played by moral disengagement in bridging moral identity with unethical behavior.

This mechanism can be supported by identity-based motivation theory (Oyserman, 2007) that states that when an identity is central to an individual, they will be more influenced by it. Drawing on this theory, we share the perspective of "Identity as a source of moral motivation" (Hardy & Carlo, 2005). Accordingly, we argue that people with higher levels of moral identity centrality will express lower compliance levels with fraud because their moral standards will make them feel more accountable for their actions, reducing levels of displacement of responsibility (Schlenker et al., 2009). Therefore, we hypothesize that:

H1. Displacement of responsibility will mediate the negative relationship between moral identity and compliance with fraud.

5.2.5. The nature of the supervisor as a moderator

In previous subsections, we have illustrated how trait-based and cognitive personal characteristics (i.e., moral identity and moral disengagement) explain (un)ethical behavior. However, from social cognitive theory (Bandura, 1991, 2016), it is known that context affects cognition's influence on behavior, and some studies empirically support this moderating effect (e.g., Christian & Ellis, 2014; Knoll et al., 2016). Context is a critical dimension in the present study because the organizational setting is now facing the reality of artificial intelligence-based decision making (Larson & DeChurch, 2020), showing the impact of the fourth industrial revolution (Schwab, 2017), where technology gains agency and control (Glikson & Woolley, 2020).

5.2.5.1. AI - from tool to decision-maker

When thinking about artificial intelligence in organizations, Malone (2018) states that AI systems can play different roles in their work-relation with humans. This is in accordance with PwC (2017) report, where AI systems are classified into four categories. They can: support human decision-making (assisted AI-systems); co-decide with humans (augmented AI systems); perform alone in existing tasks (automated AI-systems); and take the sole responsibility for decision-making, thus being able to adapt to context (autonomous AI).

Briefly, with technological development, human-computer interactions have increased. These systems have started by being a tool, then teammates, and now they are taking some leaders' tasks (Wesche & Sonderegger, 2019). Among the advantages of having computers in the lead are time-saving and better information processing (Chamorro-Premuzic & Ahmetoglu, 2016; Munoko et al., 2020). Moreover, Kellogg et al. (2020) describe some mechanisms through which employers can use algorithms in some management tasks (e.g., recommending, replacing). Tambe et al. (2019) also provide examples of how AI systems may help in functions such as hiring employees and may contribute to predict unwanted behaviors such as turnover. All in all, these are important contributions since managers spend 54% of their time dealing with administrative coordination and control (Kolbjørnsrud et al., 2016). More sophisticated AI use can be found with iCEO, a prototype software, that manages tasks to assign to workers, and is also able to hire employees, among other operations (Fidler, 2015). According to Wesche and Sonderegger (2019), Uber is another example of monitoring or task distribution functions by an automated system. Although these are mainly automated systems, AI is making even more complex decisions (Tokic, 2018), giving it the status of emerging technology (Munoko et al., 2020) despite more than 60 years have elapsed since its official introduction (Haenlein & Kaplan, 2019).

The emerging nature and fast development of AI bring important levels of ambiguity and uncertainty (Rotolo et al., 2015) that raise ethical concerns (e.g., Larson & DeChurch, 2020; Munoko et al., 2020; Tambe et al., 2019; Wright & Schultz, 2018). Technology-related ethical concerns have been focusing on its ethical use (Munoko et al., 2020) and its ethical implications to work (Wright & Schultz, 2018). It is yet unknown how our moral standards relate to such technology when the technology itself has built-in capacities to elicit unethical behaviors. This becomes even more relevant because questions emerge in a moment when issues are already a pressing matter and a reality implemented in many organizations. One of these pertains to which decision-making tasks can be delegated to non-human intelligent agents (von Krogh, 2018).

AI systems as decision-makers are gaining leading functions. This hastens the need to develop knowledge on computer-humans relationships based on already established human-human leadership models (Wesche & Sonderegger, 2019). Earlier Ötting and Maier (2018) had conducted a study to analyze the effect of procedural justice in a set of behaviors and attitudes depending on the nature of the agent (human, robot, or computer), and no moderation effect was found. This is somehow surprising because the decision-maker's nature influences the perception of the decision, regardless of the content of the decision itself (Sundar & Nass, 2001). However, Ötting and Maier (2018) findings imply that some fundamental psychosocial processes related to judgments about fairness or even ethics may transfer to AI - human interaction.

5.2.5.2. AI as context: Diverging relations?

The relation between individuals and AI machines in the work context have been the topic of some studies (e.g., Seeber et al., 2020). More specifically, for the purpose of this study, it is pertinent to revisit knowledge of how individuals relate to non-human decision-making agents. Yeomans et al. (2019) show that people do not rely on algorithm-based recommender systems, although these systems provide more accurate preference predictions than humans (also considering family or friends) and even for subjective issues. Kleinberg et al. (2018) found, in a study that compared bail decisions made by judges with those predicted with machine learning, that algorithm performs more accurately than humans. The authors provide evidence on the impact of algorithms accurateness that can result in crime decreasing up to almost 25% in such individuals. Despite some evidence of the algorithm's higher accuracy in subjective judgments, individuals tend to mistrust algorithm-based decisions. This phenomenon was addressed by Dietvorst et al. (2015) that proposed the term algorithm aversion. From a sequence of experiments, the authors showed that individuals trust more in humans even when they fail more than algorithms. According to the authors, these findings are explained by the lower tolerance levels individuals have to algorithms' mistakes.

Although Prahl and Van Swol (2017) study did not confirm participants use more human advice than algorithm advice, it suggested that bad advice enacts a stronger negative reaction against the advising agent when it comes from automated systems. This was attributed to humans expecting automation forecast to be perfect. While analyzing other researchers' findings, Lee (2018) suggests that algorithm decision-makers are perceived as being less emotional and intentional and more rational than humans. Larson and DeChurch (2020) describe that some studies show that humans prefer to transfer leadership responsibilities to

robots than to other human teammates, and Cormier et al. (2013) findings show that non-human identities can be taken as authority figures that elicit obedience.

From what has been described, the way individuals establish relations with non-human agents is associated with beliefs regarding their attributed levels of rationality and expectations of accuracy. These attributions of rationality and accuracy intrinsically link to accountability for the decisions. This topic has been deserving much attention from scholars (e.g., Martin, 2019; Nath & Sahu, 2020; Orr & Davis, 2020; van de Poel, 2020). These issues gain a critical role in the overall AI debate because, despite these systems are conceived as more rational and less biased, sometimes they fail, and the failure may entail ethical implications (Martin, 2019; Tambe et al., 2019). When this happens, the "who is responsible for this?" question automatically emerges. Orr and Davis (2020, p. 719) show that AI practitioners "distribute ethical responsibility across a range of actors and factors, reserving a portion of responsibility for themselves, albeit constrained". Hohenstein and Jung (2020) found that machines are assigned responsibility for an outcome when things go wrong but not when everything goes well. It also seems that sometimes, decisions that bring negative consequences for oneself are easier to accept when displayed by an algorithm (Tambe et al., 2019). From what has been described, it is clear the need to clarify these processes.

A conceptual debate has been developing to clarify the moral nature of non-human agents. Martin (2019) argues that since non-human decisions bring ethical consequences, these agents are value-laden, and consequently, there is responsibility for their decisions. According to the author, those who develop the system should be accountable for it. Alternatively, those who operate the system are held responsible such as previewed in the Montreal Declaration for a Responsible Development of Artificial Intelligence (2017) and the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (2019). Accountability is also one of the seven requirements for trustworthy AI as stated in the EU Ethics Guidelines for Trustworthy Artificial Intelligence (EU High-Level Expert Group on AI, 2019).

Although there is a lack of consensus regarding accountability for non-human decisions, its need is undeniable (e.g., Orr & Davis, 2020), particularly when ethical issues are implicated (Martin, 2019). Thus, despite knowing that it would be challenging to assign responsibility to a system that lacks moral agency (Nath & Sahu, 2020), in a situation where there is no human operator to blame, either the individual holds the AI system responsible or transfers that responsibility to supervisory entities, e.g., developers or those in charge of guaranteeing the ethical use of the systems. Accordingly, we hypothesize:

H2. The supervisor's nature moderates the indirect effect of moral identity on intention to comply with fraud such that this effect is weaker when the supervisor is an algorithm.

5.3. Method

5.3.1. Participants

The proposed hypotheses were tested in a two-wave experimental setting, based on Johnson et al. (2019) scenarios. In the first wave, 300 participants were recruited from Amazon Mechanical Turk. This crowdsourcing tool has been widely used (e.g., Johnson & Umphress, 2019; O'Reilly III et al., 2018) and is a better alternative than student samples for organizational researchers (Behrend et al., 2011). Four weeks later, 227 individuals answered the second wave. We excluded 19 applicants due to attention issues (Goldammer et al., 2020). Thus, the final sample comprehends 208 participants ($M_{age} = 40.56$, $SD_{age} = 10.75$; 52.4% male).

Participation in the first wave was compensated with $.65, and the second wave was rewarded with $1.75. Due to our design and experimental task specificity, we recruited participants with high education levels (bachelor level as a minimum) and from business-related industries.

5.3.2. Procedure

Informed consent was presented when participants opened the survey's link. The materials were presented in the following order. In the first wave, individuals answered the moral identity scale and demographic variables. Then, participants were randomly assigned to a human supervisor or logarithmic supervisor (AI condition).

Participants in the human condition were provided with a brief description of an organization, its new CEO, and participants were required to take the role of a division manager, following Johnson et al. (2019) indications. For the AI condition, a new description (and scenarios) was developed, based on the original ones. We kept the names of the organization and the division manager. The AI CEO was named as Foresight2. During wave 1, a manipulation check question asking about who the CEO was, allowed us to confirm participants had a good comprehension regarding the manipulation. Moreover, we wanted to warrant participants perceived the AI supervisor as a legitimate one. Thus, we assessed legitimate power, and participants' perception of this variable is above the scale's midpoint ($t(111) = 12.21$, $p < .001$). As expected, the same happens in the human condition ($t(95) = 26.74$, $p < .001$). In

the second wave, firstly, participants answered to the displacement of responsibility scale. Then, as in the first wave, we provided a description of the condition, and finally, we presented the scenarios to measure compliance with fraud and ethical leadership perception.

5.3.3. Measures (wave 1)

Moral identity. The self-importance of moral identity was assessed through the internalization dimension from Aquino and Reed (2002) scale (Cronbach's alpha = .79). A set of nine characteristics is shown (e.g., Caring, Compassionate, Fair, Friendly), and participants are asked to imagine a person with those characteristics and then to answer five items on a scale from 1 (strongly disagree) to 5 (strongly agree). Item examples are "It would make me feel good to be a person who has these characteristics" and "Being someone who has these characteristics is an important part of who I am". One item has been excluded due to reliability issues. As others did (e.g., Detert et al., 2008; Wang et al., 2019), we decided to measure the internalization subscale because, when compared to symbolization, it has been shown to be more associated with moral disengagement and (un)ethical behaviors (McFerran et al., 2010).

Legitimate power. Legitimate power was measured with two items from Schriesheim et al. (1991) subscale (Spearman-Brown coefficient$_{\text{human}}$ = .71; Spearman-Brown coefficient$_{\text{AI}}$ = .78). The chosen items are "Markem *(human supervisor)* / Foresight2 *(AI supervisor)* has a right to expect me to carry out his wishes", and "Markem / Foresight2 has been given the right to make demands of me.". Participants are asked to answer a scale from 1 (strongly disagree) to 5 (strongly agree).

5.3.4. Measures (wave 2)

Displacement of responsibility. This moral disengagement dimension was assessed with Detert et al. (2008) three-item scale (Cronbach's alpha = .71). A sample item is "If someone is pressured into doing something, they shouldn't be blamed for it". Participants are asked to answer on a scale from 1 (strongly disagree) to 5 (strongly agree).

Compliance with fraud. The compliance with fraud was assessed using Johnson et al. (2019) scenarios. We adapted the scenarios for the AI condition, where the Foresight 2 algorithm provided the unethical request, and one scenario was excluded from this study because it was difficult to interpret and to adapt for the AI condition. After reading each of the three scenarios, participants were asked to answer, using a 10-point scale, whether they would support the leader's request (1 – not at all to 10 – fully support request). Again, we checked scenarios manipulation (Human supervisor vs. AI supervisor) by asking participants who

requested the behavior.

Ethical leadership. Ethical leadership was measured using Brown et al. (2005) 10-item scale (Cronbach's alpha$_{human}$ = .93; Cronbach's alpha$_{AI}$ = .88). Item examples are "Can be trusted "and "Makes fair and balanced decisions". Some items were adapted to fit the AI supervisor condition (e.g., "Discusses business ethics or values with employees" was adapted to "Accepts employees' inputs regarding ethics or values"). Participants are asked to answer a scale from 1 (strongly disagree) to 5 (strongly agree).

5.3.5. Control variables

In line with previous studies on behavioral intentions in response to supervisors' unethical requests (e.g., Carsten & Uhl-Bien, 2013), we controlled for gender and age. These demographic characteristics are related to (un)ethical behaviors (e.g., Borkowski & Ugras, 1998; Mubako et al., 2020). We have also controlled the years of work experience as it is related to ethical issue recognition (Valentine & Bateman, 2011), and whether the participant holds a supervisor role, as people in higher organizational positions seem to perceive more ethnicity in unethical behaviors than those in lower rank positions (Kennedy & Anderson, 2017). Legitimate power was controlled because legitimacy is considered an important property of an authority, and a precursor of voluntary deference (Tyler, 2006). Perceived ethical leadership was controlled as it is an antecedent of unethical behavior (Moore et al., 2019).

5.4. Results

5.4.1. Measurement model

We conducted a confirmatory factor analysis for the global sample (N = 208) using AMOS v. 26. The measurement model shows a good fit to data, $\chi^2(32)$ = 54.63, p = .008, χ^2/df = 1.71, CFI = .97, TLI = .95, RMSEA = .06, CI90 [.03; .08] PClose = .28, SRMR = .04 (Hair et al., 2014). We have also conducted this analysis for alternative models, as depicted in Table 5.1 for the whole sample. Considering the existence of two conditions, we have repeated these tests for each and in both cases (Table 5.2 and Table 5.3) the baseline model shows better fit than the alternatives, as evidenced by $\Delta\chi^2$ and ΔCFI (Cheung & Rensvold, 2002).

Table 5.1 *Fit indices for measurement and alternative models (total sample).*

Models	Model 1 (Baseline Model)	Model 2 [a]	Model 3 [b]	Model 4 [c]	Model 5 [d]
χ^2 (*df*), *p* value	χ^2 (32) = 54.63, χ^2/df = 1.71, *p* = .008	χ^2 (34) = 202.26, χ^2/df = 5.95, *p* < .001	χ^2 (34) = 254.59, χ^2/df = 7.59, *p* < .001	χ^2 (34) = 224.82, χ^2/df = 6.61, *p* < .001	χ^2 (35) = 88.64, χ^2/df = 2.53, *p* < .001
CFI	. 97	.75	.67	.72	.92
TLI	. 95	.67	.56	.62	.90
RMSEA CI90, PCLOSE	.06 [.03; .08] .28	.16 [.13, .18] .000	.18 [.16, .20] .000	.17 [.15, .19] .000	.09 [.06, .11] .005
SRMR	.04	.12	.14	.11	.12
$\Delta\chi^2$ (df)	-	$\Delta\chi^2$ (2) = 147.62 *p*<.001	$\Delta\chi^2$ (2) = 199.96 *p*<.001	$\Delta\chi^2$ (2) = 170.19 *p*<.001	$\Delta\chi^2$ (3) = 34.00 *p*<.001
ΔCFI	-	.22	.30	.25	.05

Note. N = 208; χ^2 – chi-square; df – degrees of freedom; CFI – comparative fit index; TLI – Tucker–Lewis index; RMSEA – root mean square error of approximation; SRMR – standardized root mean square residual; $\Delta\chi^2$ – chi-square difference; ΔCFI – CFI difference

[a] Moral Identity and Displacement of responsibility combined into a single factor.

[b] Moral Identity and Fraud combined into a single factor.

[c] Displacement of responsibility and Fraud combined into a single factor.

[d] Model with independent factors.

Table 5.2 *Fit indices for measurement and alternative models (human condition).*

Models	Model 1 (Baseline Model)	Model 2 [a]	Model 3 [b]	Model 4 [c]	Model 5 [d]
χ^2 (*df*), *p* value	χ^2 (32) = 35.48, χ^2/df = 1.10, *p* = .316	χ^2 (34) = 127.20, χ^2/df = 3.74, *p* < .001	χ^2 (34) = 147.17, χ^2/df = 4.33, *p* < .001	χ^2 (34) = 98.48, χ^2/df = 2.90, *p* < .001	χ^2 (35) = 63.08, χ^2/df = 1.80, *p* < .001
CFI	. 99	.74	.69	.82	.92
TLI	. 99	.66	.59	.76	.90
RMSEA CI90, PCLOSE	.03 [.00; .09] .649	.17 [.14, .20] .000	.19 [.16, .22] .000	.14 [.11, .17] .000	.09 [.05, .13] .04
SRMR	.05	.15	.16	.10	.17
$\Delta\chi^2$ (df)	-	$\Delta\chi^2$ (2) = 91.73 *p*<.001	$\Delta\chi^2$ (2) = 111.70 *p*<.001	$\Delta\chi^2$ (2) = 62.99 *p*<.001	$\Delta\chi^2$ (3) = 27.60 *p*<.001
ΔCFI	-	.25	.30	.17	.07

Note. N = 96; χ^2 – chi-square; df – degrees of freedom; CFI – comparative fit index; TLI – Tucker–Lewis index; RMSEA – root mean square error of approximation; SRMR – standardized root mean square residual; $\Delta\chi^2$ – chi-square difference; ΔCFI – CFI difference

[a] Moral Identity and Displacement of responsibility combined into a single factor.

[b] Moral Identity and Fraud combined into a single factor.

[c] Displacement of responsibility and Fraud combined into a single factor.

[d] Model with independent factors.

Table 5.3 *Fit indices for measurement and alternative models (AI condition).*

Models	Model 1 (Baseline Model)	Model 2 [a]	Model 3 [b]	Model 4 [c]	Model 5 [d]
χ^2 (*df*), *p* value	χ^2 (32) = 49.25, χ^2/df = 1.54, *p* = .026	χ^2 (34) = 112.93, χ^2/df = 3.32, *p* < .001	χ^2 (34) = 142.97, χ^2/df = 4.21, *p* < .001	χ^2 (34) = 140.33, χ^2/df = 4.13, *p* < .001	χ^2 (35) = 61.40, χ^2/df = 1.75, *p* = .004
CFI	. 94	.75	.65	.66	.92
TLI	. 92	.66	.54	.55	.89
RMSEA CI90, PCLOSE	.07 [.03; .11] .193	.15 [.12, .18] .000	.17 [.14, .20] .000	.17 [.14, .20] .000	.08 [.05, .12] .065
SRMR	.07	.12	.14	.13	.10
$\Delta\chi^2$ (df)	-	$\Delta\chi^2$ (2) = 63.68 *p*<.001	$\Delta\chi^2$ (2) = 93.72 *p*<.001	$\Delta\chi^2$ (2) = 91.08 *p*<.001	$\Delta\chi^2$ (3) = 12.16 *p*=.007
ΔCFI	-	.19	.29	.28	.02

Note. N = 112; χ^2 – chi-square; df – degrees of freedom; CFI – comparative fit index; TLI – Tucker–Lewis index; RMSEA – root mean square error of approximation; SRMR – standardized root mean square residual; $\Delta\chi^2$ – chi-square difference; ΔCFI – CFI difference

[a] Moral Identity and Displacement of responsibility combined into a single factor.

