TECHNISCHE UNIVERSITÄT MÜNCHEN

TUM School of Management

Professur für Forschungs- und Wissenschaftsmanagement

Power in Organizations: Empirical and Theoretical Advances

Ulf Steinberg

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Philosophie (Dr. phil.) genehmigten Dissertation.

Vorsitzende/r: Prof. Dr. Michael Kurschilgen

Prüfende der Dissertation: 1. Prof. Dr. Claudia Peus

2. Prof. Dr. Peter Fischer

Die Dissertation wurde am 18.03.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15.07.2019 angenommen.

ACKNOWLEDGEMENT

Many supported me in the process of writing my dissertation and I am grateful to all of those who did. First, I would like to thank you, Prof. Dr. Claudia Peus, for accepting me as a doctoral student, for being patient, and for providing room to discuss even surprising ideas. I would also like to thank you, Prof. Dr. Peter Fischer, for providing your honest views in discussions about the dissertation and Prof. Dr. Michael Kurschilgen for heading the dissertation committee.

I thank all my collaborators and co-authors for their continued support and their invaluable contributions. To you, Dr. Kristin Knipfer, I would like to express my gratitude for your supervision and mentorship that helped me grow academically, for never tiring to give feedback and for always having a word of encouragement.

An important part of the experience of my dissertation was working with a great team of researchers at the Chair of Research and Science Management. Thank you, Maxim Egorov, for being a friend. Thank you to all who supported me in their very own way: Susanne Braun, Anna Brzykcy, Jakub Cichor, Leidy Cubillos Pinilla, Regina Dutz, Franziska Emmerling, Xuan Feng, Brooke Gazdag, Elke Graf, Tanja Hentschel, Lisa Horvath, Sylvia Hubner, Anna Kraft, Hannah Kunde, Kerstin Löwe, Sebastian Mangold, Armin Pircher-Verdorfer, Stephanie Rehbock, Ellen Schmid, and Selina Stracke.

Many good friends of mine were an important source of strength in the years of writing my dissertation and beyond. Thank you for all the discussions, the great experiences we shared and continue to share, and for putting everything back into perspective every once in a while. Thank you, Rouven Kanitz, for going through this whole process with me as a friend.

Without my family I would not have been able to finish my dissertation and may not even have started it in the first place. Thank you so much for giving me your unconditional support, for providing me with opportunities for growth, and for being open-minded. This work is dedicated to you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

In everyday conversations and articles in the popular press about work, power often serves as an explanation for behavior in organizations. Reliable experimental evidence allowing for the conclusion that power causes certain organizational behaviors, however, is scarce (Fast, Sivanathan, Mayer, & Galinsky, 2012). The present dissertation aims at clarifying when, how, and why it is important to consider power in research on management and organizational behavior. Specifically, two experimental studies and a conceptual study serve to analyze the applicability of social psychological conceptions of power to research on management and organizational behavior.

The research reported in chapter two addresses the question if power influences time planning in projects. This line of research features a novel combination of approaches from management science and psychology, a decision task with a normatively optimal solution, and real-life consequences for the participants in the form of incentives. Two preregistered experiments show that cost considerations, but not psychological power, influence time planning in projects. Exploratory analyses in chapter two demonstrate that the presence of an effect is contingent on the representation of participants as a random effect in the statistical model.

The experiments reported in chapter three examine whether power influences the evaluation of the relevance of different types of arguments. This line of research aims at testing predictions derived from two theories of power about how power influences the evaluation of arguments about the desirability and feasibility of projects. The results show that high-power individuals favor positively formulated arguments, irrespective of whether they concern the desirability or feasibility of projects. Low-power individuals, on the other hand, favor feasibility arguments over desirability arguments, irrespective of whether they are positively or negatively formulated.

The conceptual research in chapter four analyzes the relationship of digitized performance management and power. Specifically, a novel conceptualization of power, power concentration, is proposed to capture the influence of digitization on performance management in organizations. Power concentration describes how power concentrates in one individual or a small group of individuals over time, thus awarding greater opportunities for goal achievement (Pratto, 2016). Departing from this notion, propositions are developed to guide research on the relationship and consequences of digitized performance management and power concentration on multiple levels of research.

ZUSAMMENFASSUNG (GERMAN ABSTRACT)

In Gesprächen und Zeitungsartikeln über die Arbeit wird Macht oft als ein Faktor herangezogen, der das Verhalten von Angestellten und Führungskräften in Organisationen erklären kann. Für diese Ansicht gibt es jedoch nur wenige reliable wissenschaftliche Belege (Fast, Sivanatan, Mayer, & Galinsky, 2012). Die vorliegende Dissertation zielt darauf ab, unser Verständnis zu verbessern, wann, wie und warum es wichtig ist, Macht in der Forschung zu Management und Organizational Behavior zu betrachten. Dazu wird in zwei experimentellen Studien und einer konzeptuellen Arbeit die Anwendbarkeit von sozialpsychologischen Machtkonzeptionen in der Forschung zu Management und Organizational Behavior untersucht.

Die Forschung, über die in Kapitel zwei berichtet wird, beschäftigt sich mit der Frage, ob Macht die Zeitplanung in Projekten beeinflusst. Dabei kommen eine neuartige Kombination von Ansätzen aus den Management-Wissenschaften und der Psychologie, eine Entscheidungsaufgabe mit einer aus normativer Sicht optimalen Lösung und reale Konsequenzen für die Teilnehmenden in Form von leistungsabhängiger Entlohnung zum Einsatz. Zwei präregistrierte Experimente zeigen, dass Kostenerwägungen, aber nicht psychologische Macht, die Zeitplanung in Projekten beeinflussen. Explorative Analysen in Kapitel zwei zeigen, dass ein signifikanter Einfluss von Macht von der Frage abhängt, ob die Teilnehmenden in der statistischen Analyse als Zufallsfaktor konzeptualisiert werden.

In den Experimenten in Kapitel drei wird untersucht, ob Macht die Evaluation der Relevanz von verschiedenen Typen von Argumenten beeinflusst. Diese Forschungslinie zielt darauf ab, Vorhersagen von zwei unterschiedlichen Macht-Theorien über den Einfluss von Macht auf die Bewertung von Erwünschtheits- und Machbarkeitsargumenten zu testen. Die Ergebnisse zeigen, dass Individuen mit hoher Macht positiv formulierte Argumente bevorzugen, unabhängig davon ob sie die Erwünschtheit oder die Machbarkeit eines

Projektes betreffen. Die Ergebnisse zeigen weiter, dass Individuen mit niedriger Macht Machbarkeitsargumente Erwünschtsheitsargumenten vorziehen, unabhängig davon, ob sie positiv oder negativ formuliert sind.

Die konzeptuelle Arbeit in Kapitel vier analysiert die Beziehung von digitalem Performance Management und Macht. Dazu wird eine neue Konzeption von Macht, Machtkonzentration, vorgeschlagen, die den Einfluss der Digitalisierung auf das Performance Management in Organisationen abbilden kann. Machtkonzentration beschreibt, wie sich Macht in Form von Möglichkeiten, Ziele zu erreichen (Pratto, 2016) über die Zeit in einzelnen Individuen oder relativ kleinen Gruppen von Individuen konzentriert. Ausgehend von diesem Gedanken werden Propositionen entwickelt, die als Orientierung für Forschung zu der Beziehung von und den Auswirkungen von digitalem Performance Management und Machtkonzentration auf mehreren Analyseebenen dienen können.

CHAPTER 1

"[…] the fundamental concept in social science is Power, in the same sense in which Energy is the fundamental concept in physics." – Bertrand Russell[1]

1. Introduction and Research Question

While the importance of power for research on organizational behavior appears to be self-evident, empirical evidence specifically linked to problems in research on management and organizational behavior is limited (Fast et al., 2012). It is probably not unusual to explain a manager's behaviors with reference to power in private conversations about work or in the popular press. Indeed, if your boss planned shorter times for your project or if he didn't seem to be all too interested in your cautionary notes about the project's feasibility, explanations alluding to how power affects people would seem natural. Still, robust and reliable empirical evidence to evaluate this seemingly natural way to explain a manager's behavior is scarce (Fast et al., 2012; Sturm & Antonakis, 2015). In this dissertation, I report novel empirical and conceptual research that draws on research from social psychology, management science, and sociology to advance the understanding of the role of power in organizational behavior.

One reason for the relative lack of reliable empirical evidence for the importance of power in research on organizational behavior may be the fact that the literature on power spans several scientific disciplines from philosophy, sociology, psychology, and management science to biology (Anderson & Brion, 2014; Fleming & Spicer, 2014; Galinsky, Rucker, & Magee, 2015; K. M. Smith & Apicella, 2016). Even though there is some overlap and integration, this diversity partly accounts for the relative lack of unequivocal empirical evidence (cf. Neely, 2005). Ultimately, however, robust and replicable empirical evidence

[1] (Russell, 1957, p. 10)

will be required to decide whether the concept power is actually important for research on management and organizational behavior and whether interventions and policy advice should be derived from this stream of research (Antonakis, 2017). Among the diverse epistemic approaches, common in the disciplines involved in the study of power, only the experiment allows for making causal claims (e.g., Antonakis, 2017). As a consequence, over the last two decades, experiments on power have been at the center of both, a fruitful stream of research (Anderson & Brion, 2014; Galinsky et al., 2015) and a lively scientific debate (Pratto, 2016; Sturm & Antonakis, 2015; Tost, 2015).

The empirical research in this dissertation relies on a definition of power as a sense of power, that is, a subjective sense of being powerful based on asymmetric control over valued resources (Galinsky et al., 2015; Keltner, Gruenfeld, & Anderson, 2003; Pratto, 2016). The stream of research built on this definition departs from the notion that the realization of being powerful, rather than an objective assessment of power, is what actually matters for explaining the actions of powerholders (Galinsky, Gruenfeld, & Magee, 2003). This definition lead to the discovery of many effects of power in social psychology (Galinsky et al., 2015), such as, that power leads to a propensity for action (Galinsky et al., 2003) or more abstract thinking (P. K. Smith & Trope, 2006), and has been applied to research in organizational behavior (Anderson & Brion, 2014; Tost, 2015). In spite of the growing body of research on power, specific questions, for instance, with regard to the influence of power on project management, have not been answered yet (Stingl & Geraldi, 2017). In addition, the conceptualization of power as a sense of power and some methodological aspects of the experiments on power have been criticized (Pratto, 2016; Sturm & Antonakis, 2015).

In this dissertation, I report novel experimental and conceptual research to contribute to answering the questions of when, how, and why it is important to consider power in research on management and organizational behavior. Specifically, I investigate the

previously unacknowledged role of power in two common project management scenarios and I develop theory to advance future empirical research on power and performance management. Gaining a more detailed understanding of the role of power in project management (Stingl & Geraldi, 2017) and performance management is important for research on organizational behavior, because organizational behavior is determined by both projects, as a form of organization (Hobday, 2000), and performance management, as a management process aimed at eliciting desired behaviors (Neely, Gregory, & Platts, 2005; Stone, Deadrick, Lukaszewski, & Johnson, 2015). Identifying the role of power as a cause and mechanism of organizational behavior will then help to design management systems and interventions that allow organizations to achieve their goals more successfully.

In the following paragraphs, I briefly outline the chapters that report empirical and conceptual research, and I detail the dissertation's contributions to research. The rest of this dissertation is organized as follows: Chapter two reports experiments on the influence of power on time planning, chapter three reports experiments on the influence of power on ratings of relevance of different types of information, and chapter four reports conceptual research on the relationship of digitized performance management and power concentration. Additional materials can be found in the appendix.

1.1. The Influence of Power on Time Planning

The second chapter focuses on the influence of cost considerations and power on time planning in projects. It is easy to imagine how costs influence time planning in projects and there are models available from operations research that can be readily applied to address the influence of costs on planning (Schiffels, Fügener, Kolisch, & Brunner, 2014). It may be more difficult to imagine, however, how a manager's or leader's power influences time planning in projects and there are no models to address this influence yet (Stingl & Geraldi, 2017). Indeed, no studies on the influence of power in time planning in projects are available,

despite explicit calls for this type of research (Stingl & Geraldi, 2017) and the fact that predictions can be derived from current theories of power. The second chapter reports experiments designed to answer the question of how power influences time planning in projects that others execute. Answering this question advances theory in psychology by applying recommendations from the experimental rigor debate, advances theory in management science by testing the influence of an individual variable on time planning (as opposed to situational variables that have been tested extensively), and has practical relevance due to the costs associated with planning projects too long or too short.

Previous empirical research on power and time planning found that high-power individuals planned shorter times for tasks they execute themselves (Weick & Guinote, 2010) and that they acted under the impression of having greater control over time (Moon & Chen, 2014). There was no empirical evidence in the literature on how power influences time planning for tasks that others execute, which is much more common in project management than planning times for one's own tasks. I derived predictions about the influence of power on time planning in projects from the social distance theory of power (Magee & Smith, 2013). The social distance theory of power (Magee & Smith, 2013) is based on construal level theory (Trope & Liberman, 2010) and establishes a link between the powerfuls' perceptions of social distance and the degree of abstraction in their thinking (cf. Berson, Halevy, Shamir, & Erez, 2015). According to this line of reasoning, high-power individuals will focus on high-level concerns, such as, winning a project tender or finishing a project in time, rather than on low-level cost-related (Fiedler, 2007) concerns. Two experiments aimed to test these predictions by employing a power manipulation from psychology and a time-planning task with performance-based incentives adapted from management science.

The research reported in chapter two responds to several calls for research on time planning that combines methods from management science and psychology (Crusius, van

Horen, & Mussweiler, 2012; Fiedler, 2007; Halkjelsvik & Jørgensen, 2012; Katsikopoulos & Gigerenzer, 2013; Stingl & Geraldi, 2017; Svejvig & Andersen, 2015). The combination of these approaches has also been called for specifically with regard to power, as it is thought to address several issues that have been noted as potential problems with the predominant experimental paradigm of the last two decades of power research (Sturm & Antonakis, 2015). The experiments on the influence of power on time planning in projects contribute to the literature by applying recommendations from the experimental rigor debate, such as, employing a dependent variable with a normative optimum and performance-based incentives (Sturm & Antonakis, 2015), and by adding a focus on choices in data analysis to the debate. Recently, a number of recommendations on how to conduct experimental research on power, such as employing an experimental task with real-life consequences for the participants, have been published (Sturm & Antonakis, 2015) that did not receive empirical support themselves yet. Furthermore, the present research extends the experimental rigor debate by contrasting traditional ways to design experiments and analyze data, between-subjects ANOVA designs, with mixed-effects designs and linear mixed models that conceptualize participants as random effect (Baayen, Davidson, & Bates, 2008; Barr, Levy, Scheepers, & Tily, 2013; Judd, Westfall, & Kenny, 2012).

1.2. The Influence of Power on Information Evaluation

The third chapter reports experiments about the influence of power on the evaluation of arguments for or against the pursuit of specific projects. Power is considered to be a very important, but under-studied, determinant of organizational decision-making (Fast et al., 2012). To decide what projects an organization should pursue, information about the potential projects has to be aggregated to serve as a basis for decision making for managers and leaders. The aggregated information needs to respond to two general questions: First, do we want to do the project, that is, is it desirable, and second, do we have the necessary

capabilities, that is, is it feasible (Liberman & Trope, 1998; Magee & Smith, 2013)? The research reported in chapter three therefore aims to answer the question whether power influences what type of information managers and leaders think is relevant for making a decision: desirability, undesirability, feasibility, or unfeasibility information (Magee & Smith, 2013). Finding an answer to this question has important implications for theories of power, because it identifies boundary conditions for two major theories of power (Magee & Smith, 2013), and high practical relevance – think of the economic consequences of poor decision-making for organizations and society as a whole.

According to Magee and Smith (2013), empirical studies tested the influence of power on loss aversion (Inesi, 2010), which reflects (un-) desirability concerns, and on goal-constraining vs. goal-facilitating information (Whitson et al., 2013), which reflects (un-) feasibility information (Magee & Smith, 2013). No study employed a fully crossed design to test the influence of power on all four types of information yet (Magee & Smith, 2013). Predictions about what type of information appears relevant to high-power individuals can be derived from two different theories of power (Magee & Smith, 2013). On the one hand, the approach/inhibition theory of power (Keltner et al., 2003) predicts that high-power individuals rate positive information more relevant than negative information, regardless of domain (Magee & Smith, 2013). On the other hand, the social distance theory of power predicts that high-power individuals rate desirability information more relevant than feasibility information, regardless of valence (Magee & Smith, 2013).

Four experiments and a qualitative study tested the predictions I derived from the approach/inhibition theory, as well as from the social distance theory. The five studies featured different experimental manipulations of power, student as well as professional samples, and were conducted in the laboratory and online. A single-paper meta-analysis was

conducted to test the central hypotheses and the inclusion of a qualitative study provided indications that the central effects are robust across different dependent variables.

The research reported in chapter three contributes to the literature on organizational decision-making in several important ways. First, it explains what type of information managers consider to be relevant in making decisions about future projects. Second, it highlights that undesirability information may be overlooked systematically and thus points to possible interventions. Third, it identifies boundary conditions for the applicability of different theories of power to the problem of information evaluation in organizations.

1.3. Digitized Performance Management and Power Concentration

The question underlying the conceptual research reported in chapter four is: How and why does the design of digitized performance management affect power concentration on multiple levels of analysis of research on management and organizational behavior? Addressing this question is very important since performance management affects virtually every individual in the workplace and currently there is little to no guidance on how to design digitized performance management to leverage its potentials, instead of cementing its downsides.

To further the integration of the literatures on power and performance management, I conceptually analyze the interplay between design choices in digitized performance management and power concentration, notably with regard to the realization of transparency, accountability, and legitimacy. From this analysis I derive propositions to guide future research on digitized performance management and power concentration on different levels of analysis, as well as initial design advice for practitioners seeking to design digitized performance management systems.

I use the term digitized performance management to describe performance management (Neely et al., 2005; Stone et al., 2015) in its social context, characterized by an

increased appetite for evaluation (Dahler-Larsen, 2012), and its technological context, characterized by digitization (Brynjolfsson & McAfee, 2014; Loebbecke & Picot, 2015; Newell & Marabelli, 2015). In chapter four, I introduce the concept *power concentration*, defined as the active or passive accumulation of power in one individual or a relatively small group of individuals over time. For instance, many app-based services offer the buyer the opportunity to rate the seller, but also offer the seller the opportunity to rate the buyer, which effectively re-distributes power between them and thus reduces power concentration on the interpersonal level. This very process, however, increases power concentration on the organizational level, since the app provider gains power over their customers, who previously had power over the provider.

As this example indicates, performance management is currently undergoing profound changes and many consider this to have been one of the most important developments in the workplace in the past decade (Levy, Tseng, Rosen, & Lueke, 2017). Nevertheless, there is not much guidance on how to design digitized performance management systems in the scientific literature (but see Neely et al., 2005). Power research generated important insights, which could help devising performance management system designs that realize the potential benefits of digitization. However, the literatures on power and performance management are not well integrated (but see Townley, 1993).

The conceptual research in chapter four makes several important contributions. It serves as a guide for research on digitization in an important area of management and organizational behavior. It thus responds to calls for greater contextualization of research on performance management (Levy & Williams, 2004). In addition, it links two streams of research, considers multiple levels of analysis, and has practical implications by offering initial design advice for digitized performance management systems.

1.4. Contributions

The present dissertation contributes to the scientific literature on power in five important ways. First, across different studies, it identifies boundary conditions for power research, both in terms of theory and in terms of methodology. Second, by comparing predictions derived from different theories of power, it helps consolidating the field of power research. Third, by proposing a novel and useful conceptualization of power it advances theory on power in research on management and organizational behavior. Fourth, due to the interdisciplinary nature of the present research, this dissertation also addresses current challenges in research on human resource management. Fifth, the experiments presented in this dissertation inform research on behavioral operations management.