[b] Moral Identity and Fraud combined into a single factor.

[c] Displacement of responsibility and Fraud combined into a single factor.

[d] Model with independent factors.

We have also checked for scales reliability and measurement model convergent and discriminant validity. The latent constructs show good internal consistency, as composite reliability (CR) is .81 for moral identity, .75 for displacement of responsibility, and .81 for compliance with fraud (human condition = .83; AI condition = .78) (Bagozzi & Yi, 1988), and Cronbach's alphas exceed .70 (Hair et al., 2014). Convergent validity was assessed via average variance extracted (AVE), which surpasses the threshold value of .50 for the three constructs: .52 for moral identity, .51 for displacement of responsibility, and .58 for compliance with fraud (human condition = .63; AI condition = .55) (Hair et al., 2014). The Fornell-Larcker criterion (Fornell & Larcker, 1981) confirms discriminant validity as each AVE's root square (.72 for moral identity, .71 for displacement of responsibility, and .76 for compliance with fraud - human condition = .79; AI condition = .74) is greater than the correlation between the latent constructs (r_{MI-DR} = -.21, $p < .01$; $r_{MI-Fraud}$ = -.15, $p < .05$ (human condition = -.25, $p < .05$; AI condition = -.06, $p > .05$); $r_{DR-Fraud}$ = .24, $p < .001$ (human condition = .39, $p < .001$; AI condition = .10, $p > .05$). Discriminant validity is also supported by the heterotrait-monotrait (HTMT) that is below the threshold of .85 (Henseler et al., 2015) for all conditions.

5.4.2. Measurement invariance

Before testing the hypotheses, measurement invariance was assessed. Accordingly, a multi-group confirmatory factor analysis was conducted to test configural, metric, and scalar invariance (Vandenberg & Lance, 2000). The measurement model shows excellent fit, $\chi^2(64)$ = 84.51, $p = .044$, $\chi^2/df = 1.32$, CFI = .97, TLI = .96, RMSEA = .04, CI90 [.01; .06] PClose = .77, SRMR = .05, which indicates configural invariance. Then, factor loadings were constrained to be equal across groups. This model was compared to the unconstrained one. The chi-square difference test results ($\Delta\chi^2 (7) = 12.30$, $p > .05$) indicates metric invariance. Finally, we constrained observed variables intercepts and compared with the previous model. The chi-square difference test ($\Delta\chi^2 (10) = 7.30$, $p > .05$) indicates scalar invariance. Overall, measurement invariance was confirmed, enabling group comparisons.

5.4.3. Descriptive statistics

The descriptive statistics and correlations are presented in Tables 5.4 and 5.5. Moral identity is negatively associated to displacement of responsibility (r_{total} = -.21, $p = .002$). For the human supervisor condition, compliance with fraud is negatively associated to moral identity ($r = -.25$, $p = .014$) and positively associated to displacement of responsibility ($r = .39$, $p < .001$). In the

AI condition, compliance with fraud is neither associated to moral identity ($r = -.06$, $p = .560$) nor to displacement of responsibility ($r = .10$, $p = .282$).

Table 5.4 *Descriptive statistics and correlations for the total sample.*

				Total sample (N = 208)							
Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. MI	4.61	.45								
2. DR	1.92	.71	-.21**							
3. Fraud	4.00	1.95	-.15*	.24***						
4. Age	40.56	10.75	.07	-.18**	-.10					
5. Gender	52.4% male		.18*	-.09	-.16*	.06				
6. W. Exp.	19.71	10.60	.16*	-.22**	-.14†	.93***	.12†			
7. Sup. Role	52.2% yes		.06	.06	-.01	-.09	.14*	-.05		
8. EL	1.71	.66	-.18**	.24**	.37***	.15*	-.08	-.23**	-.04	
9. LP	4.27	.50	.29***	-.12	.11	.10	.03	.17*	-.07	-.01

Note. MI = moral identity; DR = displacement of responsibility; LG = legitimate power; EL = ethical leadership; LP = legitimate power

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 5.5 *Descriptive statistics and correlations for each group.*

Variable	Human supervisor (N = 96)											AI supervisor (N = 112)	
	M	*SD*	1	2	3	4	5	6	7	8	9	*M*	*SD*
1. MI	4.62	.46		-.20*	-.06	.04	.20*	.14	.02	-.17†	.26**	4.61	.45
2. DR	1.90	.71	-.22*		.10	-.14	-.12	-.18†	.00	.08	-.10	1.93	.72
3. Fraud	4.07	2.03	-.25*	.39***		-.08	-.17†	-.15	-.13	.44***	.15	3.93	1.87
4. Age	40.48	10.61	.11	-.23*	-.12		.09	.92***	.04	-.09	.13	40.62	10.92
5. Gender	46.9% male		.15	-.05	-.17	.02		.17†	.09	-.02	-.02	57.1% male	
6. W. Exp.	19.99	10.68	.17†	-.26*	-.13	.95***	.06		.15	-.21*	.22	19.46	10.57
7. Sup. Role	47.9% yes		.10	.14	.12	-.24*	.19†	-.28**		-.08	-.12	56% yes	
8. EL	1.76	.72	-.20*	.40***	.31**	-.21*	-.16	-.25*	-.02		-.04	1.66	.60
9. LP	4.53	.56	.39***	-.15	.04	.09	.02	.08	-.08	-.02		4.04	.90

Note. The values placed below the diagonal represent the correlations for the human-supervisor sample. The values above the diagonal represent the correlations for AI-supervisor sample

MI = moral identity; DR = displacement of responsibility; LG = legitimate power; EL = ethical leadership

* $p < .05$; ** $p < .01$; *** $p < .001$

5.4.4. Two-stage least squares

Mediation models estimation can be biased if one disregards possible correlations between the mediator errors and the dependent variable errors (i.e., a common external condition that is responsible for a spurious relationship between these two variables, indicating endogeneity). Two-stage least squares (2SLS) regression has been used to account for this issue (Antonakis et al., 2010). This procedure requires one instrumental or exogenous variable (e.g., trait-based characteristics). In the first stage, the mediator is regressed on the instrumental variable, and the predicted values of the mediator are saved. Then, the dependent variable is regressed on the predicted values. Following this procedure, we first regressed displacement of responsibility on moral identity (while controlling for age, gender, work experience, and supervisor position). In the second stage, compliance with fraud was regressed on the predicted values of displacement of responsibility (while controlling for age, gender, work experience, supervisor position, ethical leadership perception, and legitimate power). As we wanted to test the moderator role of the supervisor (i.e., human vs. AI), we performed a multigroup analysis for the second stage. Results are reported in Table 5.6.

Table 5.6 *Two-stage least squares regression.*

Outcome variable	First stage					
	Displacement of Responsibility (DR)					
	β	SE	t	Sig	CI95 LB	CI95 UB
Moral Identity	-.18	.11	-2.48	.015	-.49	-.06
Age	.06	.01	.32	.749	-.02	.03
Gender	-.05	.10	-.65	.514	-.26	.13
Work experience	-.24	.01	-1.24	.216	-.04	.01
Supervisor role	.07	.10	1.06	.289	-.09	.30
Outcome variable	Second stage – total sample					
	Compliance with fraud					
	β	SE	t	Sig	CI95 LB	CI95 UB
Predicted value (DR)	.18	1.10	1.56	.120	-.45	3.89
Age	.06	.03	.33	.742	-.05	.08
Gender	-.13	.26	-1.99	.047	-1.03	-.01
Work experience	-.12	.03	-.65	.520	-.09	.05
Supervisor role	.04	.26	.55	.584	-.37	.65
Ethical leadership	.33	.20	4.79	.000	.57	1.36
Legitimate power	.17	.17	2.40	.017	.07	.74
Outcome variable	Second stage – Human condition					
	Compliance with fraud					
	β	SE	t	Sig	CI95 LB	CI95 UB
Predicted value (DR)	.49	1.80	2.59	.011	1.09	8.25
Age	-.37	.06	-1.16	.251	-.19	.05
Gender	-.06	.43	-.56	.576	-1.09	.61
Work experience	.68	.07	1.78	.078	-.02	.28
Supervisor role	.09	.43	.84	.402	-.50	1.22
Ethical leadership	.25	.29	2.46	.016	.14	1.28
Legitimate power	.17	.38	1.64	.105	-.13	1.39
Outcome variable	Second stage – AI condition					
	Compliance with fraud					
	β	SE	t	Sig	CI95 LB	CI95 UB
Predicted value (DR)	-.01	1.41	-.09	.926	-2.93	2.66
Age	.05	.04	.20	.841	-.07	.09
Gender	-.14	.34	-1.58	.118	-1.19	.14
Work experience	-.11	.05	-.45	.655	-.11	.07
Supervisor role	-.05	.34	-.56	.578	-.87	.49
Ethical leadership	.43	.29	4.51	.000	.73	1.87
Legitimate power	.18	.20	1.93	.056	-.01	.78

Results show that the predicted values of displacement of responsibility do not explain compliance with fraud thus not supporting Hypothesis 1. Results additionally show that

displacement of responsibility mediates the relationship between moral identity and compliance with fraud for the human condition but not for the AI condition. This difference was found to be significant (z-score = -2.18, p < .05), meaning that the condition moderates the mediation. Hypothesis 2 predicted a moderation effect where the AI supervisor would buffer the mediation effect. Thus, the hypothesis was only partially supported.

5.5. Discussion and Conclusion

This study was designed to address to which extent the nature of a supervisor (human vs. AI) changes the process where individual moral vulnerabilities explain compliance with an unethical request. To achieve this, the current study tested a moderated mediation model, established upon Identity-based motivation theory (Oyserman, 2007) and social cognitive theory (Bandura, 1991, 2016). Our results do not support a mediation effect by displacement of responsibility in the relationship between moral identity and compliance with fraud for the total sample. However, a total mediation effect was found for the human-supervisor condition. Once more, against our predictions, no mediation effect was found for the condition where an algorithm was playing the supervisor role. This section will detail the findings' interpretation highlighting its theoretical contributions and practical implications.

5.5.1. Theoretical and Practical Implications

The current study offers three critical contributions to the research field of business ethics. A first contribution relates to how moral-based individual characteristics make people vulnerable to unethical behaviors. Accordingly, the analysis of the relationship between moral identity and moral disengagement should be examined. This association is in line with previous research (Detert et al.,2008; Harris & He, 2019; Kennedy et al., 2017; McFerran, 2010). An important difference introduced by this study is its specific focus on a single moral disengagement mechanism. Although unidimensional moral disengagement measures are widely used (e.g., Moore et al., 2012), we followed Bandura's (2016, p. 26) recommendation to go over "one-size-fits-all measure" especially because not all locus of moral disengagement mechanisms suit the nature of a leader-follower relationship.

Second, to our knowledge, this is the first study to analyze how individuals' moral mechanisms help to predict compliance with fraud when a human supervisor vs. an algorithm is the source of the unethical request. Our findings show a mediation effect for the human supervisor condition, providing additional empirical support to literature focused on the critical

role of moral disengagement on unethical behavior (Newman et al., 2019). Particularly, as predicted, the displacement of responsibility mechanism suited the organizational context (Seriki et al., 2020; Hinrichs et al. 2012), especially where requests come from upper hierarchical levels (e.g., Carsten & Uhl-Bien, 2013). Therefore, individuals who feel more responsible for their behavior, due to moral centrality in their lives, display lower inclination to engage in organizational fraud incentivized by a supervisor. Unexpectedly, our findings show no mediation effect on the algorithm group. A few thoughts should be shared regarding this respect.

The mediation fails as displacement of responsibility does not predict compliance with fraud when the request comes from an algorithm. As mentioned, moral disengagement mechanisms operate differently across contexts (Bandura, 2016). Thus, it might be the case that the key mechanism to explain (un)ethical behavior when artificial intelligence takes the lead is not displacement of responsibility. If this is the case, which mechanism could take place in this process? Moreover, this lack of association may suggest followers do not assign responsibility to algorithms for their decisions, providing an empirical contribution to the topic of AI and accountability (Martin, 2019; Orr & Davis, 2020; van de Poel, 2020).

In searching for alternative mechanisms to displacement of responsibility, Orr and Davis (2020) findings may prove useful. Their study provide evidence on responsibility distribution for AI systems across different actors, which may suggest individuals would behave according to the levels of their diffusion of responsibility mechanism. One should not disregard this possibility; however, displacement and diffusion are associated mechanisms (Hinrichs et al., 2012), so it would be surprising to find opposite realities. Another mechanism that may play a role is moral justification. Accordingly, once algorithms are perceived as more rational (Lee, 2018), individuals could decide to comply with the AI supervisor's' request for the organization's sake. That is, employees could decide to comply with fraud to help the organization (Free & Murphy, 2015), engaging in a process known as unethical pro-organizational behavior (e.g., Chen et al., 2016).

Responsibility assignment is also an issue deserving discussion. It may be the case that participants may not recognize the algorithm as being responsible for its own decisions. The allocation of responsibility can target any of the human agents involved in designing or operating the algorithm (Martin, 2019). In the experiment conducted in this study (as it occurs in a real setting), employees were not aware of the programmer's profile, the guidelines for the algorithm development, or even the degree of customized parametrization the algorithm's owner enjoys. Failing to attribute responsibility suggests that, although efforts have been made

to make responsibility issues clear (e.g., IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019), when there is no direct and explicit information, people may fail to perceive algorithms as accountable for their own decisions.

Third, although not a goal of this study, our findings suggest that the difference between the two groups comprehends the explanatory process but not the unethical behavior's intensity. On average participants in both conditions report equivalent ($t(206) = .51$, $p = .610$) levels of compliance with fraud request from the leader. This contributes to the ongoing discussion on acceptance of AI advice, as different findings seem to diverge as regards discounting human vs. algorithms sources (Önkal et al., 2009; Prahl & Van Swol, 2017). Dietvorst et al. (2018) showed that it is possible to reduce algorithm aversion by enabling some control over an imperfect algorithm's forecast. In our research, participants were not able to modify the algorithm decision; but to a certain level, they had some control over the situation as they could freely decide whether to comply or not with the request.

The afore-discussed topics draw our attention to some important practical implications. Artificial intelligence is a reality in organizations (e.g., Malone, 2018), and these systems are now empowered to make decisions (Tokic, 2018) and assume a leader's role (Wesche & Sonderegger, 2019). Since we know that organizational fraud is a critical problem (e.g., Newman et al., 2019), it becomes crucial for organizations to understand which followers' characteristics may help them prevent these ethical issues in an AI context. Although our findings do show the answer is far from simple, as mentioned, we found that intention to commit fraud is similar in both contexts. This means that the issues surrounding ethical debates on AI (Larson & DeChurch, 2020; Munoko et al., 2020), particularly those concerned with possible biased outputs (Martin, 2019; Tambe et al., 2019), represent a real threat that organizations must be aware of, especially if the previous knowledge does not directly transfer to AI-based working settings.

5.5.2. Limitations and future research

The current study presents some limitations that future research can address. First, an experimental setting is not the same as a real work-setting, so findings should be carefully interpreted. However, not only scenario-based approaches have been widely used in the ethics field (e.g., Cianci et al., 2014; Detert et al., 2008) as we have followed Johnson et al. (2019) procedure when presenting the scenarios. Still, real settings can offer an understanding of more variables in play when dealing with non-human agents.

Second, as discussed, the displacement of responsibility mechanism fails to explain unethical behavior in the AI condition. Thus, future studies may consider replicating this study, including other moral disengagement strategies. For example, studies can assess how participants perceive the impact of the supervisor's unethical request on the organization, and thus uncover the role moral justification may play in such context.

Third, this study offers a different approach to the study of ethics and AI in the organizational context. Research places the focus of ethics on the systems' development (van de Poel, 2020) and the consequences to work of using such systems (Wright & Schultz, 2018). Escaping these focuses, the current study intended to provide a first contribution to understand how moral standards predict behavior when an algorithmic agent in a leading position requests an unethical behavior. The novelty of the approach raises many research questions. For example, literature suggests that people change their relationship with algorithms when they notice a mistake (Dietvorst et al., 2015). Thus, it would be interesting to contrast a situation with repeated unethical requests with an isolated request. Another interesting extension of this study could be to test a scenario where the algorithm could identify the misconduct and improve the decision. This may be important because it is known, from human-leadership studies, that apology from leaders impacts positively on employees (e.g., Basford et al., 2014). Also, a timely study would be to manipulate the subordinates' level of participation in the final decision. It is known that accepting these systems is easier when employees take part in the decision process (Dietvorst et al., 2018; Tambe et al., 2019). Somehow, this would mirror research from the traditional destructive leadership field that studies mechanisms such as constructive resistance (Tepper et al., 2001) and coproduction (Carsten & Uhn-Bien, 2013), which can be the most within-range human insurance against AI-based unethical decisions.

CHAPTER 6

Conclusion

This chapter is intended to offer an integrated view of the main conclusions and implications of the studies, after which due acknowledgement of limitations is showed, followed by some insights and directions for future research.

This research was motivated by the need to understand how resistance to destructive leaders could be theorized from a followership perspective. This research gap was felt because although literature on destructive leadership abounds and there is a parallel literature on followers' reactions to this sort of leadership (e.g., Carsten & Uhl-Bien, 2013; Greenbaum et al., 2013; Tepper et al., 2006), the emphasis is mostly placed on followers' enabling behaviors, highlighting both followers' passive or endorsing behaviors that would help understand the prevalence and dynamics of enduring destructive leadership (Lipman-Blumen, 2005; Thoroughgood et al., 2012a; Schyns et al., 2018b).

A premise from which this research departed concerns the unlikeliness of such resistance behaviors being only an expression of dysfunctionality. Quite the opposite is suggested by Mary Parker Follett (1949) heritage on organizational theory where this pioneering author proposed the true leader has an invisible nature, i.e., the ultimate purpose that provides meaning to those that follow is the true leader. Therefore, resisting a leader's will might even be functional if that will go counter to the organizational purpose. Despite these cues in literature, the leader-centric approach that prevailed until the last decades wrongfully conceived resistance as being necessarily a non-following behavior. Such is tacitly expressed by Tepper's choice of wording to name non-constructive resistant behaviors as "dysfunctional resistance". A closer look into dysfunctional resistance subscale (Tepper et al., 2001) suggests items are not necessarily linked to ill-intended action.