CHAPTER 2

2. The Influence of Power on Time Planning[2]

2.1. Introduction

Leaders often get involved in project management to ensure appropriate time planning for project tenders to be accepted or projects to be finished in time. Thus, time planning is an important process for project managers and leaders (PMI, 2013) that has been found to be related to project success (Papke-Shields, Beise, & Quan, 2010). In projects with uncertain durations, planning too short or too long can be thought of in terms of more abstract, high-level concerns according to construal level theory (Trope & Liberman, 2010), such as losing a project tender or missing a deadline, and comparatively low-level concerns, such as increased costs or lost profits (Fiedler, 2007). Depending on the problem setting, different strategies seem appropriate to address the high-level concerns: Whereas at a project proposal stage planning short durations is attractive to win a project tender, at a project execution stage, after winning a project tender, planning long durations is attractive to be able to finish the project in time. At both project stages, cost-related concerns can be addressed best by plans that minimize the expected costs.

Thus, when planning uncertain project durations, leaders may focus on high-level concerns according to construal level theory (Trope & Liberman, 2010), such as winning the project tender and finishing in time, or they may focus on low-level cost-related (Fiedler, 2007) concerns. A detailed understanding of how leaders address these planning-related concerns is important for leader development and project success. However, research on time planning behavior that takes structural variables, such as costs, as well as individual

[2] Chapter 2 is based on a working paper by Steinberg, Schiffels, Fügener, and Peus (2019), currently under review at the Journal of Organizational Behavior.

variables, such as power, into account is scarce (Stingl & Geraldi, 2017). Leaders have power (Sturm & Antonakis, 2015), and recent experiments on power suggest that psychological power takes a prominent place among the individual factors that influence time-planning behavior for one's own tasks (Halkjelsvik & Jørgensen, 2012; Moon & Chen, 2014; Weick & Guinote, 2010). One cannot, however, conclude from the existing studies how high-power individuals act in projects that others execute and whether they orient their planning behavior towards low-level or high-level concerns in project management. Furthermore, there is currently a debate about experimental rigor and generalizability of experiments on psychological power (Sturm & Antonakis, 2015) that I address below.

The present research investigates the influence of *psychological power* (Galinsky et al., 2003; Magee & Galinsky, 2008; Sturm & Antonakis, 2015; Tost, 2015), generally defined in social psychology as "an individual's internal mental representations of their power in relation to others in their social environment" (Tost, 2015, p. 30). Psychological power corresponds to *structural power* (Sturm & Antonakis, 2015), defined in organizational science as "having the discretion and the means to asymmetrically enforce one's will over others" (Sturm & Antonakis, 2015, p. 139), but may have quite different effects (Bugental & Lewis, 1999; Fast et al., 2012; Flynn, Gruenfeld, Molm, & Polzer, 2011; Sturm & Antonakis, 2015; Tost, 2015).

In two preregistered experiments, I aim to analyze the causal influence of psychological power and costs on time planning at two different project stages with different high-level concerns, a project proposal stage and a project execution stage. At the same time, I aim to address several issues raised in the debate on experimental rigor and the interpretation of results in power research (Sturm & Antonakis, 2015; Sturm, Herz, & Antonakis, 2018), as well as in research on organizational behavior (Antonakis, 2017; Antonakis, Bendahan, Jacquart, & Lalive, 2010; Tost, 2015). Randomized experiments are

suitable to establish causality in research on leadership and power (Antonakis, 2017; Antonakis et al., 2010; Mumford, 2011; Sturm & Antonakis, 2015), but a range of further conditions need to be met to enhance confidence in and to ensure the generalizability of the results of experiments (Antonakis et al., 2010; Barr et al., 2013; Judd et al., 2012; Sturm & Antonakis, 2015). I address issues highlighted by Sturm and Antonakis (2015), such as the occurrence of experimental demand effects, difficulties in interpreting the results of studies due to confounding factors, and the ecological validity of the outcome measure. I further address issues that pertain to hypothesizing after the results are known (HARKing, Antonakis, 2017; Kerr, 1998) by means of preregistrations and clear reporting of findings. Lastly, I address issues that pertain to modeling random effects to increase the generalizability of results (Baayen et al., 2008; Barr et al., 2013; Judd et al., 2012) and contrast my analyses with analysis strategies common in previous studies on psychological power in an exploratory fashion.

The present research responds to recent calls for more rigorous research on power (Flynn et al., 2011; Sturm & Antonakis, 2015; Sturm et al., 2018) as well as an integrated management science and psychology approach to the study of time planning (Crusius et al., 2012; Fiedler, 2007; Halkjelsvik & Jørgensen, 2012; Katsikopoulos & Gigerenzer, 2013; Svejvig & Andersen, 2015). Stingl and Geraldi (2017) specifically called for research analyzing power relations in project management, of which time planning is an important process (PMI, 2013). Thus, the present research contributes to power research by using a combination of approaches from management science and psychology to investigate the influence of planning-related concerns and psychological power on time planning in projects. It further contributes to the experimental rigor debate by adding a focus on choices in data analysis and generalizability.

2.2. Theory

In recent reviews, several authors highlighted that individual variables, such as power, and structural variables, such as costs, are important to understand time planning behavior (Halkjelsvik & Jørgensen, 2012; Stingl & Geraldi, 2017). I derive predictions regarding the influence of power from psychological theory and predictions regarding the influence of costs from management science.

The behavioral, cognitive, and emotional effects of power are manifold and documented in a vast literature spanning psychology, management science, and related disciplines (recent reviews include Anderson & Brion, 2014; Magee & Smith, 2013; Schaerer, Galinsky, Lee, & Thau, 2017; Sturm & Antonakis, 2015; Tost, 2015). Empirical research on the effects of psychological power on time planning showed that high-power individuals thought they had greater control over their time (Moon & Chen, 2014) and planned shorter completion times for tasks they executed themselves (Weick & Guinote, 2010). Moon and Chen (2014) manipulated psychological power by having participants engage in a non-enacted role manipulation and found that high psychological power, compared to low psychological power, led to perceptions of greater time availability (Moon & Chen, 2014). This relationship was mediated by the high-power individuals' perception to have greater control over their time (Moon & Chen, 2014). In a series of studies where participants planned and executed everyday tasks, for example, correcting a text with the help of a word processing software or writing a term paper, Weick and Guinote (2010) showed that high-power individuals plan too optimistically for tasks they carry out themselves (Weick & Guinote, 2010). The authors identified a goal-oriented attentional focus as a mediator of this relationship (Weick & Guinote, 2010). These experiments, however, do not allow one to conclude whether high-power individuals orient their planning behavior towards

low-level or high-level concerns in project management, where planning is usually carried out for somebody else.

After a decade of productive research on psychological power (Galinsky et al., 2015) the field is currently undergoing a phase of consolidation, as evidenced by critical reviews (Flynn et al., 2011; Sturm & Antonakis, 2015), the failure to replicate the effects of an embodiment manipulation of psychological power (Ranehill et al., 2015; K. M. Smith & Apicella, 2016), and the failure to find effects predicted from theory (Heller, Borsay, & Ullrich, 2017). More rigor in experiments and field studies as well as more research that analyzes the effects of structural power, and the relations of structural power with psychological power has been called for (Flynn et al., 2011; Sturm & Antonakis, 2015; Sturm et al., 2018) and first responses to these calls are available (Bendahan, Zehnder, Pralong, & Antonakis, 2015). The experimental rigor debate in power research echoes similar debates in leadership research (Antonakis, 2017; Antonakis et al., 2010) and other fields of psychology (Open-Science-Collaboration, 2015; Simmons, Nelson, & Simonsohn, 2011). Common themes in these debates are the need for empirical research to be grounded in theory, calls for more rigor, and calls for more transparency in reporting findings, all of which I address in the present research. Project management is a promising context to address many of the concerns raised in the experimental rigor debate, and, in particular, the time planning task used here offers the possibility to define a normative optimum and to incentivize participants based on performance. A further strength of my approach is the combination of rigorous experimental methods with preregistered predictions based on psychological theory.

I derive the prediction that high-power individuals orient their planning towards high-level concerns from the social distance theory of power (Magee & Smith, 2013). The social distance theory states that high-power individuals will be more sensitive to high-level "why" concerns than low-power individuals, who will be more sensitive to low-level "how"

concerns, due to a stronger experience of psychological distance by high-power individuals (Fiedler, 2007; Fiedler, Jung, Wänke, Alexopoulos, & de Molière, 2015; Magee & Smith, 2013; Trope & Liberman, 2010). A number of psychological studies point to the fact that high-level concerns carry more weight in greater perceived distance (Fujita, Eyal, Chaiken, Trope, & Liberman, 2008; Liberman & Trope, 1998; Sagristano, Trope, & Liberman, 2002; Todorov, Goren, & Trope, 2007) and would thus be more important for high-power individuals (Magee & Smith, 2013). For instance, Liberman and Trope (1998) demonstrated that high-level desirability aspects, compared to low-level feasibility aspects, are more important for decisions in the far future, that is, when distance is high (Liberman & Trope, 1998). According to this line of reasoning, planning times accurately to keep costs in check is a low-level concern (Fiedler, 2007), whereas winning a project tender or finishing a project in time are high-level concerns. Note that depending on the project stage with the respective high-level concern, winning a project tender or finishing a project in time, high-power individuals could be motivated to plan longer or shorter durations.

The influence of different cost settings, that is, different costs associated with planning too long or too short that lead to different normatively optimal solutions in trade-off decisions, has been analyzed extensively in studies from behavioral operations management, a sub-discipline of management science. Different cost settings are commonly used in studies from behavioral operations management to isolate behavioral effects. For instance, different cost settings have been used to distinguish whether participants tend to over- or underorder inventory in newsvendor trade-off decisions (Schweitzer & Cachon, 2000). Typically, these studies find that depending on the respective experimental conditions participants benefit or suffer from certain systematic deviations from optimal plans (Benzion, Cohen, Peled, & Shavit, 2008; Bolton & Katok, 2008; Schweitzer & Cachon, 2000). In studies of inventory management, Schweitzer and Cachon (2000) established an effect that has since been

replicated in several studies (Moritz, Hill, & Donohue, 2013; Ren & Croson, 2013; Schiffels et al., 2014): Individuals tend to order too few high-margin products, that is, products with high opportunity costs in case of stock-outs, and too many low-margin products, that is, products with high costs in case of excess inventory (Schweitzer & Cachon, 2000). This observation is also referred to as the pulled-to-center or mean anchor effect (Schweitzer & Cachon, 2000). In terms of time planning, I would thus expect that individuals plan too short in cases where they should plan relatively long durations according to the cost setting (called long-plan condition, see Model and Hypothesis section) and plan too long in cases where they should plan relatively short durations according to the cost setting (called short-plan condition, see Model and Hypothesis section).

2.3. Model and Hypotheses

To address the question if psychological power and costs influence whether high-power individuals focus on high-level or low-level concerns in planning project durations, I designed two experiments. Each of these was set up to disentangle whether high-power individuals: (1) plan longer/shorter in general, (2) plan according to high-level concerns or according to the respective low-level cost-related concern (Fiedler, 2007), and (3) whether this was due to the influence of power as an individual variable or the cost settings as a structural variable. To achieve this, I varied the high-level concern across experiments and the cost setting within experiments: First, across experiments, I presented the time planning problem either at the proposal stage (Experiment 1) or at the execution stage (Experiment 2) of a project, both of which are characterized by different high-level concerns. At the proposal stage of projects, a high-level concern is to offer a duration short enough to win the project tender. During project execution, a major concern is to plan long enough to finish in time (PMI, 2013). Second, I implemented two cost settings to be able to find out whether the participants would adapt their planning behavior according to low-level cost-related

concerns. In the following paragraphs, I detail the planning problem with the different cost settings at the project proposal stage and at the project execution stage.

2.3.1. Project proposal model

In Experiment 1 that presents the planning problem at the project proposal stage, participants planned the durations p of project tenders, where winning the project tender leads to a revenue r. The tender is won, if the proposed planned duration p does not exceed the maximum accepted duration D of the customer, with the latter being unknown but following a known distribution F. Thus, the shorter the planned duration, the greater is the probability of winning the project tender. However, realizing shorter durations leads to additional costs. We define costs c to speed up the project per time unit based on the base project duration $D^{\max}$, such that the total costs for a planned project add up to $(D^{\max} - p)c$. In case a project tender is accepted ($p \leq D$), a profit of $\Pi(\mathrm{p}) = r - (D^{\max} - p)c$ is realized, while no profit is realized if the project tender is not won ($p > D$). Thus, at the proposal stage, participants face a trade-off between planning too short (leading to unnecessary costs) and planning too long (leading to a lower chance of winning the tender). The expected profit $E\big(\Pi(p)\big)$ can be computed as follows:

$$E\big(\Pi(p)\big) = (1 - F(p)\,(r - (D^{\max} - p)c)) \tag{1}$$

The profit maximizing planned duration is obtained by the first derivative, i.e.

$$\frac{\partial \mathrm{E}(\Pi(p))}{\partial p} = \big(1 - F(p)\big)c - f(p)(r - (D^{\max} - p)c) \tag{2}$$

and for a uniform distributed $D \sim U(D^{min}, D^{max})$, as employed for most related experimental studies from management science as well as our study, the profit maximizing planned duration p^* is as follows:

$$p^* = D^{max} - \frac{r}{2c} \tag{3}$$

I implemented two different cost settings with different profit maximizing planned durations: A long-plan condition with relatively high costs of speeding up the project (relatively low values of $\frac{r}{2c}$), and a short-plan condition with relatively low costs of speeding up the project (relatively high values of $\frac{r}{2c}$).

2.3.2. Project execution model

To analyze time planning at the project execution stage, I created an experimental framework based on the newsvendor trade-off problem that has often been applied to analyze planning behavior in inventory management (Schweitzer & Cachon, 2000) and can be applied to time planning scenarios as well (for example, in the context of health care planning, Fügener, Schiffels, & Kolisch, 2015; Olivares, Terwiesch, & Cassorla, 2008). In the experiment that presented the planning problem at the project execution stage, participants planned project durations under uncertainty about the realized project durations, but with a known cumulative distribution of realized durations and a known trade-off between different costs for planning too long or too short. Under these circumstances, it is straightforward to calculate the expected profit maximizing project duration. Depending on the planned duration p and the realized duration D the total profit $\Pi(p, D)$ for one project is

$$\Pi(p, D) = r - (cD + c_e \max\{p - D, 0\} + c_t \max\{D - p, 0\}) \tag{4}$$

where r denotes the revenue, c the costs for days planned, c_e the costs per day planned too long, and c_t the project tardiness costs per day planned too short. Given the cumulative distribution of the realized durations $F(D)$ the profit maximizing planned duration p^* that maximizes the expected profit $E\big(\Pi(p)\big)$ is obtained by the first derivative resulting in

$$p^* = F^{-1}\left(\frac{c_e}{c_e+c_t}\right). \tag{5}$$

where F^{-1} denotes the inverse cumulative distribution function of the realized duration D. For a uniform distributed $D \sim U(D^{min}, D^{max})$, as employed for our experimental studies, the profit maximizing planned duration is as follows:

$$p^* = D^{min} + \left(\frac{c_e}{c_e+c_t}\right) * \left(D^{max} - D^{min}\right) \tag{6}$$

In the project execution stage, I define a long-plan condition where the profit maximizing planned duration is longer than the average realized project duration, due to relatively severe consequences of planning too short ($c_t > c_e$), and a short-plan condition with relatively severe consequences of planning too long ($c_t < c_e$), where the profit maximizing planned duration is shorter than the average realized project duration.

2.4. Hypotheses

The hypotheses are grounded in the literature discussed in Section 2. In short, based on previous empirical studies and theoretical considerations I suggest that at the project proposal stage, high psychological power will lead individuals to orient their time planning behavior towards winning the project tender. This prediction is based on the social distance

theory of power (Magee & Smith, 2013) that proposes a positive relation between high power and a cognitive focus on central/abstract aspects of decisions. If this prediction holds, high-power individuals should generally plan shorter in the proposal stage experiment. Furthermore, in the proposal stage experiment, participants plan reductions in project durations and can be expected to plan longer durations if the costs for each day of reductions are high. I thus derive the following hypotheses:

Hypothesis 1. In the project proposal stage, high-power individuals plan shorter times than low power individuals, due to a stronger focus on the central/abstract goal of winning the project (by offering the shortest time to completion).

Hypothesis 2. *In the project proposal stage, longer times are planned in the high-cost setting than in the low-cost setting*.

At the project execution stage, high psychological power will lead individuals to orient their time planning behavior towards finishing the project in time. This prediction, too, is based on the social distance theory of power (Magee & Smith, 2013) that proposes a positive relation between high power and a cognitive focus on central/abstract aspects of decisions. If this prediction holds, high-power individuals should generally plan longer in the execution stage experiment. In the execution stage experiment, the participants plan project durations and can be expected to plan shorter durations if the costs for each day planned are high. I thus derive the following hypotheses:

Hypothesis 3. In the project execution stage, high-power individuals plan longer durations than low power individuals, due to a focus on the central/abstract goal of completing the project on time (by planning more time to completion).

Hypothesis 4. In the project execution stage, shorter times are planned in the high-cost setting than in the low-cost setting.

2.5. Experiments

2.5.1. Experiment 1: Project Proposal Stage

In Experiment 1, I aimed to test whether high-power individuals differed from low-power individuals in their time planning behavior in projects. The hypotheses, design, sample, procedure, and analysis plan had been preregistered prior to the start of the data collection (the preregistration can be found in Appendix A). I report deviations from our preregistration, how I determined the sample size, all data exclusions, all manipulations, and all measures in the study (cf., Simmons, Nelson, & Simonsohn, October 14, 2012).

2.5.2. Design and Sample

Each participant was randomly assigned to one of the six treatments defined as a full factorial design of cost setting and power. The study relied on a 3 power (high power vs. low power vs. control) x 2 cost setting (long-plan condition vs. short-plan condition) x 10 project (round 1 through 10) mixed-effects design with repeated-measures on the project factor.

303 students (136 female; mean age 23.6 years, standard deviation 4.3 years, six participants failed to provide age information) took part in the experiment at the computer laboratory of the faculty of economics and management of a large university. Due to technical problems the experimental data of five participants was not recorded, that is, the final sample consisted of 298 participants. I determined the sample size based on recommendations from social psychological research on power for cases when effect sizes are unknown (cf. Lammers, Stoker, Rink, & Galinsky, 2016). Experimental sessions lasted for about 30 minutes and on average the participants received 6.06 € (with a standard deviation of 1.77 €; approximately US $ 7.2 and US $ 2.1, respectively), which included a show up fee of 4 € (approximately US $ 4.6).

2.5.3. Procedure

Participants were recruited via ORSEE (Greiner, 2015) and placed in semi-private cubicles in front of computers upon arrival at the laboratory. All instructions (see Appendix B), experimental manipulations and measures were presented on the computers. In each session, all participants were prompted to start at the same time. On the first screen, they were informed that they would be rewarded with a minimum of 4 € (approximately US$ 4.6) and that depending on their performance in the experiment their reward could be considerably higher. Afterwards, they were asked to complete the power manipulation, to work on the time planning task, and to respond to some demographic questions before receiving feedback about their reward. All participants were asked to remain at their computer until the last one had finished and to leave the laboratory together.