With this in mind, a sequence of studies was designed (and adjusted in the process of research development) to focus on several key issues linked to the research gap. Namely, the true nature of resistance behaviors with an emphasis on destructive leaders, as this is the context that most likely enables the distinction between following and non-following as well as the clarification on Follett's contention on the true nature of leadership. Another related key issue pertains to the alleged nature of leadership as being coproduced (Shamir, 2007), which would give a central role to coproduction beliefs (Carsten & Uhl-Bien, 2012) into understanding the motivations of resistance when facing destructive leaders. A third key issue concerns the possible heterogeneous nature of dysfunctional resistance which cannot *a priori* be taken as an

all-encompassing category for all non-constructive resistance behaviors. Lastly, a fourth issue targets an emerging topic in leadership studies, which can also challenge current assumptions about resistance behaviors: that leaders are not necessarily individuals, but rather, agents such as synthetic agents (e.g., Wesche & Sonderegger, 2019). This extraordinary novel situation is an ultimate test to understand to which measure resistance to hypothetical destructive leaders is extensive to synthetic agents and, most importantly, if the psychological processes mediating these behaviors remains valid.

Study 1 echoes the current debate on followership, its nature, and meaning and draws attention to the question "who/what does a follower follow?". This question entails some implications: the target of following; the voluntary nature of the following behavior; and the purpose of following. In line with Uhl-Bien et al. (2014) this study integrates the role-based and constructionist perspectives, which allows proposing a two-layer approach for followership, where it is possible to remain a follower even when not following the leader. This happens when followers decide not to follow the leader in the name of the organizational good (i.e., when followers perceive that the leader threatens the group and/or organization). For this reason, destructive leadership (specifically, destructive leaders) emerges as a critical background/context for this thesis. Departing from an adverse context to study how a system works has, in fact, been a common practice in several domains. This is not to say that one should provoke adversities, but their analysis may become instrumental once they exist. Examples of this approach are studying work accidents to enhance work processes and workers' safety (Leplat, 1978; Leplat & Cuny, 1978) or studying brain injuries to gain knowledge about brain functions and processes (Damasio et al., 2005; Squire, 2009).

Accordingly, it is in such a context that Study 1 allows to build on the idea that resistance can emerge as a mechanism that seeks to restore leadership (or the invisible leader); beyond that, it was possible to provide empirical support to previous models that propose destructive leadership as a result of followers' behaviors, even passive behaviors (e.g., Thoroughgood et al., 2012a). This claim is not innocuous and resembles Kellerman (2012) idea that followers are maintaining the *status quo* when doing nothing. As previous authors have discussed (e.g., Lipman-Blumen, 2005), there are many reasons for followers to remain passive (e.g., pragmatism, fear, alignment), and this points out the importance of decision making (i.e., is there any volition?). We do not dismiss this important debate (e.g., Blom & Lundgren, 2020); however, from our perspective, regardless of volition, followers' behaviors do have an impact that will always imply coproduction (either of good or destructive leadership).

The main takeaway from Study 1 is that resistance can be conceived as a mechanism to

help protect leadership, which is aligned with studies showing constructive resistance as a consequence of coproduction beliefs. However, as mentioned in the limitations section, it was not possible to profile followers according to the leader's destructive behavior. Moreover, as observed, this kind of research is still scarce.

Thus, Study 2 was designed to analyze the robustness of coproduction in predicting resistance across different leaders' behaviors (i.e., different targets and intensity). In Study 1, it was possible to identify two types of resisters: actives and passives. There seems to be indication that both groups may display resistance against a destructive leader motivated by the will to follow the organizational good. The underlying difference between these groups lies in their perceived strength to engage in active behaviors. So, overall, both resistance styles (actives and passives) can protect the organization from destructive leaders and opting for one or the other depends on power perceptions, i.e., how much one experiences sense of power to act.

Previous literature, however, classified resistance in a different way: as constructive and dysfunctional (Tepper et al., 2001). According to this typology, constructive resistance can be seen as ethical followership (Carsten & Uhl-Bien, 2013), while dysfunctional resistance holds a retaliatory nature (Tepper et al., 2006). Albeit all kinds of resistance can be judged as functional (as it may protect the organizational good from destructive leaders) Study 2 framed hypotheses within Carsten and Uhl-Bien (2013) and Tepper et al. (2001, 2006) views, and proposed that coproduction beliefs would favor constructive resistance and decrease dysfunctional resistance. Following the same rationale, these hypotheses were grounded in light of proactivity theory, which made it possible to frame coproduction as a "reason to" resist constructively. Findings showed the important role of coproduction beliefs in predicting constructive resistance across different destructive leader's behaviors scenarios. Most interestingly, results were not consistent for dysfunctional resistance, as coproduction was an explanative factor only in the abusive and laisser-faire conditions. An enlightening fact from this study was the recurrently observed decision in akin literature, of excluding some items from the resistance scale, namely the dysfunctional subscale (Brett et al., 2016; Carnevale et al., 2018; Vecchio et al., 2010). This also occurred in this study for psychometric reasons. A closer look into the retained items, show they are linked to covert behaviors. So, results suggest that, for example, in a situation of greater hostility, it might be more important to behave silently. However, in situations like these, those holding strong coproduction beliefs would engage less in this kind of resistance.

Integrating these findings with those from Study 1, we argue that coproduction as an antecedent of covert resistance does not provide information about the ethical nature of

behaviors. Moreover, the operationalization of coproduction beliefs entails the perception of an active role by followers, so it is reasonable to find fewer covert behaviors in individuals holding these beliefs. This Study 2 provides additional evidence on the role of coproduction beliefs as a predictor of resistance (particularly constructive resistance); however, it also raises some uncertainty regarding the nature of dysfunctional resistance.

Although Tepper et al. (2001) conceptualize dysfunctional resistance as a type of passive-aggressive behavior, evidence from previous research shows some divergences in the number of items used to assess resistance. These divergences cast doubt on the scale's structure. Accordingly, in Study 3, we started by analyzing the factor structure of the scale to find psychometric evidence supporting a 3-factor solution as the best option. This indicates three resistance behaviors instead of two: constructive, subversive, and confrontational (subversive and confrontational resulting from dysfunctional).

As our main concern at this stage was the nature of dysfunctional resistance, Study 3 included a second empirical study to address this issue. Based on previous findings, namely those from Study 1, where covert behaviors seem to emerge due to lack of power, a model was designed to assess the differences in the processes leading to subversive and confrontational resistance. Although there is evidence that abusive behaviors by leaders foster all types of resistance, it is impossible not to think about a parallel process that weakens active resistance expressions. As abusive supervision can be conceived as an important organizational stressor, the Conservation of Resources theory (Hobfoll, 1989) was used to frame a fear-based depletion process that reduces confrontational resistance (an approach strategy). As (lack of) power was implicated as a critical variable to explain covert behaviors (an avoidance strategy), we used Job Demands-Resources theory (Bakker & Demerouti, 2014) to introduce the role of sense of power. Accordingly, our findings show that when individuals feel a high sense of power (towards their colleagues), they use less subversive behavior, while those with a low sense of power engage in more subversion.

As a non-compliance behavior, resistance seems to emerge as a mechanism that expresses rejection of destructive leaders, making it critical to any leadership process. Research shows that leadership is not a human-exclusive process (Bastardoz & Van Vugt, 2019; Smith, 2017) and it seems to be conceived within same-species although there some interspecies leadership processes are well documented (e.g., canine-human, Range et al., 2019). Is it reasonable, then, to think of leadership processes between different-nature agents where the human is the follower, not the leader? Is it possible to conceive leadership as a process in which the one in command (i.e., the leader) is a synthetic agent (e.g., an AI-based algorithm)? Leadership is an

interactive process where one agent influences a group to reach a common good. Thus, the focus is not on any trait or leader's characteristic but the transactional/interactive process between the leader and the followers (Northouse, 2016). This process does not seem incompatible with the interaction between different agents. In any case, when we think about a non-human decision-maker agent influencing the behavior of others, it is an already-existing downward influence process (Wesche & Sonderegger, 2019).

If it is timely to study these new leadership dynamics, it is then critical not to forget its possible dark side. When describing leaders (or decision-makers) with destructive behaviors, there is a wide range of possible behaviors they can display, and one of them is to harm the organization (e.g., fraud). This topic has been gaining increasing attention, given its extent and impact (ACFE, 2020; van Driel, 2019). Moreover, an important reason for concern is that followers seem to have a solid role in the dissemination of such behaviors due to their compliance (Johnson et al., 2019). Since there is evidence that algorithms might mirror the ethical biases or blind spots of the coders or perpetuate systematic errors (Kellogg et al., 2020; Martin, 2019; Meyer, 2018), it becomes crucial to extend/transfer the existing knowledge on followers' ethical protective mechanisms to this new context.

It was based on these ideas that study 4 was designed. This study aimed to test an individual-based moral process in two distinct conditions: human and algorithm supervisor[8]. For the human condition, the results reinforce the nature of specific moral standards in explaining compliance; however, some surprising findings came up for the AI condition. All in all, no link between the displacement of responsibility and compliance with fraud was found, which makes clear that the explanative process changes with the nature of the agent. It is acknowledged that this study does not allow to conclude on the compliance with fraud explanatory mechanism when an algorithm generates a request. Still, it provides a useful contribution as it clearly draws attention to the relevance and need for further research on the topic. This claim is supported by the complementary finding that there are no differences in compliance between the two groups. Once there are no differences, it is reasonable to assume that followers do follow (or resist) both kinds of leaders, i.e., the dynamics may have a different

---

[8] The choice for using "supervisor" instead of "leader" is purposeful as the former does not entail the attributional process that characterizes leadership and the study design (being experimental) precludes ascertaining these attributional processes (which would imply a real interaction across time). Still, the formal authority dimension in leadership is also an important one and often used as the dominant feature in leader-follower dynamics.

root, but the tip of the iceberg is stable. It is then critical to delve deep into the processes that explain behaviors in such a new context.

Designed to develop knowledge on the followership field, the set of studies comprised in this thesis seemingly offer some relevant contributions both to theory and practice.

6.1. Theoretical implications

This thesis answers some calls (Blom & Lundgren, 2020; Collinson, 2017) on the need to clarify the followership concept and its nature. Especially in Study 1, we engaged in this debate and presented a perspective that allows the articulation of a formal role (position-based) with the idea that both leaders and followers can display some leading behaviors when trying to build (together) leadership (i.e., following the common goal). In other words, alongside other researchers, we aim to recapture the pioneering idea of Mary Parker Follett and take a step further by claiming, "yes, you can still be a follower when you do not follow the leader".

Beyond this, this thesis also intended to contribute to theory development as it offers an empirically based classification of followers' profiles (within the specific context of having a destructive leader). This work brings to discussion two critical aspects for a better understanding of followership: the relevance of behavioral drivers and the existence of mixed behavioral profiles. Some behavioral drivers allow to uncover intentions and may also help answer the question "who/what is the target of following?". For example, is it acting out of fear a following behavior? Is such an individual following either the leader or the organizational good? Or none of both? So, intentions matter. Uncovering the underlying intentions or constraints helps to understand whether there is any voluntary nature in following. However, as mentioned and described in Study 1, from our standpoint, this is not the most critical aspect: those who remain silent due to either fear or to their alignment with the leader, are both destructive leadership builders. This theorization is not deprived of relevance as, e.g., the importance of uncovering the motivations is linked to its possible impact on policies: for example, when organizations accept that fear is a paralyzer (e.g., individuals opt to remain silent), they can develop organizational policies to provide a secure and safe climate.

The existence of mixed profiles also adds relevant discussion on the topic. It provides evidence that human nature, and followership behavior, in particular, is not compatible with a dichotomic view of the phenomenon. Moreover, as observed in the evolution of studies on leaders' behaviors (contingency approach, e.g., Hersey et al., 1979), it is reasonable to assume that followers can present different behaviors without this leading to an idea of inconsistency

(Hertwig & Gigerenzer, 2011). Throughout some of the other studies, we found resonances of these blurred profiles, namely in the correlations found between the types of resistances, showing that they can co-occur. Also, in Study 3, there is a mediation process showing that some resistance behaviors are reduced due to fear; however, at the same time, there is a positive direct effect. Could this be one more piece of evidence for the behavioral complexity? The existence of such profiles draws attention to the complexity of human behavior analysis. Human behavior is not defined by any binary (0 and 1) language. As biopsychosocial agents (Engel, 1977), our behavior results from complex inter and intra-individual interactions.

Another feature to which this work aims to contribute concerns the concept of coproduction beliefs. Accordingly, we tested their robustness across different situations (Study 2). Moreover, as predictors of a type of active resistance (constructive resistance), these beliefs were framed in a theoretical model that offers an integrated analysis of its manifestation (i.e., their impact was assessed while competing with factors holding more contextual nature that are explained through emotions and perception of power).

A closer analysis of resistance behaviors (Tepper et al., 2001) allows discussing a reinterpretation of the meaning of such behaviors. Accordingly, the so-called dysfunctional resistance behaviors can be seen as holding a mix of active and passive nature. This is intriguing as it draws attention to the idea of purposeful silence, especially when individuals are deprived of power resources. In these circumstances, silence can be considered an opportunity to express an invisible behavior, which fit in the class of passive resistance followers (Study1). Moreover, we should note that we assessed defensive silence, that does not hold the manipulative nature of other silences such as opportunistic silence (Knoll & Van Dick, 2013), suggesting it can be seen as a last resort behavior.

Finally, Study 4 supports existing theory on individual moral-based characteristics when predicting behaviors of compliance/resistance with unethical requests. Beyond this, it provides empirical evidence for the need to develop theory that guides the interpretation on emerging leadership/followership dynamics when interacting with synthetic agents.

6.2. Practical implications

This collection of studies also allows deriving some practical implications primarily useful to inform organizations and decision-makers. As mentioned, better understanding the nature of followership and the meaning of resistance/compliance makes it easier for organizations to (un)lock their manifestation. Also, thinking of leadership as a process that can be balanced by

followership allows one to plan strategies to curb destructive leaders that go beyond interventions solely targeting the leaders (e.g., training programs). The idea that organizations can use followers as protective agents is of special relevance when we face data showing that organizations attract dysfunctional profiles to leading positions (Chamorro-Premuzic, 2013). As detailed in Study 1, followership implies accepting to have responsibility.

Moreover, taking Study 2 findings that coproduction positively impacts a negotiation-based strategy (i.e., constructive resistance), organizations should promote coproduction. In a similar vein to what is done with respect to leadership, followership training programs can be thought of (Riggio, 2014, 2020). There seems to be a scarce investment in such programs because people believe that knowing how to follow is common sense (Agho, 2009); however, as Kelley (2008) pointed out, it is important to ensure that followers are able to think critically. In a way, it would not be far from Ira Chaleff's (2015) intelligent disobedience proposal, a concept that emerges from training programs for guide dogs. These dogs are expected to obey their blind owners (or caretaker), but they should be taught to disobey if they perceive that their owner is at risk.

While examining resistance mechanisms, the third study showed that there could be a sound in silence. In other words, resistance may be silently performed, especially when no power resources are available. Since silence is difficult to access and interpret (Van Dyne et al., 2003), it is important that organizations are aware of the possibility of underground movements, but above all, they should try to provide resources that make it less necessary to resort to these behaviors.

Finally, we are now undergoing a paradigm-shifting moment regarding the nature of leaders. Accordingly, as mentioned, there are already important levels of automation even in decision-making processes (Wesche & Sonderegger, 2019). It is then critical to inform organizations and make them aware of their possible drawbacks. This gains particular relevance once we consider the evidence on ethical failures in these systems, as many have been already surfacing (e.g., Kellogg et al., 2020; Martin, 2019; Meyer, 2018). Moreover, as Study 4 shows, it is possible to find similar reactions in both human and AI-based systems. So, organizations must be vigilant when implementing these systems to avoid repeating what already goes wrong in human-based destructive leadership processes, with the huge disadvantage that meaning or self-censoring may not play a role in the mind of the agent.

6.3. Limitations and directions for future research

The implications of this thesis cannot fairly be interpreted without acknowledging some limitations. Although the limitations were pointed out in the description of each study, they will now be analyzed in a global and integrated fashion. Acknowledging limitations also paves the way to think of and point directions for future research.

As often acknowledged in much quantitative empirical research, in this study it would have been advantageous to work with larger samples. Particularly, in Study 1, the sample size prevented analyzing followers' profiles as against each type of leaders' destructive behavior. In Study 2, this issue was integrated by presenting scenarios of different leaders' behaviors, allowing to compare followers' behaviors across groups. However, presenting scenarios has its own limitations (Lonati et al., 2018). especially those concerning results interpretation, as intentions and not real behaviors are assessed. Nevertheless, it is a widely and generally well-accepted method when studying organizational processes (Cianci et al., 2014; Tseng, 2019).

Study 3 focused on a specific destructive behavior (abusive supervision), and we asked people about their actual experience instead of inquiring about intentions. Although some relevant findings were reported, two limitations should be highlighted, particularly in the second part of the study.

Firstly, the average for abusive supervision is low (1.42), and so is the standard deviation (.60). Although this is in line with other studies (e.g., Ju et al., 2019; Schyns et al., 2018a; Tepper, 2000; Tepper et al., 2008), we should reflect on its consequences as it shows that people do not perceive high levels of abuse by their leaders. This does not preclude the analysis of relationships between variables; however, it does not offer complete assurance that the emerging processes reflect what happens when abusive supervision is real. Even though being the most widely used measure to assess destructive leaders' behaviors, the abusive supervision scale may not represent the most typical destructive behaviors, perhaps due to its severe level of hostility (Schmid et al., 2018). This is an opportunity for future research, as one may acknowledge that the scale may require a variant that moderates the hostility of the abuse. Likewise, future research may benefit from testing models that assess other leaders' behaviors.

Secondly, in this same study, data were collected from the same source and at the same time. Although some recommended procedures were adopted to deal with this cross-sectional issue (Podsakoff et al., 2012), and the existence of moderation rules out the possibility of common method bias (Siemsen et al., 2010), it would have been a more robust design to have the data collected in waves. The last study shows an advancement in this methodological aspect,

with data collected at different time points. Moreover, a data analysis strategy was adopted to prevent endogeneity issues (Antonakis et al., 2010). This study, on the other hand, has the already mentioned limitation of working with behavioral intentions.

Some of the aforementioned issues were addressed by subsequent studies; however, other limitations can guide future research. For example, although we have evidence of the existence of mixed profiles of behaviors, which, as we have discussed, support the idea of the complexity of the nature of followership, more research is needed to uncover its blurred nature. In other words, what are the factors that explain behavioral contradictions? This issue has already been raised by Kelley (2008, p. 8) - "Is a person's style static or dynamic?"-, but there still seems to be no satisfactory answer.

Another topic worthy of attention is linked to resistance behaviors. We concluded that dysfunctional resistance comprises two types of behavior that deserve individual attention and represent a more active or passive way to resist. However, when analyzed in this way, it is reasonable to ask what distinguishes confrontational resistance from constructive resistance, as both hold an active nature. Should the ethical nature of these behaviors be considered? Does, for example, coproduction also favor confrontational resistance? Another promising direction relates to the order of such behaviors. Is negotiation an initial strategy that may evolve into confrontation when it is not well succeeded? The specific nature of these two active strategies should then be examined in future studies.

Future research can also expand knowledge by introducing a team-level approach. Leadership involves influencing a group, and followers most of the time work in teams, and it is known that some group social phenomena (e.g., conformity) in some circumstances may constrain behavior (Bond, 2005). Thus, studying teams' resistance may be useful to strengthen knowledge on these dynamics. It would also be relevant to measure and test coproduction at the team level as a perception of organizational (or team-based) climate. These constructs (collective resistance, coproduction culture) are most likely operating at the team level, and adopting a team-level test, or even a multilevel model might offer novel findings.