Power Manipulation. I applied the same manipulation of psychological power as Moon and Chen (2014), which consists of visualizing oneself as the interviewer or the interviewee in a job interview (Moon and Chen (2014) published the manipulation as an online supplement to their paper). In line with recommendations from the field (Moskowitz, 2004), I added a no-intervention control group. The participants were provided with an image as an anchor for their visualization and bound to remain on the page for 60 seconds, after which they could proceed with the experiment (Moon & Chen, 2014). After the power manipulation, participants in the intervention groups completed an attention check and were asked whether they found the scenario, they were asked to visualize themselves in, difficult to imagine. All participants completed a manipulation check (see Measures section).

Time Planning Task. The second part of the experiment presented participants with the time planning task. Before engaging in the task, the participants received instructions that informed them about the scenario, the cost setting, and about how their compensation for the experiment depended on their performance in the time planning task (see instructions in

Appendix B). The instructions were always available on screen. The participants received feedback after planning the duration of each of ten projects, including their planned duration, the maximum accepted duration, and the cost incurred for speeding up the project. They were further provided with information about the project tender's success and their earnings.

The instructions informed the participants that they were to plan the durations of ten project tenders that would be bought by an unknown customer if they were shorter or equal to his maximum accepted duration. The stochastic accepted durations of the unknown customer ranged between 200 and 300 days and followed a discrete uniform distribution. I used the same accepted durations that were randomly drawn in advance for all participants and I kept the order of the accepted durations constant. The profit maximizing planned duration differed by cost setting only and was set to 275 days in the long-plan condition and to 225 days in the short-plan condition.

Manipulation check and demographic questions. After planning ten projects, a second manipulation check (see Measures section), was presented to verify whether the effect of the manipulation could still be detected. Only a few studies on psychological power used repeated-measures outcome variables and given the current debate around experimental manipulations of power (Sturm & Antonakis, 2015) I aimed to check the manipulations' effectiveness in a repeated-measures setting, that is, whether the manipulated effect could be maintained throughout the experiment. After completing the second manipulation check, the participants received feedback about the total cost incurred and the individual compensation that resulted from their performance. On the last pages of the experiment, I asked for the participants' age, gender, and mother tongue.

2.5.4. Measures

I used the same attention check as Moon and Chen (2014), that is, the participants were asked to indicate which role they had had in the power manipulation scenario. As a

manipulation check, the participants rated their psychological power on five forced choice items (e.g., “unimportant vs. important”, 7-point scale, Cronbach’s alpha = 0.86, Lammers, Dubois, Rucker, & Galinsky, 2013; P. K. Smith, Wigboldus, & Dijksterhuis, 2008) directly after the power manipulation and again after the time planning task (Cronbach’s alpha = 0.88) that served as the dependent variable. The question of whether participants found it difficult to visualize themselves in the scenario that we presented as a power manipulation was adopted from Schwarz et al. (1991), who identified ease of retrieval as a confound of feeling assertive (see also, Lammers, Dubois, Rucker, & Galinsky, 2017). I used the durations planned by each participant as the dependent variable of our study. The demographic questions included questions for the participants’ age, gender, and mother tongue.

2.5.5. Data Analysis

I checked the success of the power manipulation by comparing the mean of the five manipulation check items in an ANOVA followed up by a Tukey-HSD post hoc test. To analyze the dependent variable, I used a linear mixed effects model (Baayen et al., 2008; Barr et al., 2013; Judd et al., 2012) with random intercepts for participants, and random intercepts and random slopes for the time planning task (Barr et al., 2013). This data analysis strategy is in line with current suggestions on how to analyze similar data sets (Baayen et al., 2008; Barr et al., 2013; Judd et al., 2012). Conforming to good practice, but going beyond our preregistration, I tested different random effects structures in step-wise elimination procedures based on log-likelihood analyses (Judd et al., 2012; Kuznetsova, Brockhoff, & Christensen, 2017). The models with partial random effects structures performed no better than the models with a maximal random effects structure reported here. Degrees of freedom for the fixed effects in the model were derived using Satterthwaite approximations and tested for significance by means of t-tests (Judd et al., 2012; Kuznetsova et al., 2017). I performed

the data analyses in R (R Core Team, 2014), notably making use of the lme4 package (Bates, Mächler, Bolker, & Walker, 2015), and the lmerTest package (Kuznetsova et al., 2017). I report standard errors instead of standard deviations throughout the manuscript to facilitate the detection of significant differences.

2.5.6. Results

Manipulation Check. The participants in the high-power condition reported greater feelings of power (M = 5.86, SE = .08) than those in the control condition (M = 4.63, SE = .09) and the low-power condition (M = 3.99, SE = .11; F[2,295] = 120.7, p < .0001). A Tukey-HSD post-hoc test indicated that all groups differed significantly from each other. Thus, the power manipulation was successful.

The second manipulation check, aimed at verifying the effectiveness of the power manipulation after the time planning task, revealed a significant difference (F[2,295] = 3.46, p = .0326) between the low-power group (M = 3.79, SE = .11) and the control group (M = 4.23, SE = .12), as evidenced by a Tukey-HSD test. The high-power group (M = 3.94, SE = .11) did not differ significantly from either the control group or the low-power group.

To explore whether the power manipulation was effective throughout the experiment, that is, whether the feelings of power differed significantly between the two points of time, I conducted t-tests within each group. The exploratory analyses revealed that the feelings of power decreased significantly in the high-power group (t[1,101] = 13.16, p < .0001) and in the control group (t[1,100] = 3.65, p = .0004), but not in the low power group (t[1,94] = 1.50, p = .1363).

Hypothesis Tests. Means and standard errors across ten rounds of the time planning task are given in Table 1. Table 2 presents an overview over the results from the mixed effects model analyses. My analysis revealed a significant effect of cost setting. The participants planned longer times in the long-plan condition (M = 257.20, SE = 1.05) than in

the short-plan condition (M = 237.32, SE = .98; t[1,99.19] = 12.79, p < .0001). Thus, Hypothesis 2 could be confirmed. According to my analysis, there was no significant effect of power in the data from Experiment 1. Participants in the low-power condition (M = 248.24, SE = 1.58) planned marginally longer than participants in the control condition (M = 245.39, SE = 1.67) and the high-power condition (M = 247.81, SE = 1.52), but none of these differences reached statistical significance. Based on these results, I reject Hypothesis 1.

Table 1

Means and Standard Errors from the Time Planning Task of Experiment 1

Cost Setting	Power	*N*	*M*	*SE*
high costs (long-plan condition)	high power	47	258.69	1.73
	control	51	256.71	1.75
	low power	49	256.28	1.98
low costs (short-plan condition)	high power	55	238.51	1.55
	control	50	233.84	1.71
	low power	46	239.67	1.76

Note: Table 1 shows the means and standard errors of the time planning task across all ten projects by experimental group.

Table 2. Results of the Mixed-Model Analyses of Experiment 1

	Focal Analysis				Exploratory Analysis			
Parameter	Variance		Chi^2		Variance		Chi^2	
Random effects								
Participants								
Intercept	131.53		957.14					
Stimuli								
Intercept	0		0		0		0	
Low costs	0.84				3.91			
High costs	8.63		4.18		11.61		0.09	
Low-power	4.52				0			
Control	4.91				0			
High-power	4.10		0		0		0.01	
Residual	187.80				320.2			
	Estimate (*SE*)	*df*	*t*	*p*	Estimate (*SE*)	*df*	*t*	*p*
Fixed effects								
Intercept	237.95 (1.63)	67.16	145.95	$p < .0001$	237.95 (0.92)	21.00	259.15	$p < .0001$
Low costs x High costs	19.95 (1.56)	99.19	-12.79	$p < .0001$	19.95 (0.80)	16.80	25.03	$p < .0001$
Low power x Control	-2.64 (1.81)	153.96	-1.45	$p = .1480$	-2.64 (0.81)	2970.0 0	-3.26	$p = .0011$
Low-power x High-power	0.67 (1.77)	248.72	0.38	$p = .7060$	0.67 (0.81)	2970.0 0	0.83	$p = .4086$
High power x Control[a]	3.30 (1.72)	297.59	1.92	$p = .0560$	3.30 (0.79)	2969.9 0	4.16	$p < .0001$
-2log-likelihood				24712.4				26318.3

Note. Table 2 shows a comparison of the focal analysis and the exploratory analysis that performed significantly worse according to a log-likelihood comparison. [a]Taken from an identical model with recoded power group

Exploratory Analyses. I had preregistered three different exploratory analyses that aimed at ruling out alternative explanations for the results of the experiments and at contrasting common analysis strategies in studies of psychological power with analyses according to current best-practice recommendations. To rule out that the repetition in the planning task produced an effect in itself, I conducted analyses with the first three rounds and the first round of the planning task only. Two of the preregistered exploratory analyses, that is, analyzing the first round and the first three rounds only, yielded the same pattern of results (see Appendix C) as the model reported in Table 2. To rule out that the difficulty of imagining the scenario used as a manipulation of psychological power produced unintended effects (Lammers et al., 2017; Schwarz et al., 1991), I repeated the analysis without the participants who had responded yes to a dedicated control question. When I analyzed all ten rounds excluding participants who found it difficult to imagine the power manipulation, in line with my preregistration, I found a significant difference between the low-power condition (M = 249.21, SE = 1.61) and the control condition (M = 245.39, SE = 1.67; t[1,169.18] = -1.98, p = .0490), as well as between the control condition and the high power condition (M = 247.89, SE = 1.58; t[1,250.78] = -1.98, p = .0483). There was also a significant main effect of cost condition (t[1,84] = 12.52, p < .0001). Excluding participants based on preregistered criteria is an alternative way to account for confounding participant variation compared to our focal analysis that conceptualized participants as random effect.

To contrast these analysis strategies with typical analyses in previous studies of psychological power, I also analyzed a model leaving out the random intercept for participants (Baayen et al., 2008; Barr et al., 2013; Judd et al., 2012). This aspect of my non-preregistered, exploratory analysis mirrors analysis strategies in studies of psychological power that usually rely on between-subject designs and thus cannot include a random effect for participants. The mixed effects model analysis without the random effect for participants

is reported in Table 2. It revealed a significant difference in time planning between the control condition (M = 245.39, SE = 1.67) and the low power condition (M = 248.24, SE = 1.58; t[1,2970] = -3.26, p = .0011) as well as between the control condition and the high-power condition (M = 247.81, SE = 1.52; t[1,2969] = 4.16, p < .0001). The difference between the low-power condition and high-power condition was not significant. A log-likelihood analysis showed that the exploratory model without random effect for participants performs significantly worse than the focal model with random effect for participants (see Table 2).

2.5.7. Discussion of Experiment 1

Experiment 1 aimed to establish whether high-power individuals differ from low-power individuals in their time planning behavior and to seek first indications whether they oriented their time planning towards high-level concerns rather than low-level concerns. In the proposal stage, the high-level concern was to win the project tender and thus to plan short durations. I found different results in my focal analysis and my exploratory analyses. The results of my focal preregistered analyses show a significant effect of the different cost settings, but no effect of the manipulation of psychological power. That is, participants adjust their planning according to the low-level cost concern irrespective of the power condition. Exploratory analyses that did not model participants as a random factor revealed significant differences between the control group and both power groups, but not between the low-power group and the high-power group. Reasons for this will be discussed in the General Discussion section. Note that the significant first manipulation check indicates that the power manipulation had its intended effect and that the inconsistent pattern in the second manipulation check indicates that the effect of the manipulation could not be maintained throughout the experiment. This pattern of results also indicates that the participants responded truthfully and that no demand effect was present in the data (otherwise I would

have expected the second manipulation check to be significant and similar to the first one in terms of mean values). The significant decrease in feelings of power between the first and the second manipulation check in the control condition and high power condition can be interpreted as evidence that the experiment itself influenced the feelings of power as a structural variable (Schaerer et al., 2017; Sturm & Antonakis, 2015), which is in line with predictions by Sturm and Antonakis (2015). Further support for the notion that the experiment worked as intended and that the participants responded truthfully can be found in the fact that I found a clear effect of the different cost settings and replicated an effect similar to the mean anchor effect (Schweitzer & Cachon, 2000). That is, the average planned durations were between the profit maximizing planned duration and the mean acceptable duration in all experimental conditions.

2.6. Experiment 2: Project Execution Stage

Experiment 2 aimed to test Hypotheses 3 and 4. As such, Experiment 2 served to establish whether high-power individuals orient their time planning in projects towards high-level concerns rather than low-level concerns, which could not be decided based on Experiment 1 alone, where the high-level concern was to plan short durations in all experimental conditions. The hypotheses, design, sample, procedure, and analysis plan were preregistered prior to the start of the data collection (the preregistration can be found in Appendix D). I report deviations from the preregistration, how I determined the sample size, all data exclusions, all manipulations, and all measures in the study (cf., Simmons et al., October 14, 2012). I used the same design, procedure, manipulations, measures, and data analyses as in Experiment 1 and report deviations only.

2.6.1. Design and Sample

301 students (136 female; mean age 23.2 years, standard deviation 5.5 years; five participants failed to provide age information) took part in the experiment that was based on the same design as Experiment 1. On average the participants received 6.01 € (with a standard deviation of 1.05 €; approximately US $ 7.2 and US $ 1.3 respectively), which included a show up fee of 4 € (approximately US$ 4.6).

2.6.2. Procedure

I employed the same procedure as in Experiment 1, except for slight variations in the time planning task reported below (the instructions can be found in Appendix E).

Time Planning Task. After planning the duration of each of ten projects, participants received feedback, including their planned duration, the actual realization time, and the cost incurred. The information about the costs incurred were differentiated between the costs associated with the realization time, with the days planned too much, and with the days planned too little.

The participants planned the durations of ten projects with stochastic realization times that were said to be performed by an unknown third party. Realization times ranged between 200 and 300 days and followed a discrete uniform distribution. The realization times of Experiment 2 were equal to the randomly drawn acceptable durations in Experiment 1. The profit maximizing planned duration differed by cost setting only and was set to 275 days in the long-plan condition and to 225 days in the short-plan condition, as in Experiment 1.

2.6.3. Measures

The reliabilities for the manipulation checks in Experiment 2 were as follows: Cronbach's alpha = 0.86 for the first manipulation check and Cronbach's alpha = 0.89 for the second manipulation check.

2.6.4. Results

Manipulation Check. The participants in the high-power condition reported greater feelings of power (M = 5.89, SE = .09) than those in the control condition (M = 4.69, SE = .09) and the low-power condition (M = 4.03, SE = .11; F[2,298] = 99.67, p < .0001). A Tukey-HSD post-hoc test indicated that all groups differed significantly from each other. Thus, the power manipulation was successful.

The second manipulation check, after the time planning task, revealed no significant difference between the groups. The high-power condition (M = 4.18, SE = .14) did not differ from either the control condition (M = 4.33, SE = .11) or the low-power condition (M = 4.20, SE = .12; F[2,298] = .412, p = .6630). These results indicate that the effect of the manipulation could not be maintained throughout the experiment.

To explore whether the feelings of power differed significantly between the two points of time we conducted t-tests within each group. The exploratory analyses revealed that the feelings of power decreased significantly in the high-power group (t[1,100] = 10.47, p < .0001) and in the control group (t[1,99] = 3.80, p = .0003), but not in the low power group (t[1,99] = -1.30, p = .1982).

Hypothesis Tests. Means and standard errors across ten rounds of the time planning task are provided in Table 3. Table 4 presents an overview over the results from the mixed effects model analysis. The analysis revealed a significant effect of cost setting. The participants planned longer times in the long-plan condition (M = 258.89, SE = 1.03) than in the short-plan condition (M = 247.44, SE = 1.06; t[1,201.3] = -7.69, p < .0001). Thus, Hypothesis 4 could be confirmed. There was no significant effect of power in the data from Experiment 2. Participants in the low-power condition (M = 251.52, SE = 1.58) planned marginally shorter than participants in the control condition (M = 254.72, SE = 1.43) and

participants in the high-power condition (M = 253.07, SE = 1.16), but none of these differences reached statistical significance. Based on these results, I reject Hypothesis 3.

Table 3

Means and Standard Errors from the Time Planning Task of Experiment 2

Cost Setting	Power	*N*	*M*	*SE*
high costs (short-plan condition)	high power	51	249.10	1.52
	control	51	248.00	2.00
	low power	50	245.16	1.96
low costs (long-plan condition)	high power	50	257.13	1.57
	control	49	261.72	1.52
	low power	50	257.88	2.15

Note: Table 3 shows the means and standard errors of the time planning task across all ten projects by experimental group.

Table 4. Results of the Mixed-Model Analyses of Experiment 2

	Focal Analysis				Exploratory Analysis			
Parameter	Variance		Chi^2		Variance		Chi^2	
Random effects								
Participants								
Intercept	139.97		859.4					
Stimuli								
Intercept	0		0		0		0	
Low costs	1.47				1.03			
High costs	4.20		0.09		0.15		0.09	
Low-power	1.51				1.98			
Control	2.45				0			
High-power	6.29		0.01		0.50		0.54	
Residual	224.46				365.30			
	Estimate (*SE*)	*df*	*t*	*p*	Estimate (*SE*)	*df*	*t*	*p*
Fixed effects								
Intercept	257.26 (1.57)	135.61	163.96	$p < .0001$	257.26 (0.89)	15.80	290.02	$p < .0001$
Low costs x High costs	-11.48 (1.49)	201.29	-7.69	$p < .0001$	-11.48 (0.71)	90.60	-16.22	$p < .0001$
Low power x Control	3.32 (1.84)	164.55	1.80	$p = .0733$	3.32 (0.85)	2433.8 0	3.88	$p = .0001$
Low-power x High-power	1.61 (1.84)	173.58	0.87	$p = .3850$	1.61 (0.88)	49.90	1.82	$p = .0745$
High power x Control[a]	-1.71 (1.87)	137.65	-0.91	$p = .3627$	-1.71 (0.89)	40.20	-1.93	$p = .0602$
-2log-likelihood			25458.9				25669.6	

Note. Table 4 shows a comparison of the focal analysis and the exploratory analysis that performed significantly worse according to a log-likelihood comparison. [a]Taken from an identical model with recoded power group

Exploratory Analyses. The same preregistered exploratory analyses as in Experiment 1 (tables of the results can be found in Appendix F), that is, analyzing the first round, the first three rounds, and all ten rounds excluding participants who found it difficult to imagine the power manipulation, yielded the same pattern of results as the focal model reported in Table 4. To investigate the effects of conceptualizing participants as a random effect, I analyzed a model leaving out the random intercept for participants, as in Experiment 1. This non-preregistered, exploratory analysis showed different results. When analyzing a model without random effect for participants (as is common in studies on psychological power that rely on a between-subjects design), I found a significant difference in time planning between the low power group (M = 251.52, SE = 1.58) and the control group (M = 254.72, SE = 1.43; t[1,2434] = 3.88, p = .0001). Neither the difference between the high-power group and the control group, nor the difference between the low-power group and the high-power group were significant. The mixed effects model analysis without the random effect for participants is reported in Table 4. A log-likelihood analysis revealed that the exploratory model without random effect for participants performed significantly worse than the focal model with random effect for participants (see Table 4).