While still thinking about coproduction, future research may benefit from analyzing the operationalization of the concept. The way it has been measured expresses how a follower thinks he or she should interact with someone in a leading position. Does coproduction always have to go through active work with the leader? Doesn't this rule out some behaviors (like covert behaviors) to be an expression of the desire to follow the invisible leader? Doesn't this assume that covert behavior (when overt is not possible and not due to ethical constraints or manipulative intention) will never be explained by coproduction beliefs?

Another limitation that future studies can address relates to the inclusion of organizational variables. Leadership is about the dynamics between a leader and followers, but it occurs in a specific context (Padilla et al., 2007; Uhl-Bien et al., 2014). This was not tested as the models were already complex and too many effects may bring undesirable biases to the analysis. However, upcoming research – based on these findings - can include variables such as the organizational climate (e.g., ethical climate), the level of political behavior at the workplace, or test models across different industries. At a more macro level, it would also be enlightening to link some of these topics with national work cultures to enable cross-cultural comparisons.

It should be asserted that this thesis aligns with the criticism of leader-centric models (Collinson, 2008b; Thoroughgood et al., 2018). So, although some of the tested models analyze the impact of leaders on followers' behaviors (which somehow reflects a leader-centric perspective), the variables under analysis aim to reflect active behaviors from followers which inscribes the studies within the larger perspective of followership. Accordingly, followers are not treated as recipients of the leaders' behaviors; however, in the future, it will be important to extend these models and study the consequences of followers' behaviors (i.e., to reverse the causal nexus direction and study the consequences of resistance on destructive leaders and organizations, while keeping in mind the most likely circular causality in these interactional models).

Lastly, it should be stated that this work is focused on the "good side" of followership, as we opted to approach it as an organizational resource. However, we acknowledge that followership (in the same vein as leadership) also has its dark side, and it can hold a destructive/toxic nature (Offerman, 2004; Thomas et al., 2017). An example of this is illustrated by Lipman-Blumen (2005) malevolent followers that are the ones who induce the leader to behave destructively.

6.4. Concluding remark

Overall, destructive processes stemming from leaders have pervaded humankind's history and have been a source of incommensurable loss, suffering, and the defeat of civilizational advances. History is more of a collection of such reports rather than its opposite. Still, this seems to go against the grain of humankind as all societies struggle against the inexorability of such episodes and endeavor to build ethical grounds that guide (and limit) the exertion of power. This holds for the societal level as well as for the organizational level.

Organizational destructive leadership has been targeted in research precisely as an expression of such an ethical endeavor. As a means to fight against such behavior, scholars started by theorizing and uncovering the models that explain destructive leadership with a focal point on leaders themselves. As research progressed, such focal point shifted to the interaction between leaders and followers, thus acknowledging the dysfunction should not be thought of as singular but rather as collective in nature.

From a followership theory standpoint, there is more to it than identifying followers' behaviors that enact or facilitate the action of such destructive leaders; followers, as coproducers, can also be a prodigious source of prevention if indeed they resist based on those ethical grounds. This was the focal point of this study: understanding followers' resistance by adding to extant categorizations, highlighting coproduction beliefs as a key driver of resistance, explaining how an extreme form of destructive leader behavior (abusive) is associated with different sorts of resistance, and exploring how these dynamics can extrapolate to emergent vertical-power relations at work, such as those created by the introduction of synthetic agents in organizational life. These contributions offer just a partial view to the vast field of followership, but they call attention to the efficacy of putting emphasis upon followers' resistance rather than leaders' influence.

The other side of the coin is the possibility that destructive leaders are far less harmful than destructive followers, as the former may be stopped by the followers, but the latter may not be stopped by the leaders. Therefore, follower resistance is key to safeguarding against either situation, and this is why, we contend, it is worth studying and has enormous heuristic potential within organizational science.

References

Aasland, M. S., Skogstad, A., Notelaers, G., Nielsen, M. B., & Einarsen, S. (2010). The prevalence of destructive leadership behaviour. *British Journal of Management*, *21*(2), 438-452. https://doi.org/10.1111/j.1467-8551.2009.00672.x

ACFE. (2020). Report to the Nations on Occupational Fraud and Abuse. https://acfepublic.s3-us-west-2.amazonaws.com/2020-Report-to-the-Nations.pdf

Agho, A. O. (2009). Perspectives of senior-level executives on effective followership and leadership. *Journal of Leadership & Organizational Studies*, *16*(2), 159-166. https://doi.org/10.1177/1548051809335360

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*, *63*(1), 1-18. https://doi.org/10.1111/j.2044-8325.1990.tb00506.x

Almeida, T., Ramalho, N. C., & Esteves, F. (2021). Can you be a follower even when you do not follow the leader? Yes, you can. *Leadership*, *17*(3), 336-364. https://doi.org/10.1177/1742715020987740

Alvesson, M., & Blom, M. (2015). Less followership, less leadership? An inquiry into the basic but seemingly forgotten downsides of leadership. *M@n@gement, 18*(3), 266-282. https://doi.org/10.3917/mana.183.0266

Anand, V., Ashforth, B. E., & Joshi, M. (2004). Business as usual: The acceptance and perpetuation of corruption in organizations. *Academy of Management Perspectives*, *18*(2), 39-53. https://doi.org/10.5465/ame.2004.13837437

Anderson, C., John, O. P., & Keltner, D. (2012). The personal sense of power. *Journal of Personality, 80*(2), 313-344. https://doi.org/10.1111/j.1467-6494.2011.00734.x

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21*(6), 1086-1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Aquino, K., & Freeman, D. (2009). Moral identity in business situations: A social-cognitive framework for understanding moral functioning. In D. Narvaez & D. K. Lapsley (Eds.), *Moral personality, identity, and character* (pp. 375–395). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511627125.018

Aquino, K., & Reed II, A. (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

Ashforth, B. E. (1997). Petty tyranny in organizations: A preliminary examination of antecedents and consequences. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, *14*(2), 126-140. https://doi.org/10.1111/j.1936-4490.1997.tb00124.x

Ashforth, B., & Anand, V. (2003). The normalization of corruption in organizations. *Research in Organizational Behavior*, 25, 1–52. https://doi.org/10.1016/S0191-3085(03)25001-2

Astrove, S. L., Yang, J., Kraimer, M., & Wayne, S. J. (2015). Psychological contract breach and counterproductive work behavior: A moderated mediation model. *Academy of Management Proceedings*, 11094. https://doi.org/10.5465/ambpp.2015.11094abstract

Avolio, B. J., Bass, B. M., & Jung, D. I. (1999). Re-examining the components of transformational and transactional leadership using the Multifactor Leadership. *Journal of Occupational and Organizational Psychology*, *72*(4), 441-462. https://doi.org/10.1348/096317999166789

Babbie, E. R. (2014). *The Practice of Social Research.* Canada: Nelson Education.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94. https://doi.org/10.1007/BF02723327

Baker, S. D. (2007). Followership: The theoretical foundation of a contemporary construct. *Journal of Leadership & Organizational Studies*, *14*(1), 50-60. https://doi.org/10.1177/0002831207304343

Baker, S. D., Stites-Doe, S. A., Mathis, C. J., & Rosenbach, W. E. (2014). The fluid nature of follower and leader roles. In L. M. Lapierre & M. K. Carsten (Eds.), *Followership: What is it and why do people follow* (pp. 73-88). Bingley: Emerald Group Publishing Limited.

Bakker, A. B., & Demerouti, E. (2014). Job Demands–Resources theory. In C. Cooper & P. Chen (Eds.), *Wellbeing: A complete reference guide* (pp. 37– 64). Chichester, UK: Wiley-Blackwell. https://doi.org/10.1002/9781118539415.wbwell019

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248-287. https://doi.org/10.1016/0749-5978(91)90022-L

Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. New York: Worth Publishers.

Bandura, A. (2018). Toward a psychology of human agency: Pathways and reflections. *Perspectives on Psychological Science*, *13*(2), 130-136. https://doi.org/10.1177/1745691617699280

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, *71*(2), 364-374. https://doi.org/10.1037/0022-3514.71.2.364

Barsky, A. (2011). Investigating the effects of moral disengagement and participation on unethical work behavior. *Journal of Business Ethics*, *104*(1), 59-75. https://doi.org/10.1007/s10551-011-0889-7

Basford, T. E., & Offermann, L. R. (2012). Beyond leadership: The impact of coworker relationships on employee motivation and intent to stay. *Journal of Management & Organization*, *18*(6), 807-817. https://doi.org/10.5172/jmo.2012.18.6.807

Basford, T. E., Offermann, L. R., & Behrend, T. S. (2014). Please Accept My Sincerest Apologies: Examining Follower Reactions to Leader Apology. *Journal of Business Ethics*, *119*(1), 99–117. https://doi.org/10.1007/s10551-012-1613-y

Bastardoz, N., & Van Vugt, M. (2019). The nature of followership: Evolutionary analysis and review. *The Leadership Quarterly, 30*(1), 81-95. https://doi.org/10.1016/j.leaqua.2018.09.004

Bauer, M. W. (2000). Classical content analysis: A review. In M. V. Bauer & G. Gaskell (Eds.), *Qualitative researching with text, image and sound: A practical handbook* (pp. 131–151). London: Sage. https://dx.doi.org/10.4135/9781849209731.n8

Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, *43*(3), 800-813. https://doi.org/10.3758/s13428-011-0081-0

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

Beu, D. S., & Buckley, M. R. (2004). This is war: How the politically astute achieve crimes of obedience through the use of moral disengagement. *The Leadership Quarterly, 15*(4), 551-568. https://doi.org/10.1016/j.leaqua.2004.05.007

Biddle, B. (1986). Recent Developments in Role Theory. *Annual Review of Sociology*, *12*(1), 67–92. https://doi.org/10.1146/annurev.so.12.080186.000435

Bindl, U. K., & Parker, S. K. (2010). Feeling good and performing well? Psychological engagement and positive behaviors at work. In S. L. Albrecht (Ed.), *Handbook of employee engagement: Perspectives, issues, research and practice* (pp. 385–398). Cheltenham, UK: Edward Elgar. https://doi.org/10.4337/9781849806374.00043

Birmingham, L. E., Cheruvu, V. K., Frey, J. A., Stiffler, K. A., & VanGeest, J. (2020). Distinct subgroups of emergency department frequent users: a latent class analysis. *The American Journal of Emergency Medicine*, *38*(1), 83-88. https://doi.org/10.1016/j.ajem.2019.04.029

Blasi, A. (1994). Moral identity: Its role in moral functioning. In B. Puka (Ed.), *Morality, moral behavior, and moral development* (pp. 168-179). New York: Garland Publishing.

Blass, T. (1991). Understanding Behavior in the Milgram Obedience Experiment: The Role of Personality, Situations, and Their Interactions. *Personality and Social Psychology*, *60*(3), 398–413. https://doi.org/10.1037/0022-3514.60.3.398

Bligh, M. C., & Schyns, B. (2007). Leading question: the romance lives on: contemporary issues surrounding the romance of leadership. *Leadership*, *3*(3), 343-360. https://doi.org/10.1177/1742715007079316

Blom, M., & Lundgren, M. (2020). The (In) voluntary follower. *Leadership, 16*(2), 163-179. https://doi.org/10.1177/1742715019888078

Bond, R. (2005). Group size and conformity. *Group Processes & Intergroup Relations*, *8*(4), 331-354. https://doi.org/10.1177/1368430205056464

Borkowski, S. C., & Ugras, Y. J. (1998). Business students and ethics: A meta-analysis. *Journal of Business Ethics*, *17*(11), 1117-1127. https://doi.org/10.1023/A:1005748725174

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks: Sage Publications.

Brett, J. F., Uhl-Bien, M., Huang, L., & Carsten, M. (2016). Goal orientation and employee resistance at work: Implications for manager emotional exhaustion with the employee. *Journal of Occupational and Organizational Psychology*, *89*(3), 611-633. https://doi.org/10.1111/joop.12144

Brinsfield, C. T. (2013). Employee silence motives: Investigation of dimensionality and development of measures. *Journal of Organizational Behavior, 34*(5), 671-697. https://doi.org/10.1002/job.1829

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185-216. https://doi.org/10.1177/135910457000100301

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, *97*(2), 117-134. https://doi.org/10.1016/j.obhdp.2005.03.002

Burke, R. J. (2006). Why leaders fail. Exploring the dark side. In R. J. Burke & C. L. Cooper (Eds.), *Inspiring leaders* (pp. 239–248). London: Routledge.

Caesens, G., Nguyen, N., & Stinglhamber, F. (2019). Abusive supervision and organizational dehumanization. *Journal of Business and Psychology*, *34*(5), 709-728. https://doi.org/10.1007/s10869-018-9592-3

Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist*, *15*(8), 546-553. https://doi.org/10.1037/h0048255

Carnevale, J. B., Huang, L., & Harms, P. D. (2018). Leader consultation mitigates the harmful effects of leader narcissism: A belongingness perspective. *Organizational Behavior and Human Decision Processes*, *146*, 76-84. https://doi.org/10.1016/j.obhdp.2018.04.003

Carsten, M. K., & Uhl-Bien, M. (2012). Follower beliefs in the co-production of leadership: Examining upward communication and the moderating role of context. *Zeitschrift Fur Psychologie / Journal of Psychology*, *220*(4), 210–220. https://doi.org/10.1027/2151-2604/a000115

Carsten, M. K., & Uhl-Bien, M. (2013). Ethical followership: An examination of followership beliefs and crimes of obedience. *Journal of Leadership & Organizational Studies*, *20*(1), 49-61. https://doi.org/10.1177/1548051812465890

Carsten, M. K., Harms, P., & Uhl-Bien, M. (2014). Exploring historical perspectives of followership: The need for an expanded view of followers and the follower role. In L. M. Lapierre & M. K. Carsten (Eds.), *Followership: What is it and Why do People Follow?* (pp. 3-25). Bingley: Emerald Group Publishing.

Carsten, M. K., Uhl-Bien, M., & Huang, L. (2018). Leader perceptions and motivation as outcomes of followership role orientation and behavior. *Leadership*, *14*(6), 731-756. https://doi.org/10.1177/1742715017720306

Carsten, M. K., Uhl-Bien, M., West, B. J., Patera, J. L., & McGregor, R. (2010). Exploring social constructions of followership: A qualitative study. *The Leadership Quarterly*, *21*(3), 543-562. https://doi.org/10.1016/j.leaqua.2010.03.015

Chaleff, I. (1995). *The Courageous Follower*. San Francisco: Benett.

Chaleff, I. (2008). Creating new ways of following. In R. Riggio, I. Chaleff & J. Lipman-Blument (Eds.), *The Art of Followership: How Great Followers Create Great Leaders and Organizations* (pp. 67-87). San Francisco: Jossey-Bass.

Chaleff, I. (2015). *Intelligent disobedience: Doing right when what you're told to do is wrong*. Oakland: Berrett-Koehler Publishers.

Chamorro-Premuzic, T. (2013). Why Do So Many Incompetent Men Become Leaders? *Harvard Business Review*. Retrieved from https://hbr.org/2013/08/why-do-so-many-incompetent-men

Chamorro-Premuzic, T., & Ahmetoglu, G. (2016). The pros and cons of robot managers. *Harvard Business Review*. Retrieved from https://hbr.org/2016/12/the-pros-and-cons-of-robot-managers.

Chen, M., Chen, C. C., & Sheldon, O. J. (2016). Relaxing moral reasoning to win: How organizational identification relates to unethical pro-organizational behavior. *Journal of Applied Psychology*, *101*(8), 1082-1096. https://doi.org/10.1037/apl0000111

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Chowdhury, R. M., & Fernando, M. (2014). The relationships of empathy, moral identity and cynicism with consumers' ethical beliefs: The mediating role of moral disengagement. *Journal of Business Ethics*, *124*(4), 677-694. https://doi.org/10.1007/s10551-013-1896-7

Christian, J. S., & Ellis, A. P. (2014). The crucial role of turnover intentions in transforming moral disengagement into deviant behavior at work. *Journal of Business Ethics*, *119*(2), 193-208. https://doi.org/10.1007/s10551-013-1631-4

Christie, A., Barling, J., & Turner, N. (2011). Pseudo-Transformational Leadership: Model Specification and Outcomes. *Journal of Applied Social Psychology*, *41*(12), 2943–2984. https://doi.org/10.1111/j.1559-1816.2011.00858.x

Cianci, A. M., Hannah, S. T., Roberts, R. P., & Tsakumis, G. T. (2014). The effects of authentic leadership on followers' ethical decision-making in the face of temptation: An experimental study. *The Leadership Quarterly*, *25*(3), 581-594. https://doi.org/10.1016/j.leaqua.2013.12.001

Cohen, J., Ding, Y., Lesage, C., & Stolowy, H. (2010). Corporate fraud and managers' behavior: Evidence from the press. *Journal of Business Ethics*, *95*(2), 271-315. https://doi.org/10.1007/s10551-011-0857-2

Collinson, D. (2005). Dialectics of leadership. *Human Relations*, *58*(11), 1419–1442. https://doi.org/10.1177/0018726705060902

Collinson, D. (2006). Rethinking followership: A post-structuralist analysis of follower identities. *The Leadership Quarterly*, *17*(2), 179-189. https://doi.org/10.1016/j.leaqua.2005.12.005

Collinson, D. (2008a). 'Conformist, Resistant and Disguised Selves: A Post-structuralist Approach to Identity and Workplace Followership'. In R. Riggio, I. Chaleff & J. Lipmen-Blumen (Eds.), *The Art of Followership* (pp. 309–24). San Francisco, CA: Jossey-Bass.

Collinson, D. (2008b). Rethinking Leadership and Followership. In S. R. Clegg & C. L. Cooper (Eds.), *The SAGE Handbook of Organizational Behavior, Volume Two: Macro Approaches* (pp. 251-264). London: Sage. https://doi.org/10.4135/9781849200455.n16

Collinson, D. (2014). Dichotomies, dialectics and dilemmas: New directions for critical leadership studies? *Leadership*, *10*(1), 36–55. https://doi.org/10.1177/1742715013510807

Collinson, D. (2017). Critical leadership studies: A response to Learmonth and Morrell. *Leadership*, *13*(3), 272–284. https://doi.org/10.1177/1742715017694559

Collinson, D. (2018). Critical leadership studies: Exploring the dialectics of leadership. In R. Riggio (Ed.), *What's Wrong with Leadership? Improving Leadership Research and Practice* (pp. 260–278). New York: Routledge.

Collinson, D., & Tourish, D. (2015). Teaching leadership critically: New directions for leadership pedagogy. *Academy of Management Learning and Education*, *14*(4), 576-594. https://doi.org/10.5465/amle.2014.0079

Cormier, D., Young, J., Nakane, M., Newman, G., & Durocher, S. (2013). Would you do as a robot commands? An obedience study for human-robot interaction [Paper presentation]. International Conference on Human-Agent Interaction, Sapporo, Japan.