2.6.5. Discussion of Experiment 2

Experiment 2 aimed to establish whether high-power individuals oriented their time planning behavior in projects towards high-level concerns rather than low-level concerns by designing an experimental framework with a different high-level concern than in Experiment 1. In the execution stage, the high-level concern was to finish the project in time and thus to plan long durations. Across experiments, an indication for an effect of the high-level concerns emerged, in that participants planned longer durations on average in Experiment 2. While the results show a clear effect of the different cost settings, no significant effect of the

manipulation of psychological power could be detected. Exploratory analyses that did not model participants as a random factor (a common analysis strategy in previous studies of psychological power) revealed a significant difference between the control group and the low-power group, while there was no significant difference between the high-power group and the control group, nor between the low-power group and the high-power group. Reasons for this will be discussed in the General Discussion section. Note that the significant first manipulation check indicates that the power manipulation had its intended effect and that the inconsistent pattern in the second manipulation check indicates that the effect of the manipulation could not be maintained throughout the experiment. This pattern of results also indicates that the participants responded truthfully and that no demand effect was present in the data (otherwise I would have expected the second manipulation check to be significant and similar to the first one in terms of mean values). As in Experiment 1, the significant decrease in feelings of power between the first and the second manipulation check in the control condition and high power condition can be interpreted as evidence that the experiment itself influenced the feelings of power as a structural variable (Schaerer et al., 2017; Sturm & Antonakis, 2015). This is in line with predictions by Sturm and Antonakis (2015). Further support for the notion that the experiment worked as intended and that the participants responded truthfully can be found in the fact that I found a clear effect of the cost setting and replicated the mean anchor effect (Schweitzer & Cachon, 2000). As in Experiment 2, planned durations were between the profit maximizing durations and the mean realized duration in all experimental conditions.

2.7. General Discussion

Previous research on psychological power recently found itself in the midst of a debate on experimental rigor that concerned certain ways to manipulate power, the potential

occurrence of experimental demand effects, difficulties in interpreting power effects due to confounding factors, and the ecological validity of the outcome measures (Sturm & Antonakis, 2015). More general considerations were highlighted, too, such as hypothesizing after the results are known (HARKing, Antonakis, 2017; Kerr, 1998). I address many of these issues by means of preregistrations, the use of a repeated-measures design, and an outcome variable with a clearly defined normatively optimal solution. I further identify the modeling of random effects to increase the generalizability of results (Baayen et al., 2008; Barr et al., 2013; Judd et al., 2012) as an issue to be added to the experimental rigor debate.

The present research combines approaches from management science and psychology to investigate the previously unexplored influence of psychological power and different planning-related concerns on time planning in projects. Specifically, I compared the effects of experimentally induced psychological power on time-planning in two experiments with different low-level cost-related concerns (Fiedler, 2007) and different high-level concerns, such as winning a project tender or finishing a project in time. Please note, while in both experiments focusing on low-level concerns should lead to profit maximizing planned durations, the focus on high-level concerns should lead to shorter planned durations in the project proposal stage and to longer planned durations in the project execution stage.

The results demonstrate that low-level concerns, that is, the different cost settings, influence time-planning behavior, and that a mean anchor effect could be replicated in both experiments, both of which is in line with related literature from management science. Contrary to our preregistered predictions, no significant effect of psychological power on the time-planning behavior of high-power individuals occurred in the experiments reported here. I conducted exploratory analyses to investigate the influence of the decision to model participants as a random factor, which is in line with best-practice recommendations for data analysis (Barr et al., 2013; Judd et al., 2012), but not with current practice in studies on

psychological power that tend to employ between-subjects designs. The exploratory analyses without random effect for participants revealed a significant difference in time planning between the control group and the low-power group in both studies. The difference between the control group and the high-power group was only significant in exploratory analyses in the proposal stage (Experiment 1) and the difference between the low-power group and the high-power group failed to reach significance in both studies.

2.7.1. Discussion of the Null-Result

Several aspects of the present research merit a detailed discussion: First, the statistical difference between my focal and exploratory analyses, second, the fact that even the exploratory analyses revealed significant differences for the high-power group only in the proposal stage, and, third, the null-effect of psychological power that I found in two preregistered experiments with relatively large sample sizes and a rigorous experimental approach from management science.

First, in line with best-practice recommendations from psychology (e.g. Barr et al., 2013), I included random effects in my model not only for the stimuli that were randomly sampled from a larger population of potential stimuli, but also for the participants who were randomly sampled as well (Baayen et al., 2008; Barr et al., 2013; Judd et al., 2012). To maximize the generalizability of results, Barr et al. (2013) suggest researchers specify maximal random effects structures based on the research design when using linear mixed effects models to analyze the data. This approach allows for taking per-participant and per-item random variation into account simultaneously, which is an advantage over analysis of variance (ANOVA) based approaches (Barr et al., 2013). ANOVAs cannot take both sources of random variation into account at the same time and therefore perform worse than linear mixed effects models with maximal random effects structures when the variance confounded with the effect of interest is relatively small (Barr et al., 2013). The argument ties in well with

other contributions that highlight the importance of controlling both, random effects of participants and random effects of stimuli, to be able to generalize study results to the general populations of participants and stimuli, despite having sampled a random number of participants out of the general population and a number of stimuli out of the population of stimuli (Baayen et al., 2008; Judd et al., 2012).

I argue that this approach is best suited to account for per-participant random effects likely to occur in studies of psychological power due to the sampling of participants. Psychological power has been found to be contingent on previous experience and cultural background (Gruenfeld, Inesi, Magee, & Galinsky, 2008; Schaerer et al., 2017; P. K. Smith & Trope, 2006), everyday experiences (P. K. Smith & Hofmann, 2016), and task demands (Aime, Humphrey, Derue, & Paul, 2014), all of which makes it likely that confounds exist that cannot be controlled for by randomization of participants and experimental conditions. For instance, if the sample consists of students, they share certain characteristics related to previous experience with psychological power (Gruenfeld et al., 2008; P. K. Smith & Trope, 2006) that limit generalizability, and that still exist after the randomized assignment to different experimental conditions. The difference between the results of my focal analysis, where I modeled participants as a random effect, and the exploratory analyses can count as an example for this line of argumentation. My exploratory analyses revealed sample-specific significant differences due to my power manipulation that cannot, however, be generalized to the population level. My focal analyses, where participants are modeled as a random factor, control for the variance in participant reactions to the power manipulation and demonstrate that there is no effect beyond the effects of confounding variables.

Further note, that a data driven approach, too, supports the inclusion of the random effect for participants (Barr et al., 2013; Judd et al., 2012): the models with a random effect for participants perform better than the models without random effect for participants both, in

log-likelihood analyses and in terms of the Akaike Information Criterion (AIC), which indicates that models with a random effect for participants represent the data better and should be preferred over models without random effects for participants (Judd et al., 2012). In summary, this difference alone can account for the null-effect of power that is present if participants are (correctly) identified as a random effect in the model.

Second, the argument for modeling participants as a random effect may also offer an explanation for the inconsistent results regarding the main effects of power across both experiments. If per-participant random-variation is not controlled for by a random effect in the model, the difference in time planning between the control group and the high-power group in the proposal stage is significant (and in line with my hypothesis), but the same difference fails to reach significance in the execution stage. Since all main effects of power failed to reach significance when participants were modeled as a random effect, it is likely that the inconsistent results were due to confounding variables. Similar patterns of results, where the results in the control group were not significantly different from either high-power group or low-power group, have been found in previous studies of psychological power as well (Moon & Chen, 2014; P. K. Smith & Trope, 2006). Whether confounding variables are causing these effects, or whether they are due to the fact that the relational nature of power (P. K. Smith & Magee, 2015) makes it difficult to devise meaningful control groups, remains to be answered by future research. To explore these effects, experiments on power should make sure to include a control group regardless of conceptual difficulties (Moskowitz, 2004) and demonstrate that between-group differences exist even when participants are modeled as a random effect.

Third, besides the differences in analyses discussed above, aspects of the time planning task, such as the fact that I employed a repeated-measures design or the fact that it involved mathematical calculations and feedback, may have contributed to producing the

null-result. Current studies demonstrate that psychological power varies according to task demands and everyday experiences (Aime et al., 2014; P. K. Smith & Hofmann, 2016). After each round of the time planning problem the participants received feedback about the outcome of their planning, which may have been a low power experience (Keltner et al., 2003). In this context, the side finding that the participants in the control condition and the high-power condition reported significantly lower feelings of power in the second manipulation check deserves attention. This result may be seen as a further indication that power varies depending on task demands (Aime et al., 2014), every day experiences (P. K. Smith & Hofmann, 2016), and is influenced by structural variables (Sturm & Antonakis, 2015). Albeit only a side-finding, this is informative for the psychological literature on power, where power has rarely been examined as a dependent variable (Galinsky et al., 2015) and where there are not many examples of studies that employ repeated-measures designs. An explanation attributing the null result to the repeated-measures design and structural variables, however, would have to take the fact into account that in the first round alone there was no effect of power to be found. Nevertheless, this also leads to the question how long effects of power manipulations last may be raised and should be considered in attempts to explain null-results in studies of psychological power. Furthermore, I encourage future research on the effects of psychological power to make use of a larger variety of research designs than just the predominant between-subjects design.

2.7.2. Limitations

Interpretations of the results reported here need to consider the limitations of the present research. Demonstrations of (null-)effects from single research projects should be interpreted with care, even if conceptual replications within the paper show the same null-result and careful control has been exercised over experimental conditions and statistical analyses to ensure generalizability. The power manipulation I employed has been used in

studies of psychological power before and is in line with current suggestions for the research on psychological power (Moon & Chen, 2014; Tost, 2015). Psychological power has been shown to produce distinct (Bugental & Lewis, 1999), but also parallel effects (Inesi, 2010; Pitesa & Thau, 2013; Rus, van Knippenberg, & Wisse, 2012) to structural power (Tost, 2015). Given rather large situational variance in states of having power (Aime et al., 2014; P. K. Smith & Hofmann, 2016), studying psychological power is very important to understanding human behavior (Fast et al., 2012), but cannot be thought without also studying the effects of structural power (Sturm & Antonakis, 2015; Tost, 2015). I chose to focus on the effects of psychological power, which can be seen as a limitation of the present research, and I encourage future research to analyze the effects of structural power on the same ecologically valid outcome measure. A further limitation naturally lies in the choices I made in designing the experimental framework. Finally, while the ecological validity of my experiments is high compared to other experiments, I encourage future research to analyze the hypothesized relationships in field settings, where effects may be present that I have been unable to detect in the laboratory.

2.7.3. Situating the Present Research in the Experimental Rigor Debate

By combining approaches from management science and psychology, I address several concerns highlighted by Sturm and Antonakis (2015) with regard to experimental rigor in power research. First, the significant differences in the results of the repeated manipulation checks show that there were no experimenter demand effects present (Zizzo, 2010) and provide initial evidence for effects predicted by Sturm and Antonakis (2015). If the participants had guessed my hypotheses and responded accordingly, I would have expected to see the same pattern of results in both manipulation checks. Instead the significant decrease within the control and high-power condition indicates that the experiment itself was a low power experience. Moreover, I assume potential experimenter demand effects to be

uncorrelated with the objectives of our experiments (Zizzo, 2010), as it is unlikely that participants guessed the exact hypothesized relationships between psychological power, planning-related concerns, and time planning. Second, modelling a maximal random effects structure is a way to enhance the confidence one can have in interpretations of the effects, because it better accounts for the influence of confounding variation than traditional ways of data analyses (e.g., ANOVA). Third, the ecological validity and practical relevance of my experiments is relatively high for experimental studies, because I applied performance-based incentives, that is, real-world consequences for participants, and the experimental framework allows benchmarking behavioral outcomes against a normative optimum. This contrasts with psychological studies of time planning as cited above, where the participants' plans are usually compared to actual outcomes or other participants' estimates, but rarely to a normatively optimal solution that generalizes to many different situations. Time planning in projects, thus, is a context particularly well-suited to the study of the effects of power, because many concerns raised in the experimental rigor debate (Sturm & Antonakis, 2015) can be addressed.

2.7.4. Conclusion

The present research makes several important contributions to different streams of literature. First, my approach to testing the influence of psychological power in the context of project management brings together psychological theory with the methodological rigor of management science – an integration often called for (Crusius et al., 2012; Fiedler, 2007; Halkjelsvik & Jørgensen, 2012; Katsikopoulos & Gigerenzer, 2013; Stingl & Geraldi, 2017; Sturm & Antonakis, 2015; Svejvig & Andersen, 2015). Second, the present research is one of the first to provide empirical support for theoretical issues raised by Sturm and Antonakis (2015) about the (in-)effectiveness and potential lack of replicability of manipulations of psychological power in connection with ecologically valid decision outcomes. Third, by

demonstrating how modeling maximum random effects structures affects interpretation of results, my experiments support points raised by Barr et al. (2013) and Judd et al. (2012) about the design and analysis of confirmatory research and offer directions for the design and analysis of future experiments on social power. Finally, the present research contributes to an emerging stream of research that empirically identifies boundary conditions of the effects of psychological power (Heller et al., 2017; Ranehill et al., 2015; K. M. Smith & Apicella, 2016)

CHAPTER 3

3. The Influence of Power on Information Evaluation[3]

3.1. Introduction

The internal investigations into the illegal use of software to manipulate emissions of diesel engines at the car manufacturer VW revealed two likely scenarios of how the decision to use the software came about (Lossie & Hadem, 2016): The leadership may have ordered the use of the software or employees may have acted on their own initiative. In both scenarios, the decision was probably made to be able to complete the development of the diesel engine within time, budget, and regulatory constraints. The two scenarios put forward by Lossie and Hardem (2016) consider leaders and employees as two separate groups involved in the scandal, implying that leaders may have other criteria for making decisions than employees (see also Berson et al., 2015; Magee & Smith, 2013): Developing the new engine was probably seen as highly *desirable* while violating regulations was for sure *undesirable* for the leadership. At the same time, reducing carbon oxide emissions and nitrogen oxide emissions while facing budget constraints was likely seen as *unfeasible*; using a software that manipulates emissions during tests must have appeared as the only *feasible* means to the employees in order to comply with EU and US regulations eventually. This example shows that seemingly minor differences in the evaluation of arguments may have profound ethical and economic consequences, consider the US$ 10,033,000,000 that VW agreed to pay in an initial court ruling ("Executive Summary of Final Class Settlement Program (2.0 Liter Engine Vehicles)," 2016).

[3] Chapter 3 is based on a working paper by Steinberg, Knipfer, and Peus (2019).

In this research, I posit and test the assumption that desirability and feasibility arguments are evaluated differently by leaders and non-leaders in making decisions based on complex and ambivalent information as in the VW example (Berson et al., 2015; Magee & Smith, 2013). Desirability refers to outcome-related aspects of positive or negative valence (i.e., do I want to do it?; Liberman & Trope, 1998; Magee & Smith, 2013), while feasibility refers to means-related aspects of positive or negative valence (i.e., do I have the capacity to do it?; Liberman & Trope, 1998; Magee & Smith, 2013). When considering actions, decision-makers focus either on desirability or on feasibility (Lu, Xie, & Xu, 2013). I argue that power is key in understanding which types of arguments from these domains will be preferred by leaders versus non-leaders (see also Berson et al., 2015; Magee & Smith, 2013).

The power perspective in organizational decision-making is needed as power influences many different aspects of decision-making (e.g., Fast et al., 2012; Fischer, Fischer, Englich, Aydin, & Frey, 2011; Inesi, 2010; See, Morrison, Rothman, & Soll, 2011). A surge of studies have revealed a broad range of effects of power, from overconfidence (Fast et al., 2012), less advice-taking (Tost, Gino, & Larrick, 2012a), and loss-aversion (Inesi, 2010) to confirmation bias in information evaluation (Fischer et al., 2011). Partial investigations of the question at hand show that power leads to decreased attention as well as sensitivity to negative desirability information (Inesi, 2010) and negative feasibility information (Whitson et al., 2013). Yet, no study employed a fully crossed design of valence and argument domain to analyze the influence of power on the evaluation of (un-)desirability and (un-)feasibility arguments (Magee & Smith, 2013). Thus, there is no conclusive evidence regarding the differences of leaders' and non-leaders' argument evaluations, even though this may concern the decision-making in countless companies and projects.

In the present research, I posit and find that high-power individuals, such as leaders, prefer different types of arguments than low-power individuals, such as employees. In

developing my arguments, I refer to two major theories that make predictions about how power influences the evaluation of desirability and feasibility arguments, namely the approach/inhibition theory of power (AIT, Keltner et al., 2003) and the social distance theory of power (SDT, Magee & Smith, 2013). Specifically, I rely on the predictions derived from approach/inhibition theory (Keltner et al., 2003) and social distance theory (Magee & Smith, 2013) to investigate the influence of power on the evaluation of (un-)desirability and (un-)feasibility arguments (Magee & Smith, 2013). To test my hypotheses, I conducted a series of four quantitative studies and a qualitative study with different samples (students as well as employees with and without leadership responsibility), both in the laboratory and online.

The present research takes a power perspective to expand our understanding of leaders' and non-leaders' evaluation of desirability and feasibility arguments within the context of complex decisions. It makes important contributions to the literature on organizational decision-making by showing that power systematically influences the evaluation of different types of arguments, i.e. leaders and non-leaders process information differently. My studies also inform power research by identifying boundary conditions for the different theories of power with regard to the evaluation of desirability and feasibility arguments.

Finally, knowledge about differences in leaders' and non-leaders' evaluation of desirability and feasibility arguments has important practical implications. It allows to establish procedures aimed at raising awareness of the influence of power on decision-making and its consequences in terms of ethical decision-making and economic outcomes.

3.2. Power and the Evaluation of Desirability and Feasibility Arguments

Power is a defining feature of organizations and has been found to influence information processing and decision-making in many different ways (Tost, 2015). However,

to date there are only partial investigations of the influence of power on the evaluation of desirability and feasibility arguments (Magee & Smith, 2013).

Power is defined as asymmetric control over resources (Keltner et al., 2003; Tost, 2015) and viewed as a "psychological property of the individual" (Galinsky et al., 2003, p. 454). High-power individuals have been found to take less advice (See et al., 2011; Tost et al., 2012a), to discount expert advice in competitive settings (Tost et al., 2012a), and to make less accurate judgements (See et al., 2011). A high sense of power (Anderson, John, & Keltner, 2012) leads to overconfident decisions (Fast et al., 2012), even more so for narcissistic individuals (Macenczak, Campbell, Henley, & Campbell, 2016). Power has also been found to increase confirmatory information processing (Fischer et al., 2011). However, it has to be noted that recent research failed to replicate effects of certain power manipulations (Ranehill et al., 2015; K. M. Smith & Apicella, 2016), which is why these results have to be interpreted with caution. In short, there is a substantial number of studies on the effects of power on different aspects of information processing (e.g. Galinsky et al., 2015), but there is no conclusive evidence regarding the impact of power on the evaluation of arguments (Magee & Smith, 2013). Specifically, it remains an open question how power influences the evaluation of (un-)desirability and (un-)feasibility arguments (Magee & Smith, 2013), even though this issue is central to understanding differences in leader and non-leader decision-making.

I derive predictions about how power impacts the evaluation of desirability and feasibility arguments (for a summary, see Figure 1) from two prominent theories of power and the argumentation outlined in Magee and Smith (2013). The approach/inhibition theory (Keltner et al., 2003) proposes that high power goes along with an activation of the behavioral activation system (BAS, Carver & White, 1994) as well as a promotion focus (Higgins, 1997). Low power, in contrast, is said to entail an activation of the behavioral

inhibition system (BIS, Carver & White, 1994) and a prevention focus (Higgins, 1997). With regard to information processing the approach/inhibition theory makes the following prediction: "Increased BAS relative to BIS activation would make high-power individuals more selectively attentive to positive than negative information in both the desirability and the feasibility domains" (Magee & Smith, 2013, p. 171). This implies that high-power individuals are likely to prefer desirability and feasibility arguments to undesirability and unfeasibility arguments. On the other hand, low power leads to increased sensitivity to punishments and threats (Keltner et al., 2003). Hence, I posit that low-power individuals will prefer undesirability and unfeasibility arguments to positive desirability and feasibility arguments. To sum it up, the approach/inhibition theory would predict that the valence (positivity vs. negativity) of an argument makes a difference, but not the particular domain (desirability vs. feasibility, Magee & Smith, 2013).