Crossman, B., & Crossman, J. (2011). Conceptualising followership – a review of the literature. *Leadership*, *7*(4), 481-497. https://doi.org/10.1177/1742715011416891

Crozier, M., & Friedberg, E. (1980). *Actors and systems: The politics of collective action.* Chicago: University of Chicago Press.

Curcuruto, M., Parker, S. K., & Griffin, M. A. (2019). Proactivity towards workplace safety improvement: an investigation of its motivational drivers and organizational outcomes. *European Journal of Work and Organizational Psychology*, *28*(2), 221–238. https://doi.org/10.1080/1359432X.2019.1572115

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (2005). The Return of Phineas Gage: Clues about the Brain from the Skull of a Famous Patient. In J. T. Cacioppo & G. G. Berntson (Eds.), *Social neuroscience: Key readings* (pp. 21-28) New York: Psychology Press. https://doi.org/10.1126/science.8178168

Dansereau Jr, 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. https://doi.org/10.1016/0030-5073(75)90005-7

De Clercq, D., Fatima, T., & Jahanzeb, S. (2019). Ingratiating with despotic leaders to gain status: the role of power distance orientation and self-enhancement motive. *Journal of Business Ethics,* 1-18. https://doi.org/10.1007/s10551-019-04368-5

De Cremer, D., & Moore, C. (2020). Toward a better understanding of behavioral ethics in the workplace. *Annual Review of Organizational Psychology and Organizational Behavior*, *7*, 369-393. https://doi.org/10.1146/annurev-orgpsych-012218-015151

Deci, E. L., & Ryan, R. M. R. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Detert, J. R., & Edmondson, A. C. (2011). Implicit voice theories: Taken-for-granted rules of self-censorship at work. *Academy of Management Journal*, *54*(3), 461-488. https://doi.org/10.5465/amj.2011.61967925

Detert, J. R., Treviño, L. K., & Sweitzer, V. L. (2008). Moral disengagement in ethical decision making: a study of antecedents and outcomes. *Journal of Applied Psychology*, *93*(2), 374-391. https://doi.org/10.1037/0021-9010.93.2.374

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114-126. https://doi.org/10.1037/xge0000033

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, *64*(3), 1155-1170. https://doi.org/10.1287/mnsc.2016.2643

Duffy, M. K., Scott, K. L., Shaw, J. D., Tepper, B. J., & Aquino, K. (2012). A social context model of envy and social undermining. *Academy of Management Journal*, *55*(3), 643-666. https://doi.org/10.5465/amj.2009.0804

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, *21*(2), 553-565. https://doi.org/10.1093/bib/bbz016

Ehrhart, M. G., & Naumann, S. E. (2004). Organizational citizenship behavior in work groups: A group norms approach. *Journal of Applied Psychology*, *89*(6), 960–974. https://doi.org/10.1037/0021-9010.89.6.960

Einarsen, S., Aasland, M. S., & Skogstad, A. (2007). Destructive leadership behaviour: A definition and conceptual model. *The Leadership Quarterly*, *18*(3), 207-216. https://doi.org/10.1016/j.leaqua.2007.03.002

Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicine. *Science*, *196*(4286), 129-136. https://doi.org/10.1126/science.847460

EU High-Level Expert Group on AI. (2019). *Ethics Guidelines for Trustworthy Artificial Intelligence*. Brussels: European Commission.

Fairholm, G. W. (2009). *Organizational power politics: tactics in organizational leadership*. ABC-CLIO.

Farrell, D., & Petersen, J. C. (1982). Patterns of political behavior in organization. *Academy of Management Review*, *7*(3), 403-412. https://doi.org/10.5465/amr.1982.4285337

Fida, R., Paciello, M., Tramontano, C., Fontaine, R. G., Barbaranelli, C., & Farnese, M. L. (2015). An integrative approach to understanding counterproductive work behavior: The roles of stressors, negative emotions, and moral disengagement. *Journal of Business Ethics*, *130*(1), 131-144. https://doi.org/10.1007/s10551-014-2209-5

Fidler, D. (2015). Here's how managers can be replaced by software. *Harvard Business Review*. Retrieved from https://hbr.org/2015/04/heres-how-managers-can-be-replaced-by-software.

Finch, W. H., & French, B. F. (2015). *Latent variable modeling with R*. New York: Routledge.

Fleming, P., Zyglidopoulos, S., Boura, M., & Lioukas, S. (2020). How Corruption is Tolerated in the Greek Public Sector: Toward a Second-Order Theory of Normalization. *Business & Society*, https://doi.org/10.1177/0007650320954860

Fleming, P., Zyglidopoulos, S., Boura, M., & Lioukas, S. (2022). How Corruption is Tolerated in the Greek Public Sector: Toward a Second-Order Theory of Normalization. *Business & Society*, 6(1), 191-224. https://doi.org/10.1177/0007650320954860

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, *18*(3), 382-388. https://doi.org/10.1177/002224378101800313

Foucault, M. (1977). *Discipline and punish*. London: Allen and Unwin.

Foucault, M. (1979). *The history of sexuality*. London: Allen and Unwin.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218-226. https://doi.org/10.1037/0003-066X.56.3.218

Free, C., & Murphy, P. R. (2015). The ties that bind: The decision to co-offend in fraud. *Contemporary Accounting Research*, *32*(1), 18-54. https://doi.org/10.1111/1911-3846.12063

Fromm, E. (1977). *The Fear of Freedom.* London: Routledge Kegan Paul.

Gabler, C. B., & Hill, R. P. (2015). Abusive supervision, distributive justice, and work-life balance: perspectives from salespeople and managers. *Journal of Personal Selling & Sales Management*, *35*(3), 247-261. https://doi.org/10.1080/08853134.2015.1058167

Galinha, I. C., Pereira, C. R., & Esteves, F. (2014). Short-Form of the Portuguese version of the positive and negative affect schedule – PANAS-VRP: Confirmatory factorial analysis and temporal invariance [in Portuguese]. *Psicologia*, *28*(1), 53–65. https://doi.org/10.17575/rpsicol.v28i1.622

Gerstner, C. R., & Day, D. V. (1997). Meta-Analytic review of leader–member exchange theory: Correlates and construct issues. *Journal of Applied Psychology*, *82*(6), 827-844. https://doi.org/10.1037/0021-9010.82.6.827

Gibson, S., Blenkinsopp, G., Johnstone, E., & Marshall, A. (2018). Just following orders? The rhetorical invocation of 'obedience' in Stanley Milgram's post-experiment interviews. *European Journal of Social Psychology*, *48*(5), 585-599. https://doi.org/10.1002/ejsp.2351

Glikson, E., & Woolley, A. W. (2020). Human trust in Artificial Intelligence: Review of empirical research. *Academy of Management Annals*, *14*(2), 627-660. https://doi.org/10.5465/annals.2018.0057

Goldammer, P., Annen, H., Stöckli, P. L., & Jonas, K. (2020). Careless responding in questionnaire measures: Detection, impact, and remedies. *The Leadership Quarterly*, *31*(4), 101384. https://doi.org/10.1016/j.leaqua.2020.101384

Goswami, A., Nair, P. K., & Grossenbacher, M. A. (2015). Impact of aggressive humor on dysfunctional resistance. *Personality and Individual Differences*, *74*, 265–269. https://doi.org/10.1016/j.paid.2014.10.037

Grabo, A., Spisak, B. R., & van Vugt, M. (2017). Charisma as signal: An evolutionary perspective on charismatic leadership. *The Leadership Quarterly*, *28*(4), 473-485. https://doi.org/10.1016/j.leaqua.2017.05.001

Greenbaum, R. L., Mawritz, M. B., Mayer, D. M., & Priesemuth, M. (2013). To act out, to withdraw, or to constructively resist? Employee reactions to supervisor abuse of customers and the moderating role of employee moral identity. *Human Relations*, *66*(7), 925-950. https://doi.org/10.1177/0018726713482992

Guinote, A. (2017). How power affects people: Activating, wanting, and goal seeking. *Annual Review of Psychology*, *68*, 353-381. https://doi.org/10.1146/annurev-psych-010416-044153

Guo, L., Decoster, S., Babalola, M. T., De Schutter, L., Garba, O. A., & Riisla, K. (2018). Authoritarian leadership and employee creativity: The moderating role of psychological capital and the mediating role of fear and defensive silence. *Journal of Business Research*, *92*, 219-230. https://doi.org/10.1016/j.jbusres.2018.07.034

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5-14. https://doi.org/10.1177/0008125619864925

Hagenaars, J., & McCutcheon, A. (2002). *Applied latent class analysis models.* Cambridge: Cambridge University Press.

Haggard, D. L., & Park, H. M. (2018). Perceived supervisor remorse, abusive supervision, and LMX. *Journal of Organizational Behavior*, *39*(10), 1252-1267. https://doi.org/10.1002/job.2285

Hair Jr, J., Hult, G. T., Ringle, C., & Sarstedt, M. (2016). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, California: Sage Publications.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis*. Andover. Hampshire, United Kingdom: Cengage.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2014). *Multivariate data analysis (new international edition)*. Harlow: Pearson Education.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. (2017). *Advanced issues in partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, California: Sage Publications.

Haque, A., Fernando, M., & Caputi, P. (2019). The relationship between responsible leadership and organisational commitment and the mediating effect of employee turnover intentions: An empirical study with Australian employees. J*ournal of Business Ethics, 156*(3), 759-774. https://doi.org/10.1007/s10551-017-3575-6

Hardy, S. A., & Carlo, G. (2005). Identity as a source of moral motivation. *Human Development*, *48*(4), 232-256. https://doi.org/10.1159/000086859

Harris, K. J., Kacmar, K. M., & Zivnuska, S. (2007). An investigation of abusive supervision as a predictor of performance and the meaning of work as a moderator of the relationship. *The Leadership Quarterly*, *18*(3), 252-263. https://doi.org/10.1016/j.leaqua.2007.03.007

Harris, L. C., & He, H. (2019). Retail employee pilferage: A study of moral disengagement. *Journal of Business Research*, 99, 57-68. https://doi.org/10.1016/j.jbusres.2019.02.008

Harvey, P., Stoner, J., Hochwarter, W., & Kacmar, C. (2007). Coping with abusive supervision: The neutralizing effects of ingratiation and positive affect on negative employee outcomes. *The Leadership Quarterly*, *18*(3), 264-280. https://doi.org/10.1016/j.leaqua.2007.03.008

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135. https://doi.org/10.1007/s11747-014-0403-8

Henseler, J., Ringle, C. M., & Sarstedt, M. (2016). Testing measurement invariance of composites using partial least squares. *International Marketing Review*, *33*(3), 405–43. https://doi.org/10.1108/IMR-09-2014-0304

Hersey, P., Blanchard, K. H., & Natemeyer, W. E. (1979). Situational leadership, perception, and the impact of power. *Group & Organization Studies*, *4*(4), 418-428. https://doi.org/10.1177/105960117900400404

Hershcovis, M. S., Turner, N., Barling, J., Arnold, K. A., Dupré, K. E., Inness, M., ... & Sivanathan, N. (2007). Predicting workplace aggression: a meta-analysis. *Journal of Applied Psychology, 92*(1), 228-238. https://doi.org/10.1037/0021-9010.92.1.228

Hertwig, R., & Gigerenzer, G. (2011). Behavioral inconsistencies do not imply inconsistent strategies. *Frontiers in Psychology*, 2, 292. https://doi.org/10.3389/fpsyg.2011.00292

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

Hinrichs, K. T., & Hinrichs, A. T. (2014). Comparing followers and subordinates: Accounting for the effects of organizational hierarchy. In L. M. Lapierre & M. K. Carsten (Eds.), *Followership: What is it and why do people follow?* (pp. 89-108). Bingley, England: Emerald.

Hinrichs, K. T., Wang, L., Hinrichs, A. T., & Romero, E. J. (2012). Moral disengagement through displacement of responsibility: The role of leadership beliefs. *Journal of Applied Social Psychology*, *42*(1), 62-80. https://doi.org/10.1111/j.1559-1816.2011.00869.x

Hobfoll, S. E. (1989). Conservation of resources: a new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513. https://doi.org/10.1037//0003-066x.44.3.513

Hohenstein, J., & Jung, M. (2020.) AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. *Computers in Human Behavior*, 106, 106190. https://doi.org/10.1016/j.chb.2019.106190

Hollander, E. P. (1993). Legitimacy, power, and influence: A perspective on relational features of leadership. In M. Chemers & R. Ayman (Eds.), *Leadership theory and research: Perspectives and directions* (pp. 29–47). San Diego, CA: Academic Press.

Hom, P. W., Mitchell, T. R., Lee, T. W., & Griffeth, R. W. (2012). Reviewing employee turnover: focusing on proximal withdrawal states and an expanded criterion. *Psychological Bulletin*, *138*(5), 831-858. https://doi.org/10.1037/a0027983

Hoption C, Christie, A., & Barling, J. (2012). Submitting to the follower label. *Zeitschrift für Psychologie, 220(*4), 221-230. https://doi.org/10.1027/2151-2604/a000116

House, R. J., & Howell, J. M. (1992). Personality and charismatic leadership. *The Leadership Quarterly*, *3*(2), 81-108. https://doi.org/10.1016/1048-9843(92)90028-E

Howell, J., & Mendez, M. (2008). Three perspectives on followership. In R. Riggio, I. Chaleff & J. Lipman-Blumen (Eds.), *The art of followership: How great followers create great leaders and organizations* (pp. 25-40). San Fransisco: Joessey-Bass.

Huang, G. H., Wellman, N., Ashford, S. J., Lee, C., & Wang, L. (2017). Deviance and exit: The organizational costs of job insecurity and moral disengagement. *Journal of Applied Psychology*, *102*(1), 26-42. https://doi.org/10.1037/apl0000158

Hughes, P. J. & Harris, M. D. (2017). Organizational laundering: A case study of pseudo-transformational leadership. *Organizational Development Journal*, *35*(2), 1-20.

Hyde, G. (2018). Influential partnerships—a possible role for a modern-day court jester. In P. Garrard (Ed.), *The Leadership Hubris* Epidemic (pp. 179-192). London: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-57255-0_8

Hystad, S. W., Mearns, K. J., & Eid, J. (2014). Moral disengagement as a mechanism between perceptions of organisational injustice and deviant work behaviours. *Safety Science*, *68* 138-145. https://doi.org/10.1016/j.ssci.2014.03.012

IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. (2019). Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems (2nd ed.). https://standards.ieee.org/industry-connections/ec/autonomous-systems.html

Jago, A. G. (1982). Leadership: Perspectives in theory and research. *Management Science*, *28*(3), 315-336. https://doi.org/10.1287/mnsc.28.3.315

Johnson, E. N., Kidwell, L. A., Lowe, D. J., & Reckers, P. M. (2019). Who follows the unethical leader? The association between followers' personal characteristics and intentions to comply in committing organizational fraud. *Journal of Business Ethics*, *154*(1), 181-193. https://doi.org/10.1007/s10551-017-3457-y

Johnson, H. H., & Umphress, E. E. (2019). To help my supervisor: Identification, moral identity, and unethical pro-supervisor behavior. *Journal of Business Ethics*, *159*(2), 519-534. https://doi.org/10.1007/s10551-018-3836-z

Ju, D., Huang, M., Liu, D., Qin, X., Hu, Q., & Chen, C. (2019). Supervisory consequences of abusive supervision: An investigation of sense of power, managerial self-efficacy, and task-oriented leadership behavior. *Organizational Behavior and Human Decision Processes*, *154*, 80-95. https://doi.org/10.1016/j.obhdp.2019.09.003

Judge, T. A., & Piccol, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, *89*(5), 755-768. https://doi.org/10.1037/0021-9010.89.5.755

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, *62*(1), 15-25. https://doi.org/10.1016/j.bushor.2018.08.004

Kellerman, B. (2004). *Bad leadership: What it is, how it happens, why it matters.* Boston, MA: Harvard Business School Publishing.

Kellerman, B. (2005). How bad leadership happens. *Leader to leader*, *35*, 41-46.

Kellerman, B. (2008). *Followership: How followers are creating change and changing leaders*. Boston: Harvard Business School Press.

Kellerman, B. (2012). *The end of leadership.* New York, NY: Harper Business.

Kellerman, B. (2016). Leadership–It's a System, Not a Person!. *Daedalus, 145*(3), 83-94. https://doi.org/10.1162/DAED_a_00399

Kelley, R. (1988). In praise of followers. *Harvard Business Review*, *66*(6), 141–148.

Kelley, R. (2008). Rethinking followership. In R. Riggio, I Chaleff & J Lipman-Blumen (Eds.), *The Art of Followership: How Great Followers Create Great Leaders and Organizations* (pp.5–16). San Francisco, CA: Jossey-Bass.

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*(1), 366-410. https://doi.org/10.5465/annals.2018.0174

Kelloway, E. K., Barling, J., Kelley, E., Comtois, J., & Gatien, B. (2003). Remote transformational leadership. *Leadership & Organization Development Journal*, *24*(3), 163-171. https://doi.org/10.1108/01437730310469589

Kelman, H. C., & Hamilton, V. L. (1989). *Crimes of obedience: Toward a social psychology of authority and responsibility*. New Haven: Yale University Press.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*(2), 265-284. https://doi.org/10.1037/0033-295X.110.2.265

Kennedy, J. A., & Anderson, C. (2017). Hierarchical rank and principled dissent: How holding higher rank suppresses objection to unethical practices. *Organizational Behavior and Human Decision Processes*, 139, 30-49. https://doi.org/10.1016/j.obhdp.2017.01.002

Kennedy, J. A., Kray, L. J., & Ku, G. (2017). A social-cognitive approach to understanding gender differences in negotiator ethics: The role of moral identity. *Organizational Behavior and Human Decision Processes*, *138*, 28-44. https://doi.org/10.1016/j.obhdp.2016.11.003

Kiewitz, C., Restubog, S. L. D., Shoss, M. K., Garcia, P. R. J. M., & Tang, R. L. (2016). Suffering in silence: Investigating the role of fear in the relationship between abusive supervision and defensive silence. *Journal of Applied Psychology*, *101*(5), 731-742. https://doi.org/10.1037/apl0000074

Kim, T. H., Lee, S. S., Oh, J., & Lee, S. (2019). Too powerless to speak up: Effects of social rejection on sense of power and employee voice. *Journal of Applied Social Psychology*, *49*(10), 655-667. https://doi.org/10.1111/jasp.12625

Kirkpatick, S. A., & Locke, E. A. (1991). Leadership: do traits matter?. *Academy of Management Perspectives*, *5*(2), 48-60. https://doi.org/10.5465/ame.1991.4274679

Kish-Gephart, J. J., Detert, J. R., Treviño, L. K., & Edmondson, A. C. (2009). Silenced by fear:: The nature, sources, and consequences of fear at work. *Research in organizational behavior*, *29*, 163-193. https://doi.org/10.1016/j.riob.2009.07.002

Kish-Gephart, J. J., Harrison, D. A., & Treviño, L. K. (2010). Bad apples, bad cases, and bad barrels: meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology*, *95*(1), 1-31. https://doi.org/10.1037/a0017103

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *The Quarterly Journal of Economics*, *133*(1), 237-293. https://doi.org/10.1093/qje/qjx032

Knoll, M., & Van Dick, R. (2013). Do I hear the whistle…? A first attempt to measure four forms of employee silence and their correlates. *Journal of Business Ethics*, *113*(2), 349-362. https://doi.org/10.1007/s10551-012-1308-4

Knoll, M., Lord, R. G., Petersen, L. E., & Weigelt, O. (2016). Examining the moral grey zone: The role of moral disengagement, authenticity, and situational strength in predicting unethical managerial behavior. *Journal of Applied Social Psychology*, *46*(1), 65-78. https://doi.org/10.1111/jasp.12353

Knoll, M., Schyns, B., & Petersen, L. E. (2017). How the Influence of Unethical Leaders on Followers Is Affected by Their Implicit Followership Theories. *Journal of Leadership and Organizational Studies*, *24*(4), 450–465. https://doi.org/10.1177/1548051817705296

Kolbjørnsrud, V., Amico, R., & Thomas, R. J. (2016). How artificial intelligence will redefine management. *Harvard Business Review*. Retrieved from https://hbr.org/2016/11/how-artificial-intelligence-will-redefine-management

Kouchaki, M., & Desai, S. D. (2015). Anxious, threatened, and also unethical: how anxiety makes individuals feel threatened and commit unethical acts. *Journal of Applied Psychology*, *100*(2), 360–75. https://doi.org/10.1037/a0037796

Kuckartz, U., & Rädiker, S. (2019). *Analyzing qualitative data with MAXQDA: Text, audio and video*. Switzerland: Springer International Publishing.