In contrast, different predictions can be derived from social distance theory (Magee & Smith, 2013). Specifically, the social distance theory (Magee & Smith, 2013) proposes that high-power individuals experience greater social distance compared to low-power individuals and thus show more abstract ways of thinking (Magee & Smith, 2013). This assumption builds on construal level theory (Trope & Liberman, 2010), which posits that all forms of psychological distance are tied to the level of abstraction of the construals people form. In terms of construal level theory (Trope & Liberman, 2010), desirability is a higher-level concern than feasibility, because the outcome itself is more central in the mental representation of an action than the means to achieve it (Liberman & Trope, 1998; Magee & Smith, 2013; Trope & Liberman, 2010). Consistent with this line of reasoning "[…] the social distance theory predicts that desirability, relative to feasibility, will have a greater influence on high- than low-power decision makers" (Magee & Smith, 2013, p. 171). From their statement it follows that high-power individuals will prefer desirability arguments to

feasibility arguments, irrespective of the valence of the argument (Magee & Smith, 2013). Based on the relationship between power and construal level it also follows that low-power individuals will prefer feasibility arguments to desirability arguments, for feasibility is a lower level concern than desirability (Magee & Smith, 2013; Trope & Liberman, 2010).

Empirically it is not yet clear which of the two theories better predicts the evaluation of desirability and feasibility arguments by high vs low power individuals (Magee & Smith, 2013). Specifically, there are only two empirical studies that examined the effect of power on the processing of desirability or feasibility information (Magee & Smith, 2013). Inesi (2010) used a power manipulation as well as desirable and undesirable options in a choice scenario to investigate the influence of power on loss aversion. Whitson et al. (2013) investigated the influence of power on the processing of goal constraining versus goal facilitating information, which could be interpreted as feasibility versus unfeasibility information (Magee & Smith, 2013). In both studies, high power was associated with decreased attention and sensitivity to *negative* information but not with increased attention and sensitivity to *positive* information (Magee & Smith, 2013). According to Magee and Smith (2013) neither the social distance theory nor the approach/inhibition theory can fully account for these results. The social distance theory does not make valence based predictions (Magee & Smith, 2013), and the approach/inhibition theory predicts that high power is likely to lead to a focus on positive information (Magee & Smith, 2013), rather than producing its effects via a decrease of attention to negative information. Magee and Smith (2013) further note that these two studies were limited to either desirability (Inesi, 2010) or feasibility (Whitson et al., 2013) and call for a direct investigation of the influence of power on (un-)desirability and (un-)feasibility information.

3.2.1. The Present Research

To address this issue, I conducted four quantitative studies and a qualitative study, both in the laboratory and online, with student and professional samples, different manipulations of power, and different measures of an argument's relevance. I chose this mixed-methods approach involving different samples and conceptual replications of the effect in question in seeking to produce results with high levels of reliability and external validity.

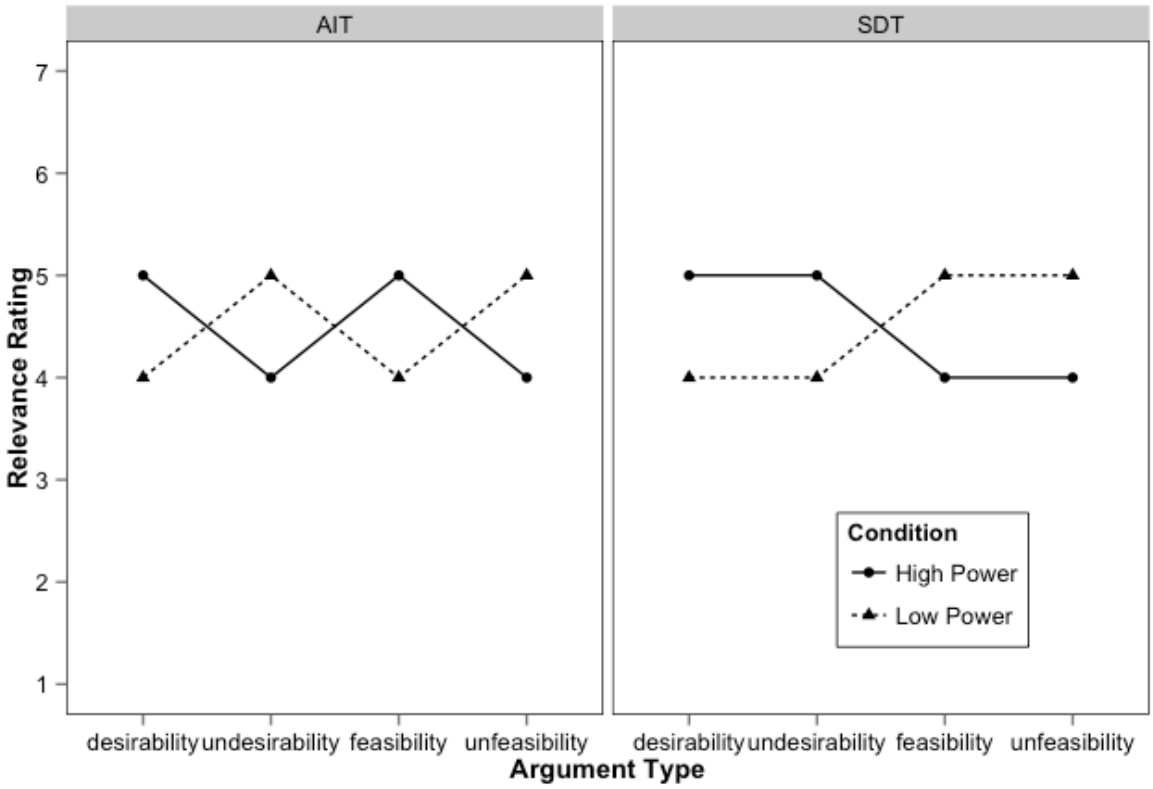

Figure 1. Idealized patterns of results. Displays idealized patterns of results that would be expected based on the AIT and SDT. All values have been chosen to match the predictions derived for this study. AIT stands for Approach/Inhibition Theory by D. Keltner, D. H. Gruenfeld, and C. Anderson (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265-284., SDT stands for Social Distance Theory by J. C. Magee and P. K. Smith (2013). The Social Distance Theory of Power. *Personality and Social Psychology Review, 17*(2), 158-186.

3.2.2. Hypotheses

Based on the argumentation above, I expect power to influence the evaluation of desirability, undesirability, feasibility, and unfeasibility arguments. Specifically, I posit the following:

H1: The evaluations of relevance of four argument types differ between the low-power and the high-power group.

Referring to the AIT, I assume that power holders favor positive information over negative information of both domains, whereas low-power individuals favor negative information over positive information of both domains. That is, I test the following two hypotheses:

H2: High-power individuals will evaluate the relevance of positive desirability and feasibility arguments higher than that of undesirability and unfeasibility arguments.

H3: Low-power individuals will evaluate the relevance of undesirability and unfeasibility arguments higher than that of positive desirability and feasibility arguments.

Based on the SDT, I suggest that power leads to preferences in information evaluation in that power holders focus more on desirability information than feasibility information, whereas low-power individuals focus on feasibility information rather than desirability information. I posit and test the following two hypotheses:

H4: High-power individuals will rate the relevance of (un-)desirability arguments higher than that of (un-)feasibility arguments.

H5: Low-power individuals will rate the relevance of (un-)feasibility arguments higher than that of (un-)desirability arguments.

3.2.3. Overview over the studies

Studies 1 to 4 are quantitative investigations of the influence of power on evaluations of (un-)desirability and (un-)feasibility arguments with student and professional samples, involving laboratory and online studies. I conducted linear mixed models to test whether the patterns of ratings differed between high vs. low power individuals in the four quantitative studies (Hypothesis 1). I analyzed the differences in the patterns of ratings, that is, I conducted the tests of Hypotheses 2 to 5, by means of planned contrasts within the

framework of a single-paper meta-analysis (McShane & Böckenholt, 2017). This method allows me to analyze the overall effect across studies and to avoid alpha error inflation at the level of individual studies. I first report the results of the linear mixed models for each study, and then the meta-analytic summary of the four studies including the tests of the specific Hypotheses 2 to 5 in the section "Meta-Analytical Results and Discussion".

While in Studies 1 to 4 participants were asked to rate the relevance of a given set of arguments, in Study 5 they were asked to generate arguments that were then rated with respect to desirability and feasibility. Study 5 thus is a qualitative investigation of the influence of power on the generation of desirability and feasibility arguments that increases the generalizability of our results by including a different operationalization of relevance.

In this article, I report how I determined the sample size, all data exclusions, all manipulations, and all measures (cf., Simmons et al., October 14, 2012). I did not derive predictions for the control group and thus focused my interpretations on the low power group and the high-power group.

3.3. Study 1

Study 1 tested how power influences the perceived relevance of desirability and feasibility arguments in a laboratory experiment with a student sample.

3.3.1. Methods

Research design. The experiment was based on a 3 (high power vs. low power vs. control) x 4 (desirable vs. undesirable vs. feasible vs. unfeasible) design with the first factor manipulated between and the second factor within participants.

Participants. I calculated a required sample size of 152 participants based on the effect sizes (two-sided $\alpha = 0.05$; power = 0.9) in studies relevant to our question (Inesi, 2010; Whitson et al., 2013). Participants were recruited on the campus of a large German university and received 6 € (approximately 7 US$) for participation. 170 undergraduates (63 female,

mean age was 20.7 years, *SD* = 2.3) participated in the study. Six participants denied the item "In your opinion, should your data be included in the analysis?" and were excluded from data analysis. The final sample included 164 participants (61 female, mean age was = 20.6, *SD* = 2.2 years).

Material and procedure. Upon arrival at the computer laboratory participants were welcomed and seated in semi-private cubicles. All further instructions, manipulations, and measures were presented on personal computers. The experiment consisted of three parts, namely the power manipulation, the information-processing task, and the manipulation check. Finally, participants were asked to provide demographic information.

The demographic information I asked for included: age, gender, education, current occupation, leadership experience, and highest number of employees ever led. On the last page of the survey the participants were asked to recall the instruction of the information-processing task and could make comments in a field for free text entry.

Power manipulation. Participants were assigned randomly to one of the power conditions. I used the same procedure as Inesi (2010, Experiment 2) who found an adaptation of the Galinsky et al. (2003) recall task to be effective for a computerized task: Participants were instructed to recall a situation in which they have had power over someone else (high power), where somebody else had had power over them (low power), or had taken a walk in the park (control condition). Participants were then asked to note some details about this situation including thoughts and feelings.

Manipulation Check. After completing the information-processing task, participants rated their sense of power in the situation recalled in the first part of the study on a 7-point scale. For this purpose, I adapted four items of the sense of power scale (Anderson & Galinsky, 2006; Anderson et al., 2012; See et al., 2011). I also included three semantic differentials based on the adjectives Lammers et al. (2013) used as a manipulation check

(e.g., “powerless” vs. “powerful”; 7-point scale; $\alpha = 0.84$). Given that the overall reliability of the seven items ($\alpha = 0.88$) was higher than the reliabilities of the separate scales, I used a composite score of all seven items as a measure for the sense of power in our manipulation check.

Information-processing task. After the power manipulation participants were instructed to rate 20 out of 40 pre-tested[4] statements with regard to their relevance for making a decision. As in Fischer et al. (2011), participants were told that the statements were summaries of longer reports and presented an argument related to a (fictitious) decision that was not specified further. Participants were asked to read each statement carefully and to rate the extent to which this statement gave them the impression that the corresponding report was relevant for the decision by answering the following question: “How relevant would a report summarized by this statement be to make a decision?” (7-point scale: 1 = not at all, 7 = very relevant). The 20 statements were presented separately on 20 consecutive screens and in random order to rule out ordering effects.

Data analysis. I used an analysis of variance to perform the manipulation check and used mixed-effects models to account for the random variance that occurred due to the item-sampling strategy (Baayen et al., 2008; Judd et al., 2012). My predictions imply different *patterns* of relevance ratings across the four argument domains (desirability / feasibility / undesirability / unfeasibility), that is I expected nonlinear relationships rather than linear relationships (see Figure 1). Mixed-effects models are necessary and adequate to investigate these nonlinear relationships because they capture the changes in direction of the curves by a

[4] In total, I created four statements for each of ten arbitrarily chosen topics. The 40 statements were pre-tested with regard to their relevance for making a decision and the strength of attitudes towards the given topics among 202 students who did not take part in any of the other studies. The results showed that no differences in attitudes towards the topics were present, and that feasibility and unfeasibility information was rated more relevant than desirability and undesirability information ($M_F = 5.05$, $SD = 1.03$ and $M_{UF} = 5.22$, $SD = 1.15$ as compared to $M_D = 4.31$, $SD = 1.09$ and $M_{UD} = 3.98$, $SD = 1.06$).

cubic term whereas linear terms only capture differences in the composite score of relevance ratings of the four domains.

To test whether the ratings of relevance of the four types of arguments (desirability, undesirability, feasibility, and unfeasibility) differed between the three experimental conditions (low power vs. high power vs. control), I estimated mixed-effects models in R (R Core Team, 2014) with the nlme package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2015). I centered the relevance ratings of the four argument types at their grand-mean, because the absolute differences in the ratings of the different items were not of interest for testing the hypotheses (Enders & Tofighi, 2007). The models contained the main effects and the interactions of power (Model 1: low vs. control, low vs. high, and Model 2: high vs. control, low vs. high, dummy-coded) with the first- to third-order polynomials of argument type. Random intercepts, and random slopes for argument type nested in participants were estimated (Barr et al., 2013) but no within-group correlations due to convergence problems (cf., Banchefsky, Westfall, Park, & Judd, 2016).

First, my hypotheses suggest different patterns of relevance ratings across the four argument domains (desirability / feasibility / undesirability / unfeasibility) in the high and low power group (H1); that is, I first looked at whether the interaction between power (low vs. high, dummy-coded) and the cubic term of argument type would be significant. Second, I expected to find a significant main effect for the cubic term of argument type, which would indicate that the relevance of the four types of arguments has been rated differently. Third, I expected to find a significant main effect for argument type, which would indicate that at least one of the four types of arguments has been rated more or less relevant than the others. As I theorized and tested *within* effects rather than *between* effects, I did not expect to find a significant main effect for power, which would indicate that power influences the mean relevance ratings *across* all argument types.

3.3.2. Results and Discussion

Manipulation check. A one-way ANOVA showed that the power manipulation successfully influenced the sense of power across conditions, $F(2, 161) = 107.23$, $p < .001$, $\eta_p^2 = 0.57$. As expected, participants in the high-power group expressed the most power ($M = 5.64$, $SD = 0.81$), and participants in the low power group expressed the least power ($M = 2.98$, $SD = 1.11$). Participants in the control group ($M = 4.61$, $SD = 0.94$) fell in between, as expected. A subsequent Tukey-HSD test revealed that groups differed significantly from each other.

Power influences the patterns of relevance ratings. Results from the mixed-effects model 1 are displayed in Table 5 (I report only significant results from Model 2 that differ from the results of Model 1) and the patterns of relevance ratings are displayed in Figure 2.

I hypothesized that the patterns of ratings of relevance would differ between the high-power group and the low power group. The significant interaction between power and the cubic term of argument type, Wald $\chi^2(1) = 5.52, p < .001$ (the Wald test is an alternative to likelihood ratio tests for small samples, Frees, 2004), indicated that the patterns of ratings of relevance differed significantly between the low power and the high power group. Figure 2 shows that the four argument types have been rated differently within each power group and that there are differences in the ratings of relevance between the groups, which is reflected in the significant interaction term. The patterns of ratings of relevance between the low power group and the control group differed significantly as well, Wald $\chi^2(1) = 5.51, p < .001$. The significant main effect in the cubic term of argument type, Wald $\chi^2(1) = 5.22, p < .001$, indicates that the curve changes its direction three times across the four conditions of argument type (see Figure 2). In other words, the relevance of the four argument types was rated differently. The mixed-effects model also revealed a significant main effect of power, which indicated higher overall ratings of relevance in the control group compared to the low

power group, Wald $\chi^2(1) = 2.03, p = .043$. All other effects did not reach statistical significance (all *p*s > .05).

Table 5

Fixed-Effects Results from the Mixed-Effects Model in Study 1

Effect	Estimate	*SE*	Wald χ^2
Intercept	-0.136	0.081	-1.689
$Power_{LCT}$	0.230	0.113	2.031*
$Power_{LH}$	0.174	0.113	1.543
Argument Type	1.379	4.557	0.303
Argument Type^2	-2.122	2.303	-0.922
Argument Type^3	-12.017	2.303	-5.219***
$Power_{LCT}$ x Argument Type	-1.126	6.386	-.176
$Power_{LH}$ x Argument Type	-2.936	6.358	-.461
$Power_{LCT}$ x Argument Type^2	3.644	3.227	1.129
$Power_{LH}$ x Argument Type^2	2.637	3.212	0.821
$Power_{LCT}$ x Argument Type^3	17.776	3.227	5.509***
$Power_{LH}$ x Argument Type^3	17.734	3.212	5.520***

Note. $Power_{LCT}$ = dummy-coded low power group vs. control group, and $Power_{LH}$ = dummy-coded low power group vs. high power group (with low power group = 0, control group = 1, and high-power group = 2). Argument Type, Argument Type^2, and Argument Type^3 represent the orthogonal first-, second-, and third-order polynomials respectively. * $p < .05$; ** $p < .01$; *** $p < .001$

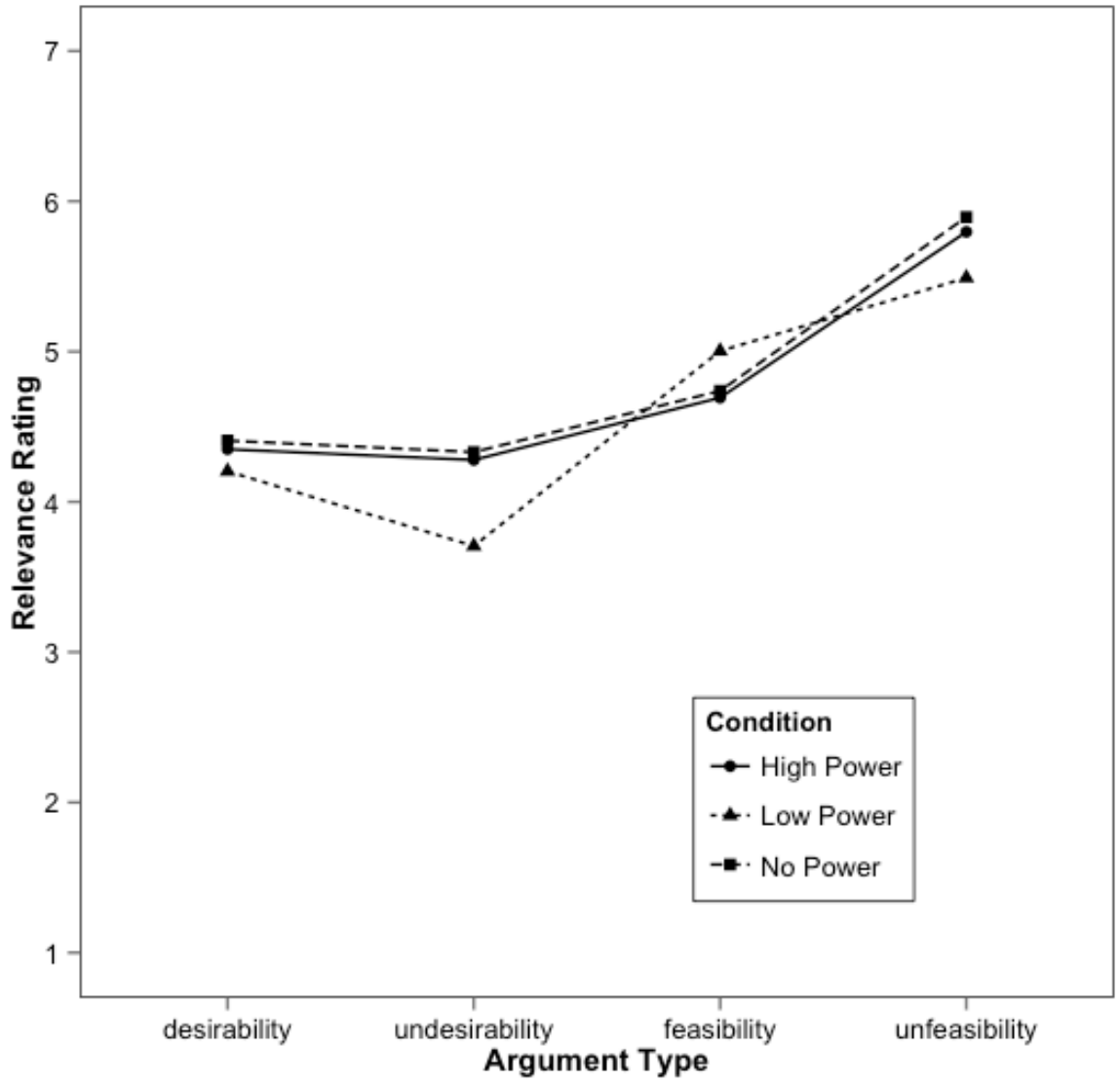

Figure 2. Mean ratings of relevance for experimental conditions and argument type from Study 1.

The results show that power influences the patterns in the ratings of relevance of (un-)desirability and (un-)feasibility arguments. Based on the propositions of two theories of power and the detailed account in Magee and Smith (2013) I hypothesized and found different patterns of ratings of relevance for low-power versus high-power individuals. The results of Study 1 provided support for Hypothesis 1. The meta-analysis reported in the section "Meta-Analytical Results and Discussion" (for an overview, see Figure 7) shows that

low-power individuals displayed a preference for feasibility arguments, which is in line with Hypothesis 5. The patterns of ratings in Study 1 were not in line with Hypothesis 2, 3, and 4.

3.4. Study 2

Study 2 was conducted as an online replication of Study 1 with a sample of high power individuals, i.e. persons in a leadership position. The methods remained unchanged, and I report only the deviations with regard to the recruitment and characteristics of the sample.

3.4.1. Methods

Participants. Participants were recruited and reimbursed via a panel firm located in Germany. The participants received a link via e-mail and completed the study on their personal computers. To be sure that the sample consisted of leaders, I included the question "Do you currently lead employees?" after the welcome screen and automatically screened out participants who were not in a leadership position. I also asked for leadership experience measured in years in the demographic questions.

209 participants completed the online study. To ensure high data quality, I excluded participants who completed the study with interruptions (that is, when the online survey tool registered 30min of inactivity or more) or said that their data should not be included ("In your opinion, should your data be included in the analysis?"). 196 participants evenly distributed across experimental conditions remained in the final sample (88 female, mean age was 43.7 years, *SD* = 11.0). On average, they had 11.7 years of experience as a leader (*SD* = 9.3 years).

3.4.2. Results and Discussion

Manipulation check. A one-way ANOVA showed that the sense of power varied across conditions, $F(2, 193) = 37.09$, $p < .001$, $\eta_p^2 = 0.28$. As expected, the participants in the high-power group expressed the most power ($M = 5.52$, $SD = 0.97$), participants in the low

power group expressed the least power (M = 3.56, SD = 1.67). Participants in the control group fell in between (M = 4.45, SD = 1.12). A subsequent Tukey-HSD test revealed that all groups differed significantly from each other.

Power influences patterns of relevance ratings. Results from the mixed-effects model are displayed in Table 6 and the patterns of relevance ratings are displayed in Figure 3. I expected to find that high-power individuals and low-power individuals rate the relevance of the four types of arguments for making a decision differently, again providing support for Hypothesis (H1). The significant interaction between power and the cubic term of argument type, Wald $\chi^2(1) = \text{-}4.06, p < .001$, indicated that the patterns of ratings of relevance differed significantly between the low-power and the high-power group. Figure 3 shows that the four argument types were rated differently within each power group and that there are differences in the ratings of relevance between the groups, which is reflected in the significant interaction term. The patterns of ratings of relevance differed significantly between the high-power group and the control group as well, Wald $\chi^2(1) = 3.13, p < .01$. The significant main effect of the cubic term of argument type, Wald $\chi^2(1) = 2.29, p < .05$, indicated that the ratings of relevance differed from each other. All other effects did not reach statistical significance (all ps > .05). Study 2 provided further support for Hypothesis 1. The meta-analysis reported in the section "Meta-Analytical Results and Discussion" (for an overview, see Figure 7) shows that high-power individuals displayed a preference for positive arguments, which is in line with Hypothesis 2. Low-power individuals displayed a preference for feasibility arguments, which is in line with Hypothesis 5. The patterns of ratings in Study 2 were not in line with Hypothesis 3 and 4.

Table 6

Fixed-Effects Results from the Mixed-Effects Model in Study 2

Effect	Estimate	*SE*	Wald χ^2
Intercept	0.069	0.118	0.588
$Power_{LCT}$	-0.136	0.169	-0.801
$Power_{LH}$	-0.076	0.171	-0.448
Argument Type	-0.008	4.984	-0.002
Argument $Type^2$	3.282	2.395	1.370
Argument $Type^3$	5.491	2.395	2.292*
$Power_{LCT}$ x Argument Type	-0.266	7.156	-0.037
$Power_{LH}$ x Argument Type	0.303	7.245	0.042
$Power_{LCT}$ x Argument $Type^2$	-5.569	3.439	-1.619
$Power_{LH}$ x Argument $Type^2$	-4.536	3.481	-1.303
$Power_{LCT}$ x Argument $Type^3$	-3.078	3.439	-0.895
$Power_{LH}$ x Argument $Type^3$	-14.131	3.481	-4.058***

Note. $Power_{LCT}$ = dummy-coded low power group vs. control group, and $Power_{LH}$ = dummy-coded low power group vs. high power group (with low power group = 0, control group = 1, and high-power group = 2). Argument Type, Argument $Type^2$, and Argument $Type^3$ represent the orthogonal first-, second-, and third-order polynomials respectively. * $p < .05$; ** $p < .01$; *** $p < .001$

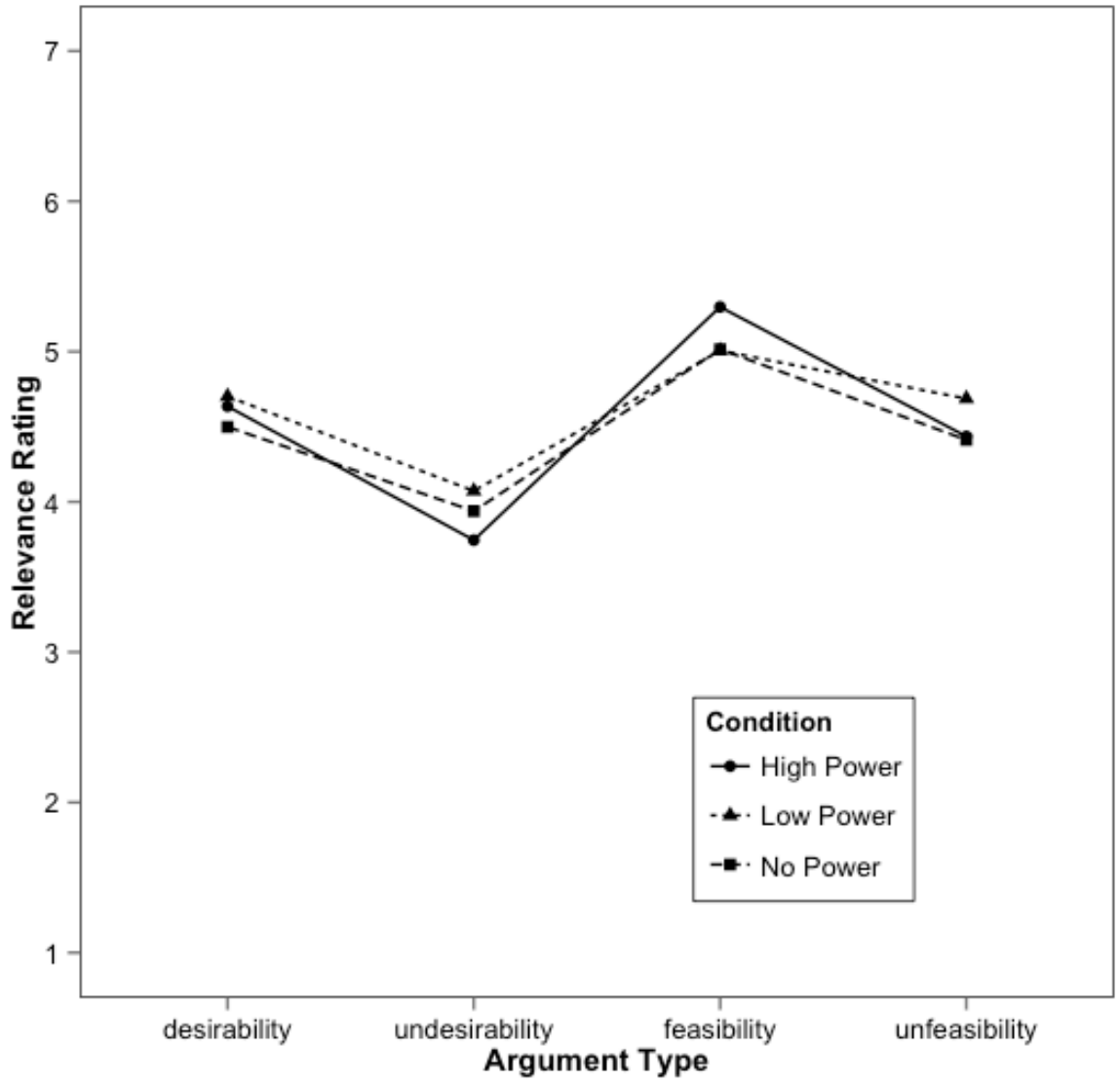

Figure 3. Mean ratings of relevance for experimental conditions and argument type from Study 2.

3.5. Study 3

In Study 3, I compared leaders to working adults without leadership responsibility. I used the same manipulation, materials, and a similar procedure as in Study 2. To analyze effects of structural power, I changed the study design to account for the new experimental factor, namely professional role.

3.5.1. Methods

Research design. The study was based on a 2 (leader vs. employee) x 2 (high power vs. low power) x 4 (desirable vs. undesirable vs. feasible vs. unfeasible) design with the first two being between-subjects factors and the third being a within-subject factor.

Participants. Participants were again recruited and reimbursed via a panel firm located in Germany. To ensure high data quality in the online sample, I included two more questions ("Is German your mother tongue?" and "Are you currently employed with or without leadership responsibility?").

222 participants completed the online study. I excluded participants who completed the study with interruptions or said that their data should not be included (self-ratings on the item "In your opinion, should your data be included in the analysis?"). 217 participants remained in the final sample (111 female, mean age was 42.32, *SD* = 9.93 years).

Materials and procedure. I used a different manipulation check compared to Study 2, because I also included job-related power in this study, and altered the procedure slightly in that the manipulation check was presented immediately after the power manipulation.

At the end of Study 3, I asked for more detailed demographic information, including: position with or without leadership responsibility, line of business, position in the hierarchy at work, degree of formal authority awarded by the position at work. I also included measures to assess potential covariates: positive and negative affect, social distance, abstract thinking, need for cognition, behavioral approach/inhibition, social dominance orientation, and job-related power.

The participants responded to two items that we included to assess if insufficient effort responding (Huang, Curran, Keeney, Poposki, & DeShon, 2012) had occurred in the information processing task ("When rating the items I focused on the last two words" and "I

have read the topics carefully", 7-point scale). I also used a range of control questions (see Table 13, Appendix G).

Positive negative affect scale. After the completing the power manipulation and the manipulation check the participants responded to the positive negative affect scale (PANAS, Krohne, Egloff, Kohlmann, & Tausch, 1996). After completing the PANAS, the participants worked on the same information evaluation task as in Studies 1 and 2.

Social distance. Participants completed the inclusion of the other in the self scale (IOS scale, Aron, Aron, & Smollan, 1992), which I used as a measure of social distance (Magee & Smith, 2013). It is a one-item measure with seven Venn-diagrams that prompts participants to choose the diagram that best describes their closest relationship.

Abstract thinking. Participants responded to 24 items from the Dyads of Triads Task (DOT, Bowers, Regehr, Balthazard, & Parker, 1990), which is a measure for abstract thinking that asks participants to judge whether a triad of words is coherent and then asks them to provide a possible solution that explains the relationship of all four words. I used the task as P. K. Smith and Trope (2006) did, but with fewer triads as Bolte, Goschke, and Kuhl (2003) used in their study. The measure of interest was the number of triads for which participants provided a correct solution (Bowers et al., 1990; P. K. Smith & Trope, 2006).

Social dominance orientation. Participants responded to the social dominance orientation scale (Pratto, Sidanius, Stallworth, & Malle, 1994).

Need for cognition. Participants responded to the scale that assesses need for cognition (Bless, Wänke, Bohner, Fellhauer, & Schwarz, 1994).

Regulatory focus. Participants completed the regulatory focus scale to assess their dispositional promotion and prevention focus (Fellner, Holler, Kirchler, & Schabmann, 2007).

Behavioral approach/inhibition. Participants responded to the BIS/BAS scale (Carver & White, 1994; Strobel, Beauducel, Debener, & Brocke, 2001) as a measure for behavioral activation and inhibition. This measure has been used in previous studies involving power (Lammers, Galinsky, Gordijn, & Otten, 2008).

Job-related power. I assessed job-related power with four items from the sense of power scale that I slightly adapted to refer to the participant's jobs (e.g., "In my job I have a great deal of power", Anderson & Galinsky, 2006; Anderson et al., 2012; See et al., 2011, $\alpha = 0.88$).

Results and Discussion

In Study 3, I assessed a range of variables to test for potential mediators. Since there were no correlations between power and the potential mediators (see Table 14, Appendix H), I focused my analyses on the influence of power on the relevance ratings of the four argument types. The analyses of the control questions revealed no significant differences between power conditions either. Controlling for insufficient effort responding yielded a pattern of results similar to the one reported here. The patterns of ratings of relevance in Study 3 are similar to those in Studies 2 and 4, where I did not use additional measures, thus the use of the PANAS scale before the argument evaluation task did not have an effect.

Manipulation check. We used a six-item measure in accordance with Lammers et al. (2013) that included the three semantic differentials used in Study 1 and three additional semantic differentials (P. K. Smith et al., 2008, e.g., "submissive vs. dominant", 7-point scale, $\alpha = 0.95$) to measure participants' sense of power.

3.5.2. Results and Discussion

Manipulation check. A 2 (role) x 2 (power) ANOVA showed a significant main effect of power, $F(1, 213) = 156.99$, $p < .001$, $\eta_p^2 = 0.42$. Neither the main effect of role, $F(1, 213) = 0.702$, $p = .403$, $\eta_p^2 = 0.002$, nor the interaction, $F(1, 213) = 0.510$, $p = .476$, $\eta_p^2 =$

0.001, were significant. That is, the power manipulation worked as intended. The leaders in the high-power group expressed more power (M = 5.10, SD = 1.40) than the leaders in the low power group (M = 2.48, SD = 1.36). The employees in the high-power group expressed more power (M = 4.80, SD = 1.38) than the employees in the low power group (M = 2.46, SD = 1.37).

Power influences the ratings of relevance. Results from the mixed-effects model 1 are displayed in Table 7 and the patterns of relevance ratings are displayed in Figure 4. I expected to find further support for Hypothesis 1 and to find the professional role to influence the ratings of relevance in a similar way as the power manipulation did.

Significant three-way interactions between professional role, power, and argument type demonstrated that professional role and power interact in how they influence the ratings of relevance of the four argument types (see Figure 4). The significant three-way interaction between power, professional role, and the cubic term of argument type, Wald $\chi^2(1) = 21.41$, $p < .001$, indicated that the patterns of ratings of relevance differed significantly between the low power and the high-power group and that these differences varied between leaders and non-leaders. The significant three-way interaction between power and the linear term of argument type, Wald $\chi^2(1) = 23.36$, $p < .05$, indicated that there also was a significant difference in the mean ratings of relevance between leaders and non-leaders, and the low power and the high-power group.

A significant two-way interaction between professional role and the cubic term of argument type, Wald $\chi^2(1) = 16.50$, $p < .001$, indicated that the patterns of ratings of relevance differed between leaders and non-leaders. A significant two-way interaction between power and the cubic term of argument type, Wald $\chi^2(1) = -8.51$, $p < .05$, indicated that that the patterns of ratings of relevance differed between the low power group and the high-power group. The significant two-way interactions between professional role and the

cubic term of argument type, and between power and the cubic term of argument type show that both factors influenced the ratings of relevance independently of each other. In the case of power, these findings are in line with the results from Study 1 and Study 2. In general, the results are in line with theorizing and empirical results that demonstrate independent effects of structural power and the sense of power (Anderson et al., 2012; Galinsky et al., 2003; P. K. Smith & Hofmann, 2016; Tost, 2015).

A significant two-way interaction between power and the linear term of argument type, Wald $\chi^2(1) = \text{-}18.95, p < .05$, indicated that there was a significant difference in the mean ratings of relevance between the low power and the high-power group. The ratings of relevance differed from each other, as indicated by the significant main effect of the cubic term of argument type, Wald $\chi^2(1) = 2.29, p < .05$. All other effects did not reach statistical significance (all *p*s > .05). The results of Study 3 provide further support for Hypothesis 1. The meta-analysis reported in the section "Meta-Analytical Results and Discussion" (for an overview, see Figure 7) shows that high-power individuals displayed a preference for positive arguments, which is in line with Hypothesis 2. Low-power individuals displayed a preference for feasibility arguments, which is in line with Hypothesis 5. The patterns of ratings in Study 3 were not in line with Hypothesis 3 and 4.