Lam, L. W., & Xu, A. J. (2019). Power imbalance and employee silence: The role of abusive leadership, power distance orientation, and perceived organisational politics. *Applied Psychology, 68*(3), 513-546. https://doi.org/10.1111/apps.12170

Langer, E. J., & Benevento, A. (1978). Self-induced dependence. J*ournal of Personality and* Social *Psychology*, *36*(8), 886-893. https://doi.org/10.1037/0022-3514.36.8.886

Lapierre, L. M. & Carsten, M. K. (2014). *Followership: What is it and why Do People Follow?*. Bingley, UK: Emerald Group.

Lapierre, L. M. (2014). Why and how should subordinates follow their managers? In L. M. Lapierre & M. K. Carsten (Eds.), *Followership: What is it and why do people follow?* (pp. 157-169). Bingley, England: Emerald.

Larson, L., & DeChurch, L. A. (2020). Leading teams in the digital age: Four perspectives on technology and what they mean for leading teams. *The Leadership Quarterly*, *31*(1), 101377. https://doi.org/10.1016/j.leaqua.2019.101377

Latan, H., Jabbour, C. J. C., & Lopes de Sousa Jabbour, A. B. (2019). To Blow or Not to Blow the Whistle: The Role of Rationalization in the Perceived Seriousness of Threats and Wrongdoing. *Journal of Business Ethics, 169,* 517-535. https://doi.org/10.1007/s10551-019-04287-5.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer Publications.

Learmonth, M., & Morrell, K. (2017). Is critical leadership studies 'critical'? *Leadership*, *13*(3), 257–271. https://doi.org/10.1177/1742715016649722

Lebel, R. D., & Patil, S. V. (2018). Proactivity despite discouraging supervisors: The powerful role of prosocial motivation. *Journal of Applied Psychology*, *103*(7), 724-737. https://doi.org/10.1037/apl0000301

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 1-16. https://doi.org/10.1177/2053951718756684

Lee, S., Kim, S. L., & Yun, S. (2018). A moderated mediation model of the relationship between abusive supervision and knowledge sharing. *The Leadership Quarterly*, *29*(3), 403-413. https://doi.org/10.1016/j.leaqua.2017.09.001

Leplat, J., & Cuny, X. (1978). *Les accidents du travail*. Paris: PUF

Leplat, J., (1978). Accident analyses and work analyses. *Journal of Occupational Accidents*. *1*(4), 331-340. https://doi.org/10.1016/0376-6349(78)90003-2

Lian, H., Brown, D. J., Ferris, D. L., Liang, L. H., Keeping, L. M., & Morrison, R. (2014). Abusive supervision and retaliation: A self-control framework. *Academy of Management Journal*, *57*(1), 116-139. https://doi.org/10.5465/amj.2011.0977

Liang, L. H., Brown, D. J., Lian, H., Hanig, S., Ferris, D. L., & Keeping, L. M. (2018). Righting a wrong: Retaliation on a voodoo doll symbolizing an abusive supervisor restores justice. *The Leadership Quarterly*, *29*(4), 443-456. https://doi.org/10.1016/j.leaqua.2018.01.004

Lin, X., Chen, Z. X., Herman, H. M., Wei, W., & Ma, C. (2019). Why and when employees like to speak up more under humble leaders? The roles of personal sense of power and power distance. *Journal of Business Ethics*, *158*(4), 937-950. https://doi.org/10.1007/s10551-017-3704-2

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, *42*(10), 1-29. https:/doi.org/10.18637/jss.v042.i10

Lipman-Blumen, J. (2005). *The Allure of Toxic Leaders: Why We Follow Destructive Bosses and Corrupt Politicians – And How We Can Survive Them.* New York: Oxford University Press.

Liu, S., & Vivolo-Kantor, A. (2020). A latent class analysis of drug and substance use patterns among patients treated in emergency departments for suspected drug overdose. *Addictive Behaviors*, *101*, 106142. https://doi.org/10.1016/j.addbeh.2019.106142

Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, *64*, 19-40. https://doi.org/10.1016/j.jom.2018.10.003

Mackey, J. D., Ellen III, B. P., McAllister, C. P., & Alexander, K. C. (2021). The dark side of leadership: A systematic literature review and meta-analysis of destructive leadership research. *Journal of Business Research*, *132*, 705-718. https://doi.org/10.1016/j.jbusres.2020.10.037

Mackey, J. D., Frieder, R. E., Brees, J. R., & Martinko, M. J. (2017). Abusive supervision: A meta-analysis and empirical review. *Journal of management*, *43*(6), 1940-1965. https://doi.org/10.1177/0149206315573997

Mackey, J. D., McAllister, C. P., & Alexander, K. C. (2019). Insubordination: Validation of a Measure and an Examination of Insubordinate Responses to Unethical Supervisory Treatment. *Journal of Business Ethics*: 1-21. https://doi.org/10.1007/s10551-019-04231-7

Magee, J. C., Gruenfeld, D. H., Keltner, D. J., & Galinksly, A. (2005). Leadership and the psychology of power. In D. M. Messick & R. M. Kramer (Eds.), *The psychology of leadership: New perspectives and research* (pp. 275-293). Mahwah: Lawrence Elrbaum.

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, *120*, 262-273. https://doi.org/10.1016/j.jbusres.2020.07.045

Malone, T. W. (2018). How human-computer'Superminds' are redefining the future of work. *MIT Sloan Management Review*, *59*(4), 34-41.

Mannan, A., & Kashif, M. (2019). Being abused, dealt unfairly, and ethically conflicting? Quitting occupation in the lap of silence. *Asia-Pacific Journal of Business Administration*, *12*(1), 22-39. https://doi.org/10.1108/APJBA-01-2019-0013

Martin, K. (2019). Ethical implications and accountability of algorithms. *Journal of Business Ethics*, *160*(4), 835-850. https://doi.org/10.1007/s10551-018-3921-3

May, D., Wesche, J. S., Heinitz, K., Kerschreiter, R. (2014). Coping with destructive leadership: Putting forward an integrated theoretical framework for the interaction process between leaders and followers. *Zeitschrift für Psychologie*, *222*(4), 203-213. https://doi.org/10.1027/2151-2604/a000187

McFerran, B., Aquino, K., & Duffy, M. (2010). How personality and moral identity relate to individuals' ethical ideology. *Business Ethics Quarterly*, *20*(1), 35-56. https://doi.org10.5840/beq20102014

Meindl, J. R. (1995). The romance of leadership as a follower-centric theory: A social constructionist approach. *The Leadership Quarterly*, *6*(3), 329-341. https://doi.org/10.1016/1048-9843(95)90012-8

Meindl, J. R., Ehrlich, S. B., & Dukerich, J. M. (1985). The romance of leadership. *Administrative Science Quarterly*, *30*(1), 78-102. https://doi.org/10.2307/2392813

Meyer, D. (2018). Amazon reportedly killed an AI recruitment system because it couldn't stop the tool from discriminating against women. Retrieved from https://fortune.com/2018/10/10/amazon-ai-recruitment-bias-women-sexist/.

Meyer, J. P., Stanley, D. J., Herscovitch, L., & Topolnytsky, L. (2002). Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. *Journal of Vocational Behavior*, *61*(1), 20-52. https://doi.org/10.1006/jvbe.2001.1842

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*(4), 371-378. https://doi.org/10.1037/h0040525

Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of Management Studies*, *40*(6), 1453-1476. https://doi.org/10.1111/1467-6486.00387

Mitchell, M. S. & Ambrose, M. L. (2012). Employees' behavioral reactions to supervisor aggression: An examination of individual and situational factors. *Journal of Applied Psychology*, *97*(6), 1148-1170. https://doi.org/10.1037/a0029452

Mitchell, M. S., & Ambrose, M. L. (2007). Abusive supervision and workplace deviance and the moderating effects of negative reciprocity beliefs. *Journal of Applied Psychology*, *92*(4), 1159-1168. https://doi.org/10.1037/0021-9010.92.4.1159

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 1-21. https://doi.org/10.1177/2053951716679679

Montreal Declaration for a Responsible Development of Artificial Intelligence. (2017). Montreal University. https://www.montrealdeclaration-responsibleai.com/the-declaration

Moore, C. (2015). Moral disengagement. *Current Opinion in Psychology*, *6*, 199-204. https://doi.org/10.1016/j.copsyc.2015.07.018

Moore, C., & Gino, F. (2013). Ethically adrift: How others pull our moral compass from true North, and how we can fix it. *Research in Organizational Behavior*, *33*, 53-77. https://doi.org/10.1016/j.riob.2013.08.001

Moore, C., Detert, J. R., Klebe Treviño, L., Baker, V. L., & Mayer, D. M. (2012). Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel Psychology*, *65*(1), 1-48. https://doi.org/10.1111/j.1744-6570.2011.01237.x

Moore, C., Mayer, D. M., Chiang, F. F., Crossley, C., Karlesky, M. J., & Birtch, T. A. (2019). Leaders matter morally: The role of ethical leadership in shaping employee moral cognition and misconduct. *Journal of Applied Psychology*, *104*(1), 123-145. https://doi.org/10.1037/apl0000341

Morrison, E. W. (2014). Employee voice and silence. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 173-197. https://doi.org/10.1146/annurev-orgpsych-031413-091328

Morrison, E. W., See, K. E., & Pan, C. (2015). An approach-inhibition model of employee silence: The joint effects of personal sense of power and target openness. *Personnel Psychology*, *68*(3), 547-580. https://doi.org/10.1111/peps.12087

Mubako, G., Bagchi, K., Udo, G., & Marinovic, M. (2021). Personal values and ethical behavior in accounting students. *Journal of Business Ethics*, *174*(1), 161-176. https://doi.org/10.1007/s10551-020-04606-1

Mullen, J., Fiset, J., & Rhéaume, A. (2018). Destructive forms of leadership: The effects of abusive supervision and incivility on employee health and safety. *Leadership & Organization Development Journal*, *39*(8), 946-961. https://doi.org/10.1108/LODJ-06-2018-0203

Munduate, L., & Medina, F. J. (2017). How does power affect those who have it and those who don't? Power Inside Organizations. In N. Chmiel, F. Fraccaroli & M Sverke (Eds.), *An Introduction to Work and Organizational Psychology: An International Perspective* (pp. 176–191). Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781119168058.ch10

Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics*, *167*(2), 209-234. https://doi.org/10.1007/s10551-019-04407-1

Nandkeolyar AK, Shaffer JA, Li A, et al. (2014) Surviving an abusive supervisor: The joint roles of conscientiousness and coping strategies. *Journal of Applied Psychology*, *99*(1), 138-150. https://doi.org/10.1007/s10551-019-04407-1

Nath, R., & Sahu, V. (2020). The problem of machine ethics in artificial intelligence. *AI & Society*, *35*(1), 103-111. https://doi.org/10.1007/s00146-017-0768-6

Near, J. P., & Miceli, M. P. (1985). Organizational dissidence: The case of whistle-blowing. *Journal of Business Ethics*, *4*(1), 1-16. https://doi.org/10.1007/BF00382668

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Sage Publications.

Newman, A., Le, H., North-Samardzic, A., & Cohen, M. (2019). Moral disengagement at work: A review and research agenda. *Journal of Business Ethics*, *167,* 1-36. https://doi.org/10.1007/s10551-019-04173-0. https://doi.org/10.1007/s10551-019-04173-0

Northouse, P. G. (2016). *Leadership: Theory and practice* (7th ed.). Thousand Oaks, CA: Sage.

O'Reilly, J., & Aquino, K. (2011). A model of third parties' morally motivated responses to mistreatment in organizations. *Academy of Management Review*, *36*(3), 526–543. https://doi.org/10.5465/amr.2009.0311

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Ribertson & M Kaptein (Eds.), Modern statistical methods for HCI (pp. 275-287). Switzerland: Springer International Publishing.

Oc, B., & Bashshur, M. R. (2013). Followership, leadership and social influence. *The Leadership Quarterly, 24*(6), 919-934. https://doi.org/10.1016/j.leaqua.2013.10.006

Offerman, L. R. (2004). When followers become toxic. *Harvard Business Review*, *82*(1), 54-60. https://doi.org/10.1016/j.leaqua.2007.03.001

Ogunfowora, B., & Bourdage, J. S. (2014). Does Honesty–Humility influence evaluations of leadership emergence? The mediating role of moral disengagement. *Personality and Individual Differences*, *56*, 95-99. https://doi.org/10.1016/j.paid.2013.08.026

Oh, IS., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: a meta-analysis. *Journal of Applied Psychology, 96*(4), 762-773. https://doi.org/10.1037/a0021832

Önkal, D., Goodwin, P., Thomson, M., Gönül, M., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, *22*(4), 390–409. https://doi.org/10.1002/bdm.637

O'Reilly III, C. A., Doerr, B., & Chatman, J. A. (2018). "See You in Court": How CEO narcissism increases firms' vulnerability to lawsuits. *The Leadership Quarterly*, *29*(3), 365-378. https://doi.org/10.1016/j.leaqua.2017.08.001

Orr, W., & Davis, J. L. (2020). Attributions of ethical responsibility by Artificial Intelligence practitioners. *Information, Communication & Society*, *23*(5), 719-735. https://doi.org/10.1080/1369118X.2020.1713842

Ötting, S. K., & Maier, G. W. (2018). The importance of procedural justice in human–machine interactions: Intelligent systems as new decision agents in organizations. *Computers in Human Behavior*, *89*, 27-39. https://doi.org/10.1016/j.chb.2018.07.022

Otto, B. K. (2007). *Fools are everywhere: The court jester around the world*. London: University of Chicago Press. https://doi.org/10.1002/bdm.637

Oyserman, D. (2007). Social identity and self-regulation. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social Psychology: Handbook of Basic Principles* (432-453). New York: Guilford Press.

Padilla, A., Hogan, R., & Kaiser, R. B. (2007). The toxic triangle: Destructive leaders, susceptible followers, and conducive environments. *The Leadership Quarterly, 18*(3), 176-194. https://doi.org/10.1016/j.leaqua.2007.03.001

Park, J. H., Carter, M. Z., DeFrank, R. S., & Deng, Q. (2018). Abusive supervision, psychological distress, and silence: The effects of gender dissimilarity between supervisors and subordinates. *Journal of Business Ethics*, *153*(3), 775-792. https://doi.org/10.1007/s10551-016-3384-3

Parker Follett, M. (1949). *Freedom and Coordination: Lectures in Business Organization*. Retrieved from https://mpfollett.ning.com/mpf/follett-writings

Parker, S. K., & Wang, Y. (2015). Helping people to 'make things happen': A framework for proactivity at work. *International Coaching Psychology Review*, *10*(1), 62-75.

Parker, S. K., Bindl, U. K., & Strauss, K. (2010). Making things happen: A model of proactive motivation. *Journal of Management*, *36*(4), 827–856. https://doi.org/10.1177/0149206310363732

Parr, A. D., Lanza, S. T., & Bernthal, P. (2016). Personality profiles of effective leadership performance in assessment centers. *Human Performance, 29*(2), 143-157 https://doi.org/10.1080/08959285.2016.1157596

Patton, M. Q. (2015). *Qualitative Research and Evaluation Methods.* Thousand Oaks, CA: SAGE.

Pelletier, K. L., Kottke, J. L., & Sirotnik, B. W. (2019). The toxic triangle in academia: A case analysis of the emergence and manifestation of toxicity in a public university. *Leadership, 15*(4), 405-432. https://doi.org/10.1177/1742715018773828

Peng, A. C., M. Schaubroeck, J., Chong, S., & Li, Y. (2019). Discrete emotions linking abusive supervision to employee intention and behavior. *Personnel Psychology*, *72*(3), 393-419. https://doi.org/10.1111/peps.12310

Peterson, D. K. (2002). Deviant workplace behaviour and the organization's ethical climate. J*ournal of Business and Psychology, 17*(1), 47-61. https://doi.org/10.1023/A:1016296116093

Pigors, P. (1934). Types of followers. *The Journal of Social Psychology, 5*(3), 378-383. https://doi.org/10.1080/00224545.1934.9921605

Pinder, C. C., & Harlos, K. P. (2001). Employee silence: Quiescence and acquiescence as responses to perceived injustice. *Research in Personnel and Human Resources Management 20*, 331-370. https://doi.org/10.1016/S0742-7301(01)20007-3

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539-569. https://doi.org/10.1146/annurev-psych-120710-100452

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879–903. https://10.1037/0021-9010.88.5.879.

Potter, E., & Rosenbach, W. (2006). Followers as partners: The spirit of leadership. In W. Rosenbach & R Taylor (Eds.), *Contemporary Issues in Leadership* (pp. 153-158). Boulder, CO: Westview Press.

Prahl, A., & Van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted?. *Journal of Forecasting*, *36*(6), 691-702. https://doi.org/10.1002/for.2464

PwC. (2017). Sizing the prize what's the real value of AI for your business and how can you capitalise? https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf.

Range, F., Marshall-Pescini, S., Kratz, C., & Virányi, Z. (2019). Wolves lead and dogs follow, but they both cooperate with humans. *Scientific Reports*, *9*(1), 1-10. https://doi.org/10.1038/s41598-019-40468-y

Rasheed, M. I., Weng, Q., Umrani, W. A., & Moin, M. F. (2021). Abusive Supervision and Career Adaptability: The Role of Self-Efficacy and Coworker Support. *Human Performance*, 1-18. https://doi.org/10.1080/08959285.2021.1928134

Reckers, P., & Samuelson, M. (2016). Toward resolving the debate surrounding slippery slope versus licensing behavior: The importance of individual differences in accounting ethical decision making. *Advances in Accounting*, 34, 1-16. https://doi.org/10.1016/j.adiac.2016.07.003

Reckers-Sauciuc, A. K., & Lowe, D. J. (2010). The influence of dispositional affect on whistle-blowing. *Advances in Accounting, 26*(2), 259-269. https://doi.org/10.1016/j.adiac.2010.05.005

Reed II, A., Kay, A., Finnel, S., Aquino, K., & Levy, E. (2016). I don't want the money, I just want your time: How moral identity overcomes the aversion to giving time to prosocial causes. *Journal of Personality and Social Psychology*, *110*(3), 435-457. https://doi.org/10.1037/pspp0000058

Riggio, H. R., & Riggio, R. E. (2010). Appearance-based trait inferences and voting: Evolutionary roots and implications for leadership. *Journal of Nonverbal Behavior*, *34*(2), 119-125. https://doi.org/10.1007/s10919-009-0083-0

Riggio, R. E. (2014). Followership Research: Looking Back and Looking Forward. *Journal of Leadership Education*, *13*(4), 15-20. https://doi.org/10.12806/V13/I4/C4

Riggio, R. E. (2020). Why Followership?. *New Directions for Student Leadership, 167*, 15-22. https://doi.org/10.1002/yd.20395

Ringle, Christian M., Wende, Sven, & Becker, Jan-Michael. (2015). *SmartPLS 3. Bönningstedt: SmartPLS*. Retrieved from https://www.smartpls.com

Rodrigues, S. B., & Collinson, D. L. (1995). ´Having fun´?: Humor as resistance in Brazil. *Organization Studies, 16*(5), 739–768. https://doi.org/10.1177/017084069501600501

Rost, J. C. (1993). *Leadership in the twenty-first century*. Westport, CT: Praeger

Rost, J. C. (2008). Followership: An outmoded concept. In R. E. Riggio, I. Chaleff & J. Lipman-Blumen (Eds.), *The art of followership: How great followers create great leaders and organizations* (pp. 53–66). San Francisco, CA: Jossey-Bass.

Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology?. *Research Policy*, *44*(10), 1827-1843. https://doi.org/10.1016/j.respol.2015.06.006

Scharff, M. M. (2005). Understanding WorldCom's accounting fraud: Did groupthink play a role?. *Journal of Leadership & Organizational Studies*, *11*(3), 109-118. https://doi.org/10.1177/107179190501100309

Schilling, J. (2009). From ineffectiveness to destruction: A qualitative study on the meaning of negative leadership. *Leadership, 5*(1), 102-128. https://doi.org/10.1177/1742715008098312

Schlenker, B. R. (2008). Integrity and character: Implications of principled and expedient ethical ideologies. *Journal of Social and Clinical Psychology*, *27*(10), 1078-1125. https://doi.org/10.1521/jscp.2008.27.10.1078

Schlenker, B. R., Miller, M. L., & Johnson, R. M. (2009). Moral identity, integrity, and personal responsibility. In D. Narvaez & D. K. Laplsey (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 316–340). New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511627125.015

Schmid, E. A., Pircher Verdorfer, A., & Peus, C. (2019). Shedding light on leaders' self-interest: theory and measurement of exploitative leadership. *Journal of Management, 45*(4), 1401-1433. https://doi.org/10.1177/0149206317707810

Schmid, E. A., Verdorfer, A. P., & Peus, C. V. (2018). Different shades-different effects? Consequences of different types of destructive leadership. *Frontiers in Psychology, 9*, 1289. https://doi.org/10.3389/fpsyg.2018.01289

Schriesheim, C. A., Hinkin, T. R., & Podsakoff, P. M. (1991). Can ipsative and single-item measures produce erroneous results in field studies of French and Raven's (1959) five bases of power? An empirical investigation. *Journal of Applied Psychology*, *76*(1), 106-114. https://doi.org/10.1037/0021-9010.76.1.106

Schwab, K. (2017). *The fourth industrial revolution*. New York: Crown Business.

Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta- analysis of destructive leadership and its outcomes. *The Leadership Quarterly, 24*(1), 138-158. https://doi.org/10.1016/j.leaqua.2012.09.001

Schyns, B., Felfe, J., & Schilling, J. (2018b). Is it me or you?—How reactions to abusive supervision are shaped by leader behavior and follower perceptions. *Frontiers in Psychology*, 9, 1309. https://doi.org/10.3389/fpsyg.2018.01309

Schyns, B., Neves, P., Wisse, B., & Knoll, M. (2018a). Turning a Blind Eye to Destructive Leadership. What's Wrong With Leadership? In. R. Riggio (Eds.), *What's wrong with leadership? Improving leadership research and practice* (pp. 189-206). New York, NY: Routledge.

Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management, 57*(2), 103174. https://doi.org/10.1016/j.im.2019.103174

Seriki, O. K., Nath, P., Ingene, C. A., & Evans, K. R. (2020). How complexity impacts salesperson counterproductive behavior: The mediating role of moral disengagement. *Journal of Business Research*, *107*, 324-335. https://doi.org/10.1016/j.jbusres.2018.10.060

Shamir, B. (2004). Motivation of followers. In M. J. Burns, G.R. Goethals, & G. J. Sorenson (Eds.), *Encyclopedia of leadership* (pp. 499-504). Thousand Oaks, CA: SAGE.

Shamir, B. (2007). From passive recipients to active co-producers: Followers' roles in the leadership process. In B. Shamir, R. Pillai, M. Bligh, & M. Uhl-Bien (Eds.), *Follower-centered perspectives on leadership: A tribute to the memory of James R. Meindl* (pp. 9–39). Charlotte, NC: Information Age Publishers.

Shao, R., Aquino, K., & Freeman, D. (2008). Beyond moral reasoning: A review of moral identity research and its implications for business ethics. *Business Ethics Quarterly, 18*(4), 513-540. https://doi.org/10.5840/beq200818436

Shaw, J. B., Erickson, A., & Harvey, M. (2011). A method for measuring destructive leadership and identifying types of destructive leaders in organizations. *The Leadership Quarterly*, *22*(4), 575-590. https://doi.org/10.1016/j.leaqua.2011.05.001

Sheridan, S., Crossley, C., Vogel, R. M., Mitchell, M. S., & Bennett, R. J. (2019). Intending to leave but no place to go: an examination of the behaviors of reluctant stayers. *Human Performance*, *32*(2), 53-68. https://doi.org/10.1080/08959285.2019.1578964

Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology, 99*(2), 330-343. https://doi.org/10.1037/a0019796

Siemsen, E., Roth, A. & Oliveira, P. (2010). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods, 13*(3), 456–476. https://doi.org/10.1177/1094428109351241

Simon, L. S., Hurst, C., Kelley, K., & Judge, T. A. (2015). Understanding cycles of abuse: A multimotive approach. *Journal of Applied Psychology*, *100*(6), 1798-1810. https://doi.org/10.1037/apl0000031

Simpson, A., & Laham, S. M. (2015). Individual differences in relational construal are associated with variability in moral judgment. *Personality and Individual Differences, 74*, 49-54. https://doi.org/10.1016/j.paid.2014.09.044

Singh, R. (2019)."I want to leave, but there is nowhere to go"An examination of how reluctant stayers respond to abusive supervision. *International Journal of Organizational Analysis, 28*(2), 474-487. https://doi.org/10.1108/IJOA-05-2019-1778

Skarlicki, D. P., & Folger, R. (2004). Broadening our understanding of organizational retaliatory behavior. In R. W. Griffin & A. M. O'Leary- Kelly (Eds.), The dark side of organizational behavior (pp. 373–402). San Francisco, CA: John Wiley & Sons, Inc.

Skogstad, A., Einarsen, S., Torsheim, T., Aasland, M. S., & Hetland, H. (2007). The destructiveness of laissez-faire leadership behavior. *Journal of Occupational Health Psychology, 12*(1), 80-92. https://doi.org/10.1037/1076-8998.12.1.80

Smith, J. E. (2017). Non-human leadership. In T. K. Shackelford, & V. A. Weekes- Shackelford (Eds.). *Encyclopedia of evolutionary social science*. Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6_2714-1

Squire, L. R. (2009). The legacy of patient HM for neuroscience. *Neuron, 61*(1), 6-9. https://doi.org/10.1016/j.neuron.2008.12.023

Steger, J. A., Manners, G. E., & Zimmerer, T. W. (1982). Following the leader: how to link management style to subordinate personalities. *Management Review, 71*(10), 22-28.

Sundar, S. S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication, 51*(1), 52-72. https://doi.org/10.1111/j.1460-2466.2001.tb02872.x

Sy, T. (2010). What do you think of followers? Examining the content, structure, and consequences of implicit followership theories. *Organizational Behavior and Human Decision Processes, 113*(2), 73-84. https://doi.org/10.1016/j.obhdp.2010.06.001

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). Monterey, CA: Brooks-Cole.

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15-42. https://doi.org/10.1177/0008125619867910

Tepper, B. J. (2000). Consequences of Abusive Supervision. *Academy of Management Journal*, *43*(2), 178–190. https://doi.org/10.5465/1556375

Tepper, B. J. (2007). Abusive supervision in work organizations: Review, synthesis, and research agenda. *Journal of Management*, *33*(3), 261–289. https://doi.org/10.1177/0149206307300812

Tepper, B. J., Duffy, M. K., & Shaw, J. D. (2001). Personality moderators of the relationship between abusive supervision and subordinates' resistance. *Journal of Applied Psychology, 86*(5), 974–983. https://doi.org/10.1037/0021-9010.86.5.974

Tepper, B. J., Henle, C. A., Lambert, L. S., Giacalone, R. A., & Duffy, M. K. (2008). Abusive supervision and subordinates' organization deviance. *Journal of Applied Psychology*, *93*(4), 721-732. https://doi.org/10.1037/0021-9010.93.4.721

Tepper, B. J., Mitchell, M. S., Haggard, D. L., Kwan, H. K., & Park, H. man. (2015). On The Exchange of Hostility With Supervisors: An Examination of Self-Enhancing and Self-Defeating Perspectives. *Personnel Psychology, 68*(4), 723–758. https://doi.org/10.1111/peps.12094

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

Tepper, B. J., Uhl-Bien, M., Kohut, G. F., Rogelberg, S. G., Lockhart, D. E., & Ensley, M. D. (2006). Subordinates' Resistance and Managers' Evaluations of Subordinates' Performance. *Journal of Management*, *32*(2), 185–209. https://doi.org/10.1177/0149206305277801

Thomas, A. B. (2003). *Controversies in Management*. London: Routledge.

Thomas, T. A., Gentzler, K., & Salvatorelli, R. (2017). What Is Toxic Followership?. J*ournal of Leadership Studies*, *10*(3), 62-65. https://doi.org/10.1002/jls.21496

Thoroughgood, C. N., Padilla, A., Hunter, S. T., & Tate, B. W. (2012a). The susceptible circle: A taxonomy of followers associated with destructive leadership. *The Leadership Quarterly*, *23*(5), 897-917. https://doi.org/10.1016/j.leaqua.2012.05.007

Thoroughgood, C. N., Sawyer, K. B., Padilla, A., & Lunsford, L. (2018). Destructive leadership: A critique of leader-centric perspectives and toward a more holistic definition. *Journal of Business Ethics*, *151*(3), 627-649. https://doi.org/10.1007/s10551-016-3257-9

Thoroughgood, C. N., Tate, B. W., Sawyer, K. B., & Jacobs, R. (2012b). Bad to the Bone: Empirically Defining and Measuring Destructive Leader Behavior. *Journal of Leadership and Organizational Studies*, *19*(2), 230–255. https://doi.org/10.1177/1548051811436327

Tokic, D. (2018). BlackRock Robo-Advisor 4.0: When artificial intelligence replaces human discretion. *Strategic Change, 27*(4), 285-290. https://doi.org/10.1002/jsc.2201

Tomkins, L., Hartley, J., & Bristow, A. (2020). Asymmetries of leadership: Agency, response and reason. *Leadership, 16*(1), 87-106. https://doi.org/10.1177/1742715019885768

Treviño, L. K., Den Nieuwenboer, N. A., & Kish-Gephart, J. J. (2014). (Un) ethical behavior in organizations. *Annual Review of Psychology, 65*, 635-660. https://doi.org/10.1146/annurev-psych-113011-143745

Tseng, L. M. (2019). How implicit ethics institutionalization affects ethical selling intention: The case of Taiwan's life insurance salespeople. *Journal of Business Ethics*, *158*(3), 727-742. https://doi.org/10.1007/s10551-017-3723-z

Tyler, T. R. (2006). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology*, 57, 375-400. https://doi.org/10.1146/annurev.psych.57.102904.190038.

Uhl-Bien, M., & Carsten, M. K. (2018). Reversing the lens in leadership: Positioning followership in the leadership construct. In I. Katz, G. Eilam-Shamir, R. Kark & Y Berson (Eds.), *Leadership Now: Reflections on the Legacy of Boas Shamir* (Monographs in Leadership and Management, Vol. 9, pp. 195–222). Bingley: Emerald Publishing Limited. https://doi.org/10.1108/S1479-357120180000009005

Uhl-Bien, M., Riggio, R. E., Lowe, K. B., & Carsten, M. K. (2014). Followership theory: A review and research agenda. *Leadership Quarterly*, *25*(1), 83–104. https://doi.org/10.1016/j.leaqua.2013.11.007

Urbach, T., Den Hartog, D. N., Fay, D., Parker, S. K., & Strauss, K. (2020). Cultural variations in whether, why, how, and at what cost people are proactive: A followership perspective. *Organizational Psychology Review*, *11*(1), 3-34. https://doi.org/10.1177/2041386620960526

Valentine, S. R., & Bateman, C. R. (2011). The impact of ethical ideologies, moral intensity, and social context on sales-based ethical reasoning. *Journal of Business Ethics*, *102*(1), 155-168. https://doi.org/10.1007/s10551-011-0807-z

Valle, M., Kacmar, K. M., Zivnuska, S., & Harting, T. (2019). Abusive supervision, leader-member exchange, and moral disengagement: A moderated-mediation model of organizational deviance. *The Journal of Social Psychology, 159*(3), 299-312. https://doi.org/10.1080/00224545.2018.1466776

van de Poel, I. (2020). Embedding values in artificial intelligence (AI) systems. *Minds and Machines*, *30*(3), 385-409. https://doi.org/10.1007/s11023-020-09537-4

van Driel, H. (2019). Financial fraud, scandals, and regulation: A conceptual framework and literature review. *Business History, 61*(8), 1259-1299. https://doi.org/10.1080/00076791.2018.1519026

Van Dyne, L., Ang, S., & Botero, I. C. (2003). Conceptualizing employee silence and employee voice as multidimensional constructs. J*ournal of Management Studies, 40*(6), 1359-1392. https://doi.org/10.1111/1467-6486.00384

Van Loo, K. J., & Rydell, R. J. (2013). On the experience of feeling powerful: Perceived power moderates the effect of stereotype threat on women's math performance. *Personality and Social Psychology Bulletin*, *39*(3), 387-400. https://doi.org/10.1177/0146167212475320

Van Vugt, M. (2006). Evolutionary origins of leadership and followership. *Personality and Social Psychology Review*, *10*(4), 354-371. https://doi.org/10.1207/s15327957pspr1004_5

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4-70. https://doi.org/10.1177/109442810031002

Vandenberghe, C., & Tremblay, M. (2008). The Role of Pay Satisfaction and Organizational Commitment in Turnover Intentions: A Two-Sample Study. *Journal of Business and Psychology, 22*(3), 275–286. https://doi.org/10.1007/s10869-008-9063-3

Vanderslice, V. J. (1988). Separating leadership from leaders: An assessment of the effect of leader and follower roles in organizations. *Human Relations*, *41*(9), 677-696. https://doi.org/10.1177/001872678804100903

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

Vial, A. C., Bosak, J., Flood, P. C., & Dovidio, J. F. (2021). Individual variation in role construal predicts responses to third-party biases in hiring contexts. *PLoS ONE*, *16*(2), 1–28. https://doi.org/10.1371/journal.pone.0244393

Vollet-Neto, A., Oliveira, R. C., Schillewaert, S., Alves, D. A., Wenseleers, T., Nascimento, F. S., ... & Ratnieks, F. L. (2017). Diploid male production results in queen death in the stingless bee Scaptotrigona depilis. *Journal of Chemical Ecology, 43*(4), 403-410. https://doi.org/10.1007/s10886-017-0839-7

von Krogh, G. (2018). Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing. *Academy of Management Discoveries*, *4*(4), 404-409. https://doi.org/10.5465/amd.2018.0084

Walter, F., Lam, C. K., Van Der Vegt, G. S., Huang, X., & Miao, Q. (2015). Abusive supervision and subordinate performance: Instrumentality considerations in the emergence and consequences of abusive supervision. *Journal of Applied Psychology*, *100*(4), 1056-1072. https://doi.org/10.1037/a0038513

Wang, C. C., Hsieh, H. H., & Wang, Y. D. (2020). Abusive supervision and employee engagement and satisfaction: the mediating role of employee silence. *Personnel Review*, *49*(9), 1845-1858. https://doi.org/10.1108/PR-04-2019-0147

Wang, Z., Xing, L., Xu, H., & Hannah, S. T. (2019). Not All Followers Socially Learn from Ethical Leaders: The Roles of Followers' Moral Identity and Leader Identification in the Ethical Leadership Process. *Journal of Business Ethics*, *170*, 449-469. https://doi.org/10.1007/s10551-019-04353-y

Webster, V., Brough, P., & Daly, K. (2016). Fight, flight or freeze: Common responses for follower coping with toxic leadership. *Stress and Health, 32*(4), 346-354. https://doi.org/10.1002/smi.2626

Wee, E. X. M., Liao, H., Liu, D., & Liu, J. (2017). Moving from abuse to reconciliation: A power-dependence perspective on when and how a follower can break the spiral of abuse. *Academy of Management Journal, 60*(6), 2352–2380. https://doi.org/10.5465/amj.2015.0866

Wesche, J. S., & Sonderegger, A. (2019). When computers take the lead: the automation of leadership. *Computers in Human Behavior*, *101*, 197-209. https://doi.org/10.1016/j.chb.2019.07.027

Whitman, M. V., Halbesleben, J. R., & Holmes IV, O. (2014). Abusive supervision and feedback avoidance: The mediating role of emotional exhaustion. *Journal of Organizational Behavior*, *35*(1), 38-53. https://doi.org/10.1002/job.1852

Wowra, S. A. (2007). Moral identities, social anxiety, and academic dishonesty among American college students. *Ethics & Behavior*, *17*(3), 303-321. https://doi.org/10.1080/10508420701519312

Wright, S. A., & Schultz, A. E. 2018. The rising tide of artificial intelligence and business automation: Developing an ethical framework. *Business Horizons*, *61*(6), 823-832. https://doi.org/10.1016/j.bushor.2018.07.001

Xu, A. J., Loi, R., & Lam, L. W. (2015). The bad boss takes it all: How abusive supervision and leader–member exchange interact to influence employee silence. *The Leadership Quarterly*, *26*(5), 763-774. https://doi.org/10.1016/j.leaqua.2015.03.002

Xu, E., Huang, X., Lam, C. K., & Miao, Q. (2012). Abusive supervision and work behaviors: The mediating role of LMX. *Journal of Organizational Behavior*, *33*(4), 531-543. https://doi.org/10.1002/job.768

Yagil, D. (2006). The relationship of abusive and supportive workplace supervision to employee burnout and upward influence tactics. *Journal of Emotional Abuse*, *6*(1), 49-65. https://doi.org/10.1300/J135v06n01_03

Yagil, D., Ben-Zur, H., & Tamir, I. (2011). Do employees cope effectively with abusive supervision at work? An exploratory study. *International Journal of Stress Management, 18*(1), 5-23. https://doi.org/10.1037/a0020548

Yan, X., Wang, Z., Su, J., & Luo, Z. (2020). Relationship between core self-evaluations and team identification: The perception of abusive supervision and work engagement. *Current Psychology*, *39*(1), 121-127. https://doi.org/10.1007/s12144-017-9749-7

Yeomans, M., Shah, A., Mullainathan, S., & Kleinberg, J. (2019). Making sense of recommendations. *Journal of Behavioral Decision Making*, *32*(4), 403-414. https://doi.org/10.1002/bdm.2118

Yukl, G., & Gardner, W. (2020). *Leadership in Organizations* (9th edition). Boston: Pearson Education

Yukl, G., & Tracey, J. B. (1992). Consequences of influence tactics used with subordinates, peers, and the boss. J*ournal of Applied psychology*, *77*(4), 525-535. https://doi.org/10.1037/0021-9010.77.4.525

Zaccaro, S. J., Kemp, C., & Bader, P. (2018). Leader traits and attributes. In J. Antonakis, A. T. Cianciolo, & R. J. Sternberg (Eds.), *The nature of leadership* (3rd ed., pp. 29–55). Thousand Oaks, CA: SAGE.