Table 7

Fixed-Effects Results from the Mixed-Effects Model in Study 3

Effect	Estimate	*SE*	Wald χ^2
Intercept	-0.121	0.127	-0.947
$Power_{LH}$	0.081	0.179	0.453
$Power_{FK}$	0.237	0.172	1.381
Argument Type	8.237	6.203	1.328
Argument $Type^2$	-2.199	2.674	-0.822
Argument $Type^3$	7.422	2.674	2.775**
$Power_{LH}$ x $Power_{FK}$	-0.183	0.241	-0.760
$Power_{LH}$ x Argument Type	-18.946	8.727	-2.171*
$Power_{LH}$ x Argument $Type^2$	6.125	3.763	1.628
$Power_{LH}$ x Argument $Type^3$	-8.512	3.763	-2.262*
$Power_{FK}$ x Argument Type	-9.401	8.353	-1.125
$Power_{FK}$ x Argument $Type^2$	1.791	3.601	0.619
$Power_{FK}$ x Argument $Type^3$	-16.503	3.601	-4.582***
$Power_{LH}$ x $Power_{FK}$ x Argument Type	23.356	11.736	1.990*
$Power_{LH}$ x $Power_{FK}$ x Argument $Type^2$	-6.745	5.060	-1.333
$Power_{LH}$ x $Power_{FK}$ x Argument $Type^3$	21.410	5.060	4.231***

Note. $Power_{LH}$ = dummy-coded low power group vs. high power group (with low power group = 0 and high-power group = 1), and $Power_{FK}$ = dummy-coded employee group vs. manager group (employees = 0 and managers = 1). Argument Type, Argument $Type^2$, and Argument $Type^3$ represent the orthogonal first-, second-, and third-order polynomials respectively. * $p < .05$; ** $p < .01$; *** $p < .001$

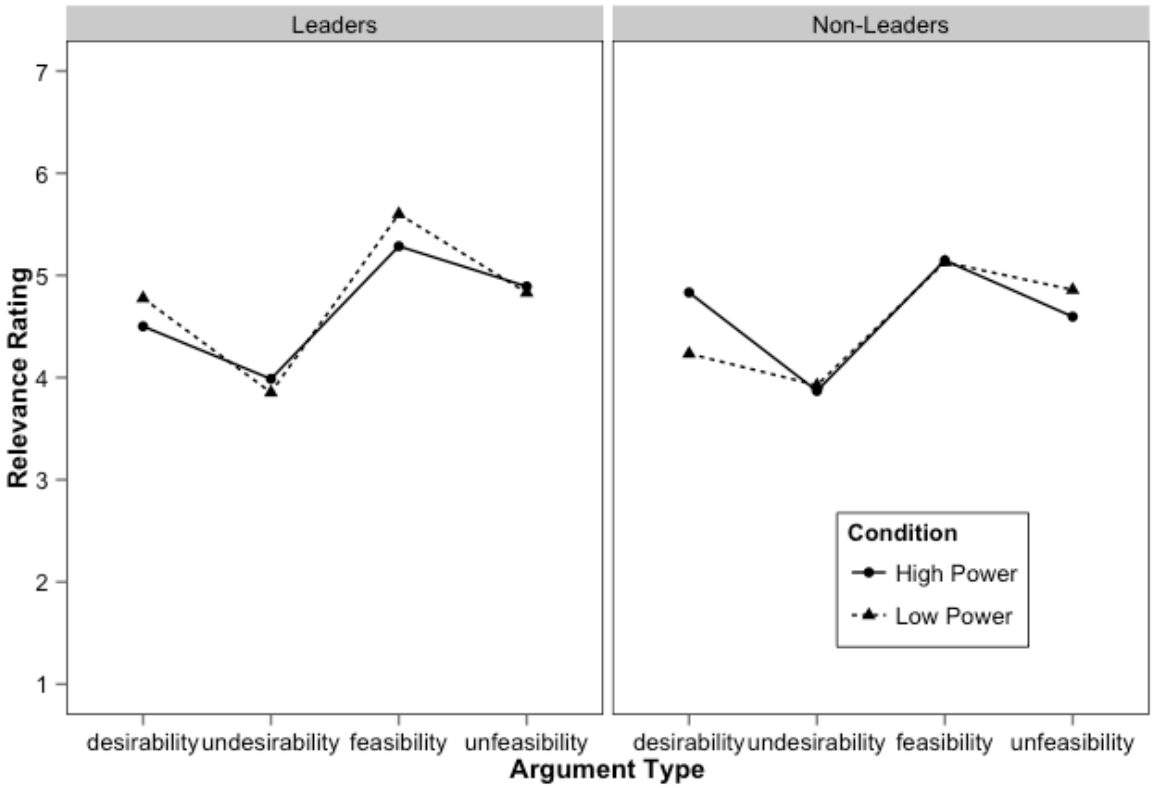

Figure 4. Mean ratings of relevance for experimental conditions and argument type from the leader sample from Study 3.

3.6. Study 4

Study 4 features a sample of non-leaders, a different power manipulation than in Studies 1-3, an additional dependent variable, and a slightly altered study design. These steps aim to increase the overall reliability and external validity of our findings. I report only deviations from the experimental protocol of Study 2.

3.6.1. Methods

Participants. In Study 4, I automatically screened out participants who were in a leadership position as I intended to have a sample of non-leaders, because Study 3

demonstrated the influence of structural power. 196 participants completed the study. To ensure high data quality, I excluded participants who completed the study with interruptions or said that their data should not be included (self-ratings on the item "In your opinion, should your data be included in the analysis?"). 179 participants evenly distributed across the experimental conditions remained in the final sample (139 female, mean age was 41.9, *SD* = 12.5 years).

I used the same demographic questions as in Study 1 and added questions for: position in the hierarchy at work, and whether the participant's job allows them to make decisions.

3.6.2. Material and procedure

Power manipulation. Participants were randomly assigned to one of the power conditions. I applied the same power manipulation as Moon and Chen (2014), who published it as an online supplement. The manipulation consisted of visualizing oneself as the interviewer or the interviewee in a job interview (Moon & Chen, 2014). Participants were provided with an image of a table with a comfortable chair for the interviewer and an uncomfortable chair for the interviewee as an anchor for their visualization. They were bound to remain on the page for 60 seconds, after which they could proceed with the experiment.

Manipulation Check. I used five semantic differentials from P. K. Smith et al. (2008, e.g., "submissive vs. dominant", 7-point scale, $\alpha = 0.90$) as a manipulation check. I also used five control items that aimed to assess whether the power manipulation had also elicited feelings of responsibility (Tost, 2015) and whether participants found it easy to imagine the scenario (Schwarz et al., 1991) to better understand the effects of the power manipulation. I used an attention check and the same items to test for insufficient effort responding (Huang et al., 2012) as in Study 3.

Information-processing task. In Study 4, each participant was presented with 20 out of 20 pre-tested statements in random order to cancel out random effects that were present in Studies 1-3 due to my item sampling strategy.

In addition to the ratings of relevance, I added a second dependent variable that was more oriented towards a concrete behavior. As in Fischer et al. (2011), I asked the participants how likely it was that they would read the entire report that was summarized by the statements ("How likely is it that you would read the entire report summarized by these statements?"; 7-point scale: 1 = not at all, 7 = very likely).

3.6.3. Results and Discussion

Manipulation check. A one-way ANOVA showed that the sense of power varied across conditions according to the power manipulation, $F(2, 176) = 29.89$, $p < .001$, $\eta_p^2 = 0.25$. As expected, the participants in the high-power group expressed the most power ($M = 5.53$, $SD = 1.08$), participants in the low power group expressed the least power ($M = 4.33$, $SD = 1.28$). Participants in the control group fell in between ($M = 3.88$, $SD = 1.20$). A subsequent Tukey-HSD test revealed that the high power and control group differed significantly from each other, while there was no significant difference between the low power group and the control group.

Power influences the ratings of relevance. Results from the mixed-effects model 1 are displayed in Table 8. As in Study 1, Study 2, and Study 3 the results show that power influences the patterns in the ratings of relevance of (un-)desirability and (un-)feasibility arguments. That is, high-power individuals and low-power individuals rate the relevance of the four types of arguments for making a decision differently. The same pattern emerged for the second dependent variable, that is, high-power individuals would likely read different types of reports than low-power individuals.

The significant interaction between power and the cubic term of argument type, Wald $\chi^2(1) = 2.62, p < .01$, indicated that the patterns of ratings of relevance differed significantly between the low power and the high-power group. The patterns of ratings of relevance differed significantly between the low power group and the control group as well, Wald $\chi^2(1) = 3.84, p < .001$. The significant main effect of the cubic term of argument type, Wald $\chi^2(1) = -3.06, p < .01$, indicated that the ratings of relevance varied across the four argument types.

The second dependent variable, the probability that particpants would read the entire report summarized by one of the arguments, showed the same significant effects. That is, a significant interaction between power and the cubic term of argument type, Wald $\chi^2(1) = 3.81, p < .001$, and a significant main effect of the cubic of argument type, Wald $\chi^2(1) = -3.53, p < .001$. None of the other effects reached statistical significance (all *p*s > .05).

Controlling for insufficient effort responding, feelings of responsibility, or the ease of imagining the scenario in the power manipulation by excluding participants yielded patterns of results similar to those reported in this section.

The results of Study 4 provide further support for Hypothesis 1. The meta-analysis reported in the section "Meta-Analytical Results and Discussion" (for an overview, see Figure 7) shows that high-power individuals displayed a preference for positive arguments, which is in line with Hypothesis 2. Low-power individuals displayed a preference for feasibility arguments, which is in line with Hypothesis 5. The patterns of ratings in Study 4 were not in line with Hypothesis 3 and 4.

Table 8

Fixed-Effects Results from the Mixed-Effects Model in Study 4

Effect	Estimate	*SE*	Wald χ^2
Intercept	-0.074	0.131	-0.562
$Power_{LCT}$	0.096	0.184	0.522
$Power_{LH}$	0.125	0.186	0.671
Argument Type	0.239	3.452	0.069
Argument $Type^2$	-0.418	2.193	-0.191
Argument $Type^3$	-6.711	2.193	-3.060**
$Power_{LCT}$ x Argument Type	-2.828	4.842	-0.584
$Power_{LH}$ x Argument Type	2.199	4.881	0.451
$Power_{LCT}$ x Argument $Type^2$	0.295	3.076	0.096
$Power_{LH}$ x Argument $Type^2$	0.963	3.102	0.311
$Power_{LCT}$ x Argument $Type^3$	11.820	3.076	3.842***
$Power_{LH}$ x Argument $Type^3$	8.141	3.102	2.625**

Note. $Power_{LCT}$ = dummy-coded low power group vs. control group, and $Power_{LH}$ = dummy-coded low power group vs. high power group (with low power group = 0, control group = 1, and high-power group = 2). Argument Type, Argument $Type^2$, and Argument $Type^3$ represent the orthogonal first-, second-, and third-order polynomials respectively. * $p < .05$; ** $p < .01$; *** $p < .001$

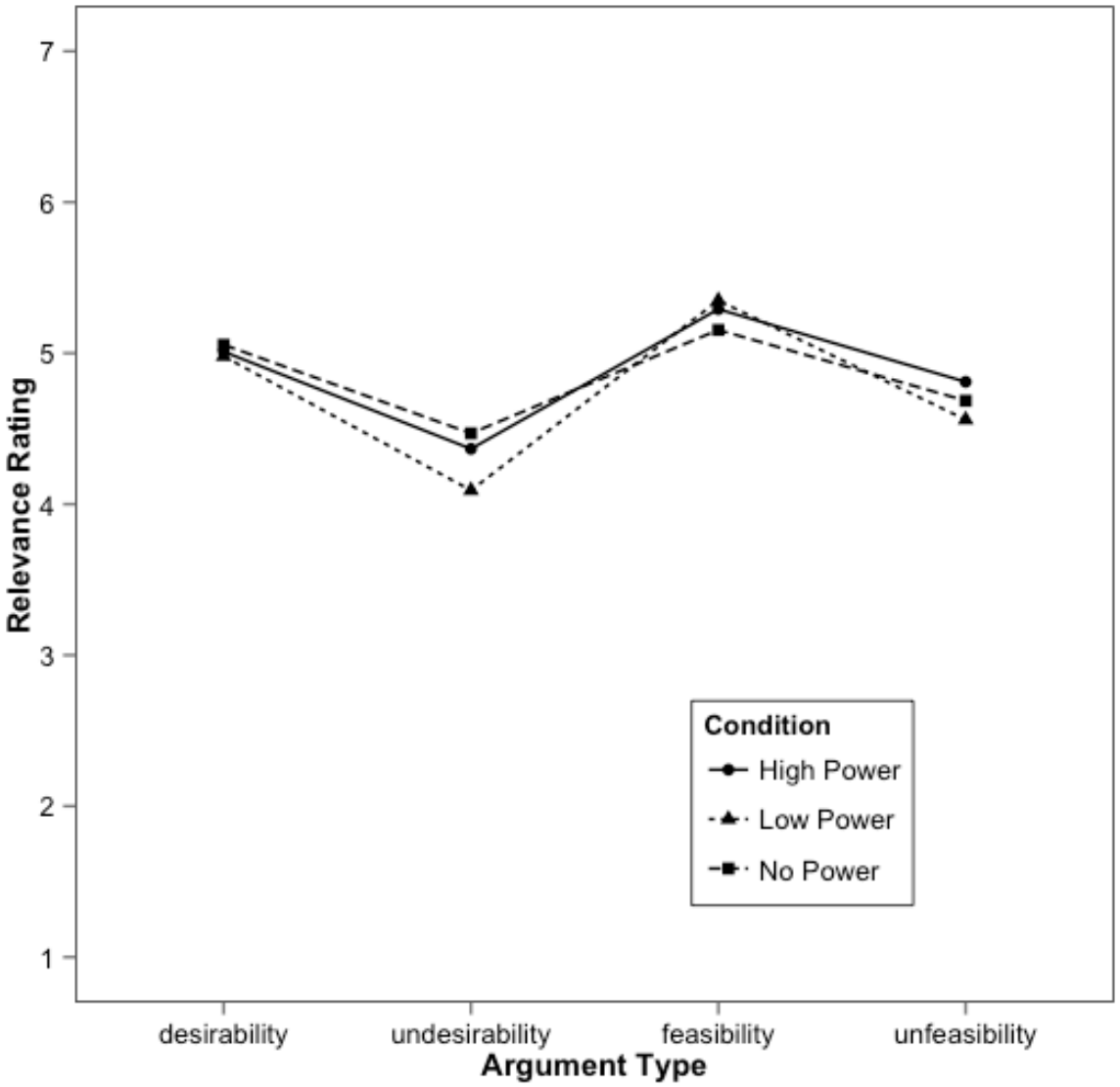

Figure 5. Mean ratings of relevance for experimental conditions and argument type from Study 4.

3.7. Meta-Analytical Results and Discussion

I conducted a meta-analysis of the four quantitative studies to test Hypotheses 2-5 and to achieve reliable insights into the differences power causes in the evaluation of (un-)desirability and (un-)feasibility arguments. Following the methodology to conduct single-paper meta-analyses proposed by McShane and Böckenholt (2017), I defined a set of contrasts to test my hypotheses across the four quantitative studies reported here. Table 9

displays the contrast weights I assigned, as well as the results from the meta-analysis. Figure 6 shows the results of the planned contrasts for each study as well as the overall meta-analytical effect.

Table 9

Contrast Weights and Results from the Meta-Analysis across four Studies

Contrast (Figure 7)	Power-Group	Contrast-weights				Estimate	*SE*
		$MEAN_D$	$MEAN_{UD}$	$MEAN_F$	$MEAN_{UF}$		
2	High	0.5	-0.5	0.5	-0.5	0.3694	0.1821
	Low	0	0	0	0		
3	High	0.5	0.5	-0.5	-0.5	-0.6725	0.1781
	Low	0	0	0	0		
4	High	0	0	0	0	-0.4790	0.1829
	Low	-0.5	0.5	-0.5	0.5		
5	High	0	0	0	0	0.7592	0.1772
	Low	-0.5	-0.5	0.5	0.5		

Note. $MEAN_D$ = desirability, $MEAN_{UD}$ = undesirability, $MEAN_F$ = feasibility, and $MEAN_{UF}$ = unfeasibility information, Estimate = weighted mean difference across four studies, *SE* = standard error. The contrast weights were chosen such that a higher positive value indicates a better match with the respective hypothesis. They are numbered according to the hypothesis they aim to test.

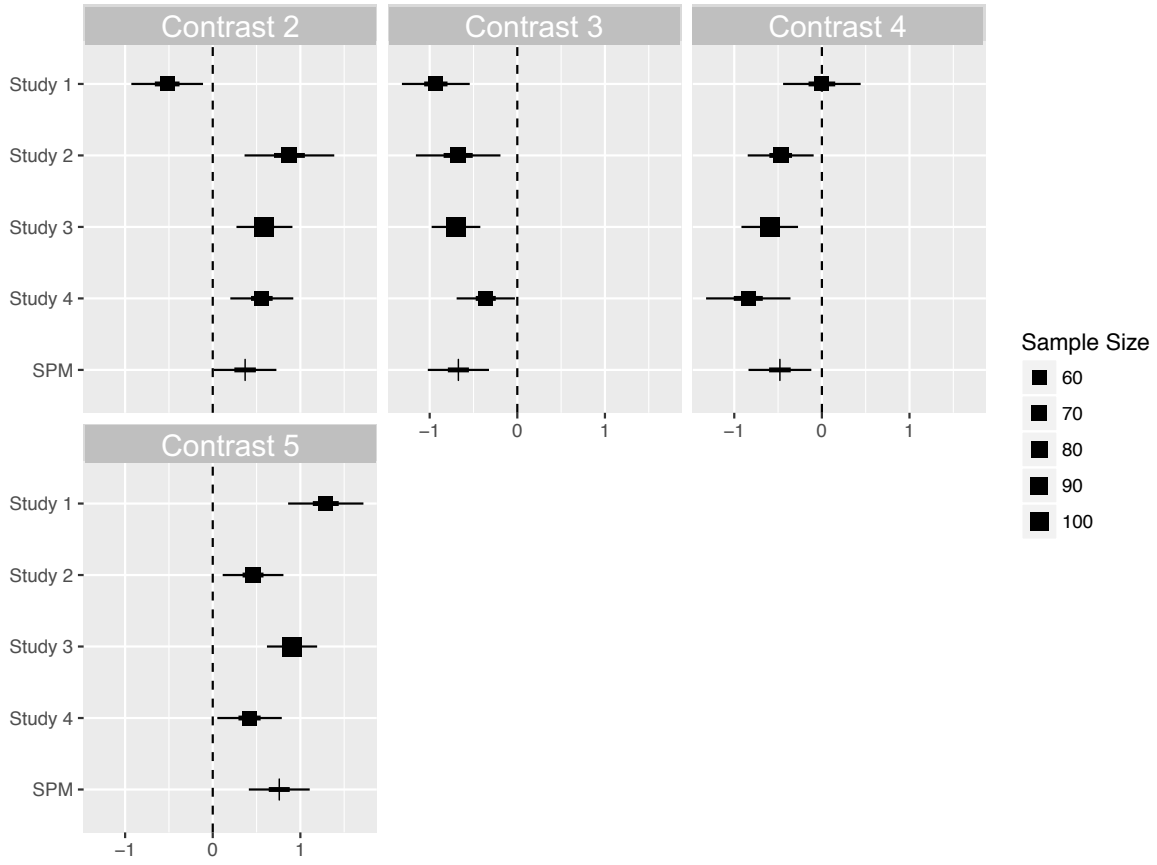

Figure 6. Results of the single-paper meta-analysis. Positive values indicate fit with the prediction derived from the respective theory. SPM stands for single-paper meta-analysis and denotes the overall mean contrast. Contrast 2 displays the results for the AIT's predictions for high-power individuals (Hypothesis 2). Contrast 3 displays the results for the SDT's predictions for high-power individuals (Hypothesis 4). Contrast 4 displays the results for the AIT's predictions for low-power individuals (Hypothesis 3). Contrast 5 displays the results for the SDT's predictions for low-power individuals (Hypothesis 5). The figure was created with the help of the online tool by McShane, B. B., & Böckenholt, U. (2017). Single-Paper Meta-Analysis: Benefits for Study Summary, Theory Testing, and Replicability. *Journal of Consumer Research*, 43(6), 1048-1063. doi:10.1093/jcr/ucw085

Across the four studies, high-power individuals rated desirability and feasibility arguments as more relevant than undesirability and unfeasibility arguments. That is, high-power individuals rated positive information as more relevant than negative information, Contrast 2 = 0.37, 95% CI [0.01, 0.73]. This supports Hypothesis 2. Low-power individuals rated (un-)feasibility arguments as more relevant than (un-)desirability arguments. That is, low-power individuals rated feasibility information more relevant than desirability information, Contrast 5 = 0.76, 95% CI [0.41, 1.11]. This supports Hypothesis 5.

The other hypotheses did not receive empirical support or I found patterns of ratings opposite to what was expected. Specifically, high-power individuals rated feasibility information as more relevant than desirability information, Contrast 3 = -0.67, 95% CI [-1.02, -0.32]. Thus, I had to reject Hypothesis 4. Also, low-power individuals rated positive information as more relevant than negative information, Contrast 4 = -0.48, 95% CI [-0.84, -0.12]. I conclude that Hypothesis 3 also has to be rejected.