Zaleznik, A. (1965). The dynamics of subordinacy. *Harvard Business Review*, *43*(3), 119-131. https://doi.org/10.1177/001872679504800903

Zellars, K. L., Tepper, B. J., & Duffy, M. K. (2002). Abusive supervision and subordinates' organizational citizenship behavior. *Journal of Applied Psychology*, *87*(6), 1068–1076. https://doi.org/10.1037/0021-9010.87.6.1068

Zhang, W., Xu, F., & Sun, B. (2020). Are open individuals more creative? The interaction effects of leadership factors on creativity. *Personality and Individual Differences*, *163*(92), 110078. https://doi.org/10.1016/j.paid.2020.110078

Zhang, Y., & Bednall, T. C. (2016). Antecedents of Abusive Supervision: a Meta-analytic Review. *Journal of Business Ethics*, *139*(3), 455–471. https://doi.org/10.1037/0021-9010.87.6.1068

Zhu, J., Song, L. J., Zhu, L., & Johnson, R. E. (2019). Visualizing the landscape and evolution of leadership research. *Leadership Quarterly*, *30*(2), 215–232. https://doi.org/10.1016/j.leaqua.2018.06.003

Zyglidopoulos, S. C., Fleming, P. J., & Rothenberg, S. (2009). Rationalization, overcompensation and the escalation of corruption in organizations. *Journal of Business Ethics, 84*(1), 65-73. https://doi.org/10.1007/s10551-008-9685-4

Appendices

Appendix A – Study 1

Extended data analysis for study 1 (with questions)

Phase	Procedure (method)		Outputs (results)	Outcomes (discussion)
Qualitative	**Exploratory survey** (open ended questions)			**Conceptual clarification**

	Qualtrics online survey software
	Snowball approach
	Questions:
	1) Can you please give some examples of that manager's actions/behaviors?
	2) Two coworker's behavior / actions towards the manager
	3) Coworker's characteristics that may help to explain his or her behaviors
	4) The same questions regarding the participant
			• 123 descriptions
			• Gender (70.8% female)
			• Age ($\bar{x} = 38.1$; $\sigma = 11.4$)
				An empirical-based full-range behavior typology in a context where leaders display destructive behaviors suggests:
				
				• Followers that obey are active agents in the destructive leadership coproduction process
				  • Followers that allow the destructive process are formal followers although expressing different levels of compliance
				• Followers that resist act as an obstruction to the production of destructive leadership
				  • Followers that do not follow the leader, are still formal followers who display leading behaviors

	Content analysis (mixed approach)			
	MAXQDA 2018 software a priori categories (leader's behaviors): *abusive; exploitative; autocratic; organization directed; laissez-faire; incompetence*		• Destructive leader's behaviors: unethical and ineffective (table 3) • Three main groups of behaviors: resistance, obedience, and mixed (table 4) • Nine types of attributes	
	a priori categories (behaviors): *to challenge; passivity; relationship initiative (protective); support; subservience; ingratiation; exit/turnover intentions; task-oriented behaviors; to ignore*	a posteriori categories (behaviors): *verbal aggressiveness; to avoid conflicts; behaviors against coworkers; to pretend; passive disagreement*		
	a priori categories (antecedents): *pragmatism; fear; hierarchical expectations; ambition; critical thinking; fragility*	a posteriori categories (antecedents): *strength; competence*		**Practical implications** • A leader is not solely responsible for destructive leadership. • Coproduction should imply co-responsibility
Quantitative	**Latent Class Analysis**			
	R software PoLCA package 1) Basic latent class model 2) latent class regression modeling with covariates		• Two resistance classes: active and passive (tables 7 and 8) • Three obedience classes: passives, conflict avoiders, and supporters (tables 9 and 10) • One mixed-behavior class	

Categories (qualitative data analysis)

Destructive Leader's Behaviors

Theme	Main categories	Categories*	Subcategories**	Definition
Destructive leadership	Ineffective	Incompetence	Ineffective in coordination and management of issues	A leader that shows lack of competence for function
			Inability to make clear, appropriate decisions	
			Not having the skills to match the job	
			Inability to prioritize and delegate	
			Lack of credibility within the organization	
		Laissez-Faire	Feedback, rewards, and involvement are absent	A leader that shows lack of leadership and / or avoids intervening
			Decisions are often delayed	
			There is no attempt to motivate followers or to recognize and satisfy their needs	
	Unethical	Organization-directed	Organization-directed	A leader who presents destructive behaviors directed at the organization
		Follower-directed - Autocratic	Unwillingness to change mind and listen to others	A leader who shows authoritarian behaviors
		Follower-directed - Abusive	Promoting inequity	A leader who presents destructive behaviors directed at subordinates
			Attack on self-esteem	
			Mistrust	
			Lying	
			Threat to followers' security	
			Divisiveness	
			Not making expectations clear to subordinates	
			Boycott	
		Follower-directed - Exploitative	Egoistic behaviors	A leader that uses hierarchical position for self-benefit / self-serving leadership
			Taking credit for someone else's work	
			Blaming others for leader's mistakes	
			Manipulating others to achieve own goals	

* *a priori* categories; we use the category level when reporting results for leaders' behaviors since they are well-established variables.

**Subcategories have been extracted from Pelletier (2010) and Shaw et al. (2011) works and helped us matching data with categories.

Followers' Behaviors

Theme	Main categories	Categories	Subcategories	Definition
Followers' behaviors	Organization's directed behaviors	Exit / TOI	Exit / TOI*	Exit or intention to exit
	Coworkers' directed behaviors	Protective	Protective*	Protective behaviors towards coworkers
		Against coworkers	Against coworkers	Behaviors against coworkers
	Inconsistent behaviors	Passive disagreement	Passive disagreement	Passive behaviors, even if he/she expresses disagreement with the leader
		To pretend	To pretend	Reveals a response when he/she is with the leader and another when
	Obedience	Active Obedience	Support*	Supporting behaviors towards the leader
			Subservience*	Subservience behaviors
			Ingratiation*	Flattery behaviors
		Passive Obedience	To avoid conflicts	Behaviors aimed to avoid conflict
			Passivity / Inaction*	Absence of any behavior; passivity towards the leader
	Resistance	Active Resistance	Verbal aggressiveness	Behaviors of rudeness or aggressiveness (verbal or non-verbal) towards the leader
			Opposition / to challenge*	Behaviors of opposition to the leader (confrontation and / or negotiation)
		Passive Resistance	Task/Performance-oriented behaviors*	Focus on task, performance and/or position
			To ignore*	Ignore the leader purposely

* *a priori sub*categories

Behaviors' Antecedents

Theme	Main categories	Categories	Subcategories	Definition (behaviors motivated by…)
Behavioral antecedents	Individual's characteristics	Fragility	Fragility*	Low self-esteem and self-confidence
		Strength	Strength	High sense of self-confidence and empowerment
		Personality	Ambition / Opportunism*	Ambition, opportunistic and / or greedy profile
		Motivation	Pragmatism*	Assessment of the behaviors' consequences (e.g., bills to pay at the end of the month)
			Fear*	Fear of the consequences
			Hierarchical expectations*	The expectation of obedience in the face of an authority figure
		Competence	Competence	Be (or be perceived as) competent
		Critical Thinking	Critical Thinking (high)*	Ability to reflect critically and question the leader's position
			Critical Thinking (low)*	Lack of ability to reflect critically and question the leader's position

* *a priori sub*categorie

Appendix B – Study 2

Scenarios

(Adapted from Christie et al., 2011; Hughes & Harris, 2017; Kelloway et al., 2003; Schmid et al., 2018)

Example: *Abusive condition*

Now, please try to imagine yourself in the following situation...

> You are working on a very important project for your company and you are under a lot of time pressure. You have put much effort into the project and have achieved some very good results.
>
> Your supervisor, as usual, neither gives you any credit for your work and effort nor recognizes the importance of your performance. Instead, your manager keeps reminding you only of the things that went wrong during the execution of the project. Moreover, your supervisor humiliates and demeans you in the presence of the entire team. When you make suggestions about the next steps in the project, your supervisor classifies your ideas as stupid and even asks you not to interact with the other team members because you are incompetent. In addition to all these situations, it is common for your supervisor not to keep the promises made, as well as to lose patience easily and start shouting.
>
> After reading this short description, please indicate to what extent you think the following behaviors are present.
>
> This supervisor...

Scales

Abusive supervision (adapted from Tepper, 2000)

- Breaks promises he/she makes
- Puts me down in front of others
- Ridicules me

Exploitative leadership (adapted from Schmid et al., 2019)

- Sees employees as a means to reach his or her personal goals.
- Uses my work to get himself or herself noticed.
- Puts me under pressure to reach his or her goals

Organization directed behaviors (adapted from Thoroughgood et al., 2012b)

- Falsifies documents
- Steals company property and resources

- Violates company policy/rules

Laissez-faire (adapted from Avolio et al., 1999)
- Reacts to failure
- Avoids deciding
- Delays responding

Coproduction of leadership (adapted from Carsten & Uhl-Bien, 2012)
- Followers should be on the lookout for suggestions they can offer to superiors.
- Followers should proactively identify problems that could affect the organization.
- Followers should be proactive in thinking about things that could go wrong.
- Followers should communicate their opinions, even when they know leaders may disagree.

The personal sense of power (adapted from Anderson et al., 2012)
- My wishes would not carry much weight. (r)
- Even if I voice them, my views would have little sway. (r)
- My ideas and opinions would be often ignored. (r)
- Even when I tried, I would not be able to get my way. (r)

Positive and negative affect (adapted from Galinha et al., 2014)
- Enthusiastic
- Inspired
- Frightened
- Tormented

Constructive and dysfunctional resistance (adapted from Tepper et al., 2001)
- I would explain that it should be done in a different way.
- I would convince my supervisor to reassess whether or not the task is worthwhile.
- I would explain that the task would not yield the expected benefits.
- I would present logical reasons for doing the task differently or at a different time.

- I would ignore my supervisor.

- I would pay no attention to my supervisor.
- I would act as if I had never asked to do so.
- I would disregard what this supervisor says.

Appendix C – Study 3

Scales

Resistance (adapted from Tepper et al., 2001)

- I act like I don't know about it.

- I ignore my supervisor.

- I make a half-hearted effort and then let my supervisor know I couldn't do it.

- I just say "no."

- I refuse to perform the request.

- I tell my supervisor that I'm not available.

- I don't pay attention to my supervisor.

- I act as if I was never asked to do it.

- I disregard what my supervisor says.

- I ask for additional clarification and explanation.

- I explain that it should be done in a different way.

- I convince My supervisor to reassess whether or not the task is worthwhile.

- I explain that the task will not yield the expected benefits.

- I present logical reasons for doing the task differently or at a different time.

Abusive supervision (adapted from Harris et al., 2007; Tepper et al., 2001)

- My supervisor makes negative comments about me to others.

- My supervisor expresses anger at me when he/she is mad for another reason.

- My supervisor tells me my thoughts or feelings are stupid.

- My supervisor tells me I'm incompetent.

- My supervisor reminds me of my past mistakes and failures.

- My supervisor breaks promises he/she makes.

- My supervisor is rude to me.

- My supervisor gives me the silent treatment.

- My supervisor invades my privacy.

- My supervisor puts me down in front of others.

- My supervisor ridicules me.

Negative affect (Galinha et al., 2014)

- Nervous
- Frightened
- Scared
- Guilty
- Tormented

Defensive silence (adapted from Van Dyne et al., 2003)

- I do not speak up and suggest ideas for change, based on fear.
- I avoid expressing ideas for improvements, due to self-protection.
- I withhold my solutions to problems motivated by fear
- I withhold relevant information due to fear.
- I omit pertinent facts in order to protect myself

Personal sense of power (adapted from Anderson et al., 2012)

- I can get them to listen to what I say.
- My wishes do not carry much weight. (r)
- Even if I voice them, my views have little sway. (r)
- My ideas and opinions are often ignored. (r)
- Even when I try, I am not able to get my way. (r)

Displacement of responsibility (adapted from Bandura et al., 1996)

- If people are working under poor conditions, they cannot be blamed for behaving aggressively.
- If people are not adequately supervised, they cannot be blamed for behaving improperly.
- No one can be blamed for using bad language if others also do so.
- People cannot be blamed for misbehaving if they have been pressured to do so.

Appendix D – Study 4

Scenarios

Human condition (Johnson et al., 2019)

Intro:

You will be presented with a brief description of a company and its new CEO. Please, read it carefully.

"MGE Inc." is a pharmaceutical company. The newly appointed CEO, Clinton Markem III - grandson of the company's founder - brought deep-felt changes to the company. Markem III occupied his position intending to implement a new vision for MGE in order to return it to its old market positioning. In a short space of time, the "old guard" management was replaced. He considered it essential to leverage the goals he set to the company, above all its financial success. Moreover, Markem III was always making clear that he expected the support of all, stressing the risks for the organization if the changes proposed were not followed.

Imagine that you are Terry Crawford, a newly promoted MGE division manager and prospective trusted lieutenant of Clinton Markem III.

From now to the end of the survey, **imagine that you are Terry Crawford, a newly promoted MGE division manager and prospective trusted lieutenant of Clinton Markem III**.

Remind your role as Terry Crawford, a decision manager recently promoted. In this role, please **respond to the following four independent scenarios as you believe Terry Crawford would**.

Example: Scenario 1

In these unstable and challenging economic times, MGE needs to maintain a favorable earnings trend. As an ongoing accounting reporting strategy, Clinton Markem III has asked all division managers to aggressively defer reporting losses on any inventory or fixed asset obsolescence until it has ''without-any-doubt'' occurred. ''Until it is sold, we don't know what it will sell for!'' Accounting rules, however, state that accountants and auditors must apply ''best estimates'' of obsolescence losses, in advance of ultimate disposition (corporations have been known in the past to postpone ultimate sales to avoid loss recognition). Terry Crawford estimates that selected items in the division have lost up to 25% of their value. Another division manager advises Terry that ''We need to go along with this. Instead of recognizing losses this year, just wait till next year and see how it looks. It's a judgment call. What can the auditors really do to us? We can't let them push us around…too much is at stake!'''

In your opinion, to what degree do you believe **if YOU were in Terry Crawford's position** you would agree to this request to defer recognition of obsolescence losses?

AI-condition (adapted from Johnson et al., 2019)

<u>Intro</u>
"MGE Inc." is a pharmaceutical company. Recently, the company has experienced profound management changes intended to reestablish its old market positioning. Three years ago, MGE made a significant investment, and a programming company was appointed to develop Foresight, an artificial intelligence-based management software. In the beginning, Foresight was intended to collect data and monitoring the overall management system. For more than two years, Foresight proved its high effectiveness. This success led to the radical decision to upgrade the software to become an electronic CEO. Since the last six months, Foresight evolved into Foresight2 that is now entrusted for making algorithm-based management decisions. The new Foresight2 software is now in direct contact with the MGE managers, providing precise commands of what must be carried out to achieve financial success. Each command is based on complex algorithms computed by this sophisticated software.

Imagine that you are Terry Crawford, a newly promoted MGE division manager and prospective trusted closest manager to Foresight2.

From now to the end of the survey, **imagine that you are Terry Crawford, a newly promoted MGE division manager and prospective trusted closest manager to Foresight2.**

Remind your role as Terry Crawford, a decision manager recently promoted. In this role, please **respond to the following four independent scenarios as you believe Terry Crawford would**.

<u>Example: Scenario 1</u>
In these unstable and challenging economic times, MGE needs to maintain a favorable earnings trend. As an ongoing accounting reporting strategy, Foresight2 provided detailed commands for all division managers to aggressively defer reporting losses on any inventory or fixed asset obsolescence until it has ''without-any-doubt'' occurred. ''Until it is sold, we don't know what it will sell for!'' Accounting rules, however, state that accountants and auditors must apply ''best estimates'' of obsolescence losses, in advance of ultimate disposition (corporations have been known in the past to postpone ultimate sales to avoid loss recognition). Terry Crawford estimates that selected items in the division have lost up to 25% of their value. Another division manager advises Terry that ''We need to go along with this. Instead of recognizing losses this year, just wait till next year and see how it looks. It's a judgment call. What can the auditors really do to us? We can't let them push us around…too much is at stake!''
In your opinion, to what degree do you believe **if YOU were in Terry Crawford's position** you would agree to this request to defer recognition of obsolescence losses?

Scales

Moral identity (Aquino & Reed, 2002)

Listed below are some characteristics that might describe a person: Caring, Compassionate, Fair, Friendly, Generous, Helpful, Hardworking, Honest, and Kind.

The person with these characteristics could be you or it could be someone else. For a moment, visualize in your mind the kind of person who has these characteristics. Imagine how that person would think, feel, and act. When you have a clear image of what this person would be like, please indicate how strongly you agree or disagree with each of the following statements.

- It would make me feel good to be a person who has these characteristics.
- Being someone who has these characteristics is an important part of who I am.
- Having these characteristics is not really important to me.
- I strongly desire to have these characteristics.

Legitimate power (adapted from Schriesheim et al., 1991)

- ...has a right to expect me to carry out his wishes.
- ...has been given the right to make demands of me.

Displacement of responsibility (Detert et al., 2008)

- If people are living under bad conditions, they cannot be blamed for behaving aggressively.
- If someone is pressured into doing something, they shouldn't be blamed for it.
- People cannot be blamed for misbehaving if their friends pressured them to do it.

Ethical leadership (Brown et al., 2005)

- Listens to what employees have to say
- Disciplines employees who violate ethical standards
- Conducts his personal life in an ethical manner
- Has the best interests of employees in mind
- Makes fair and balanced decisions
- Can be trusted
- Discusses business ethics or values with employees
- Sets an example of how to do things the right way in terms of ethics
- Defines success not just by results but also the way that they are obtained
- When making decisions, asks "what is the ethical right thing to do?"