Together, the results of Study 1 to 4 provide support for the predictions derived from approach/inhibition theory for high-power individuals and the predictions derived from social distance theory for low-power individuals.

3.8. Study 5

Study 5 aimed at increasing the ecological validity of our findings. I tested the influence of power on the active generation of arguments rather than evaluations of a given set of arguments.

3.8.1. Methods

Research design. The experiment was based on a 3 (high power vs. low power vs. control) between-groups design.

Participants. Participants were recruited and reimbursed via a panel firm located in Germany. The participants received a link via e-mail and completed the study on their personal computers. I included the question "Do you currently lead employees or have you lead employees in the past?" on the second page of the survey and automatically screened out participants with leadership experience. I also included a question about the leadership experience (in years) in the demographic questions.

129 participants completed the online study without interruption. To ensure high data quality for the qualitative data analysis, I excluded participants who typed in random letters, answered in one word, said that they knew nothing about the given topic, or stated that their data should not be included ("In your opinion, should your data be included in the analysis?"). 94 participants evenly distributed across the experimental conditions remained in the final sample (39 female, mean age = 38.0 years, *SD* = 11.3).

3.8.2. Material and procedure

Power manipulation. Participants were randomly assigned to one of the power conditions. I applied the same power manipulation as in Study 4.

Manipulation check. I used the same manipulation check as in Study 3 ($\alpha = 0.89$) and the same control items as in Study 4.

Argument generation task. After completing the manipulation check, participants were prompted to discuss the topic "autonomous vehicles" by writing arguments into the designated fields for text entry. They were bound to generate at least two arguments before they could proceed to the last page, where they were asked for their attitude towards automated vehicles and for demographic information (as in Study 4).

Data preparation. Questions, such as "will it be affordable?", were excluded from the analysis, for they are no arguments. Arguments devoid of valence, desirability, and feasibility information, such as "is the future", were excluded from the analysis, too. In total 221

arguments remained as the basis for my analyses. The arguments consisted of half sentences mostly. When two arguments were linked in one statement the raters focused on the central argument. Examples of the statements in the final sample included: “could be dangerous,” or “automated vehicles make driving safer,” or “The car takes control of the traffic by means of sensors and connections to other cars.”

Rating of valence and type of arguments. Two research assistants blind to the purpose of the study rated the valence and argument domain, that is, whether it was related to desirability or feasibility. Due to problems with achieving sufficient agreement on the argument domain, I conducted a separate study to identify prototypical arguments for each domain: 62 participants, who did not take part in the main study, rated a subset of 50 arguments on valence and domain. From this data, we identified arguments that were rated very reliably as either desirability or feasibility arguments by analyzing frequencies. Then, the first and the second author used the prototypical arguments to rate the remaining 171 arguments. Good levels of interrater agreement were achieved as indicated by Cohen’s Kappa for valence (kappa = 0.96) and argument domain (kappa = 0.84). Cases of disagreement were resolved in a discussion.

3.8.3. Results and Discussion

Manipulation check. A one-way ANOVA showed that our manipulation successfully varied the sense of power across conditions, $F(2, 91) = 14.42$, $p < .001$, $\eta_p^2 = 0.24$. As expected, participants in the high-power group expressed the most power ($M = 5.22$, $SD = 1.03$), and participants in the low power group expressed the least power ($M = 3.93$, $SD = 0.85$). Participants in the control group ($M = 4.28$, $SD = 1.05$) fell in between, as expected. A subsequent Tukey-HSD test revealed that the high power and control group differed significantly from each other, while there was no significant difference between the low power group and the control group.

Analyses of the argument types. We compared the average number of arguments for each argument type provided by the experimental groups (see Figure 7). We expected to find differences in the number of arguments per type (desirability, undesirability, feasibility, and unfeasibility) provided by low- versus high-power individuals. The results provide support for Hypothesis 1, as the number of arguments generated per type differ between the low-power and the high-power group. Due to differences in the processing of positive and negative information (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Unkelbach, Fiedler, Bayer, Stegmüller, & Danner, 2008), and the fact that participants generated arguments with regard to only one topic, which likely led to the occurrence of confirmation bias (Nickerson, 1998), I did not expect to be able to conduct tests of the more specific Hypotheses 2 to 5 in this study.

High-power individuals generated fewer desirability arguments than low-power individuals (9 vs. 16 respectively) and also fewer undesirability arguments (24 vs 31 respectively). High-power individuals generated more feasibility arguments than low-power individuals (18 vs. 8 respectively) and both groups generated nearly the same number of unfeasibility arguments (25 vs. 26 respectively).

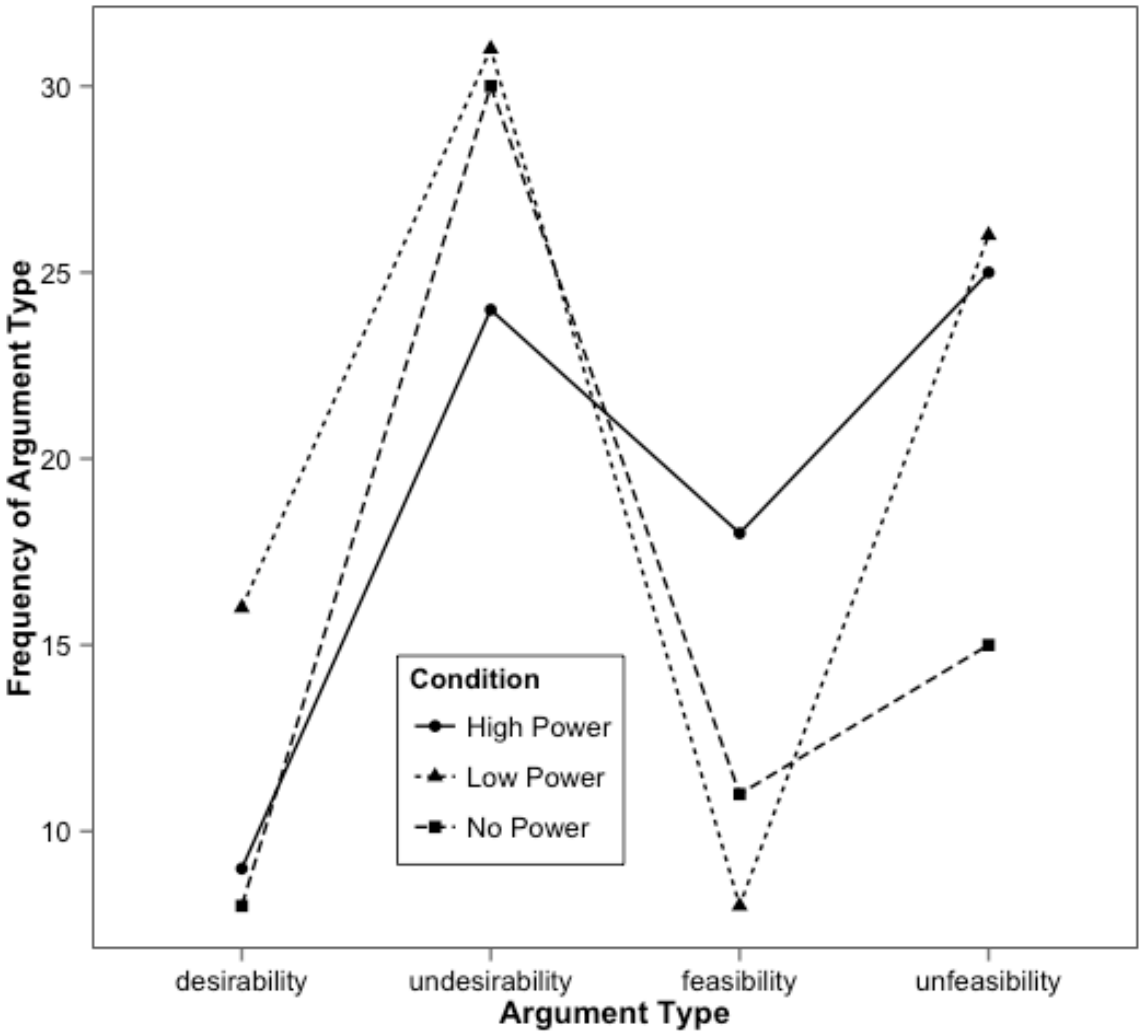

Figure 7. Frequencies of the argument types per group from Study 5.

The results of Study 5 show that power influences processes beyond the explicit evaluation of the relevance of an argument. The arguments generated by the participants can be seen as an indication of their mental model (Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999). The arguments generated also provide information about what aspects participants found to be relevant, because conversational norms (Grice, 1975) demand they make relevant statements. Study 5 increases the ecological validity of my findings by showing that power influences what aspects of mental models are seen as relevant.

3.9. General Discussion

In five studies, I show that power influences the degree to which desirability and feasibility arguments are regarded as relevant for making decisions in complex organizational scenarios. This effect was robust across different samples, power manipulations, and operationalizations of the dependent variable. A single-paper meta-analysis of the results of the quantitative Studies 1 to 4 revealed that high-power individuals preferred positive arguments, irrespective of the desirability and feasibility domain, and that low-power individuals preferred feasibility arguments, irrespective of valence. The results of an additional qualitative study, Study 5, demonstrate that power influences not only explicit ratings of the relevance of arguments, but also the active generation of arguments. Due to the experimental manipulation of power, I conclude that power caused the differences in the evaluation of arguments. Further, the momentary sense of power (Galinsky et al., 2003) produced effects independently of and beyond the effects of structural power (Tost, 2015), as in Studies 2 and 3, or the relative lack of structural power, as in Studies 1, 3, 4, and 5.

Contributions and implications for theory. The present research responds to calls for more specific examinations of the effects of power on organizational decision-making (Fast et al., 2012) and for a direct test of the influence of power on the processing of desirability and feasibility information (Magee & Smith, 2013). First, I contribute to the literature on organizational decision-making by showing that power systematically influences leaders' and non-leaders' evaluations of different types of arguments. The influence of power on organizational decision-making has received only little attention despite its obvious importance (Fast et al., 2012) and potentially profound economical and ethical consequences. The present research extends a line of inquiry that demonstrates systematic influences of power on organizational decision-makers: They are overly confident in their decisions (Fast et al., 2012) and take less advice (See et al., 2011; Tost et al., 2012a). In addition, I found that

high power entails a preference for positive arguments, irrespective of the desirability and feasibility domain, and low power entails a preference for feasibility arguments, irrespective of valence. The present research demonstrates that rather than a "more or less" on one dimension the effects of power on different dimensions need to be taken into account. In summary, it appears that power does not only tip *the* balance but many different balances simultaneously.

Second, responding to a call by Magee and Smith (2013), I identify boundary conditions for two different theories of power with regard to the processing of desirability and feasibility information. My results are in line with the predictions derived from approach/inhibition theory (Keltner et al., 2003) for high-power individuals and the predictions derived from social distance theory (Magee & Smith, 2013) for low-power individuals. Based on my empirical results, it appears that the approach/inhibition theory (Keltner et al., 2003) correctly predicts that high-power individuals prefer positive desirability and feasibility information, and that the social distance theory (Magee & Smith, 2013) correctly predicts that low-power individuals prefer feasibility and unfeasibility information. Thus, my studies constitute a step towards consolidating the field of power research that has been thriving for more than a decade (Galinsky et al., 2015).

Third, my results further contribute to research on power by suggesting that different processes determine the effects of high power and low power on information processing. I observed that a uni-dimensional view of power emerged in psychological research: More power was shown to go along with a higher activation, a greater tendency to act, and a stronger focus on goals (Anderson & Berdahl, 2002; Galinsky et al., 2003; Galinsky et al., 2015; Guinote, 2007; See et al., 2011; Tost, Gino, & Larrick, 2012b). This view implies that the effects of having power vs. having low power are explained best by an increase or decrease of the same process, such as the abstraction of construals as argued by the social

distance theory (Magee & Smith, 2013). In contrast to this view, my results imply that the effects of having high power and the effects of having low power are explained best by different processes. Indeed, the authors of the approach/inhibition theory (Keltner et al., 2003) proposed that independent processes produce the effects of having high-power, BAS and promotion focus, and having low power, BIS and prevention focus, but most empirical work so far focused on only one of the processes. There is much empirical support for the proposition that high power entails an activation of the BAS and a promotion focus, whereas the association of low power with an activation of the BIS and a prevention focus has received considerably less empirical attention (Hiemer & Abele, 2012). Based on my results, I concur with recent theorizing that (a) different processes produce the effects of having high power and having low power (P. K. Smith & Hofmann, 2016; Tost, 2015), and propose that (b) the concrete construals proposed by the social distance theory (Magee & Smith, 2013) are a better predictor of the effects of low power on information processing than the influence of the BIS and the prevention focus proposed by the approach/inhibition theory (Keltner et al., 2003). These conclusions are in line with previous work that found social considerations to carry greater weight on the low power side (Galinsky et al., 2015; Gruenfeld et al., 2008; Lammers & Stapel, 2009; Overbeck & Park, 2001). I extend previous theoretical and empirical work by proposing that we need to draw on different theories of power to explain the distinct effects of having high power and having low power.

Limitations and future research. Across different manipulations of power and different dependent variables, I found power to influence the evaluation and generation of arguments. Given the interesting results and certain factors that limit the ways in which one can interpret the results of the present studies, future research on the influence of power on information processing and decision-making in organizations is warranted.

First, inconsistencies in the findings in the high-power groups point to potential limitations of the generalizability of the effects of my power manipulation across samples. The findings in the high-power group of Study 1, a laboratory experiment drawing on a student sample, are different from the findings of the high-power groups in the studies with professional samples, Studies 2 and 3. It is possible that differences between laboratory and online studies account for the results, that students were unfamiliar with the type of argument evaluation, or that sample-dependent effects of the episodic recall task occurred (students usually never had power over somebody else, Tost, 2015). Future research on effects of power in organizational decision-making should thus differentiate between the effects of structural power (Sturm & Antonakis, 2015; Tost, 2015), and the effects of a sense of power, as Tost (2015) suggests.

Second, the generalizability of my results is limited by the repeated use of the same dependent variable. Although conceptual replications carry value in themselves and I demonstrated that power influences the generation of arguments as well, the present research is focused on the evaluation of arguments in organizational decision-making. I encourage future research to validate my findings in studies on other components of information processing, such as attention and recall, to get a fuller picture of the effects of power in organizational decision-making.

Third, I found that the predictions of the approach/inhibition theory (Keltner et al., 2003) regarding the influence of power on the evaluation of arguments fit for high-power individuals and the predictions of the social distance theory (Magee & Smith, 2013) regarding the influence of power on the evaluation of arguments fit for low-power individuals. While it seems reasonable to expect the same pattern to emerge in other areas of human cognition, one cannot draw conclusions about the validity of the one or the other theory from the present research. With the goal of consolidating power research further,

future research should continue to test the assumptions and predictions of different theories of power against each other.

Conclusion. The present research shows that power influences the evaluation of desirability and feasibility arguments: High-power individuals prefer positive arguments, irrespective of the desirability and feasibility domain, and low-power individuals prefer feasibility arguments, irrespective of valence. The studies contribute to a small but growing literature on the effects of power on decision-making in organizations and to research on power more generally.

CHAPTER 4

4. Digitized Performance Management and Power Concentration[5]

4.1. Introduction

Management research did not yet adequately address the changes in performance management (PM) due to digitization, even though a recent survey among members of the Society for Industrial and Organizational Psychology (SIOP) identified the changing nature of PM "as *the most important workplace trend*" (Levy et al., 2017, p. 157) and very basic notions, such as the boundaries of firms or the relation between employers and employees, appear to be in question (Colbert, Yee, & George, 2016; Phan, Wright, & Lee, 2017). In this paper, I analyze consequences of the reciprocity of power concentration and digitized performance management on multiple levels of analysis to guide the development of and research on digitized performance management. I derive nine theory-based propositions to guide research on digitized performance management and formulate principles for designing digitized performance management systems that foster flexibility, participation, and engagement (Gruber, de Leon, George, & Thompson, 2015; Gruman & Saks, 2011).

Performance management has two central functions for organizations: First, attracting, selecting, motivating, and retaining talented employees (Stone et al., 2015) and second, aligning individual performance with strategic goals of the organization (Neely et al., 2005). Traditional performance management encompasses annual reviews between managers and employees, dedicated to the discussion of future goals and past performance, often documented by means of formal ratings (Capelli & Tavis, 2016). In terms of power, traditional performance management is tailored mostly towards maintaining formal

[5] Chapter 4 is based on a working paper by Steinberg, Peus, and de Rijcke (2019).

hierarchies (Magee & Galinsky, 2008). Currently, performance management is changing (Levy et al., 2017; Phan et al., 2017; Stone & Deadrick, 2015) due to digitization (Loebbecke & Picot, 2015) and changes in the social context of performance management characterized by an increased societal focus on evaluation, as described by Dahler-Larsen (2012) in *the evaluation society*. In response to these trends, two extreme approaches to digitized performance management emerged: Algorithmic performance management (Lee, Kusbit, Metsky, & Dabbish, 2015), that is, delegating performance management to algorithms, and informal performance management, that is, dispensing with performance reviews and ratings altogether (Capelli & Tavis, 2016). Both approaches, however, increase power concentration due to two features inherent to digitization, the tendency to move towards winner-takes-all markets (Brynjolfsson & McAfee, 2014; Lanier, 2014) and the lack of algorithmic accountability (Couldry, 2016; Pasquale, 2016).

Besides having negative consequences, current approaches to digitized performance management expose organizations to the risk of missing out on realizing the full potential of digitization (Capelli & Tavis, 2016; Phan et al., 2017; Stone & Deadrick, 2015). In this paper, I argue that a relative concentration of power, rather than a relative distribution, increases the likelihood that organizations miss out on realizing the potential of digitized performance management to foster flexibility, participation, and engagement (Gruber et al., 2015; Gruman & Saks, 2011). Furthermore, I argue that, conversely, the design of digitized performance management influences the relative concentration vs. distribution of power and that active design choices with regard to transparency, accountability, and legitimacy thus determine favorable consequences of digitized performance management (Gruber et al., 2015). Uber, the digitized ride-sharing company, can count as an illustrative example (Lee et al., 2015):

Uber, a digitized company, matches drivers with cars and passengers seeking a ride. Uber uses algorithms to manage the performance of its drivers and, interestingly, of its

passengers too, since driver and passenger are bound to rate each other's performance at the end of a ride (Lee et al., 2015). If drivers log on to Uber's system they are bound to accept requests from passengers and are automatically barred from using the service if they decline several requests in a row (Lee et al., 2015). Uber has enjoyed great success, but its founder recently had to resign as CEO after allegations of sexual harassment pointed to a problematic workplace culture (Isaac, 2017b). Internal investigations related to the allegations of sexual harassment revealed an organizational culture that left managers in positions of power a relatively free hand as long as performance targets were met (Isaac, 2017a).

The present research makes several important contributions to research on performance management and research on digitization, as well as to practice, by analyzing digitized performance management from a power perspective. First, I identify power concentration as a mechanism that explains when and why digitized performance management increases the likelihood that organizations miss out on fostering flexibility, participation, and engagement (Gruber et al., 2015; Gruman & Saks, 2011). Second, I highlight the influence of technology and the social context on performance management, both of which have been identified as important and lacking theory (Levy & Williams, 2004). Third, the interdisciplinary analysis allows me to derive theory-based propositions on multiple levels of analysis. Fourth, I explain the theoretical foundations of seemingly disparate trends in digitized performance management in practice (Capelli & Tavis, 2016; Lee et al., 2015). Fifth, the present research also contributes to practice as I set forth as series of design principles for digitized performance management.

The article is structured as follows: based on a literature review I analyze the consequences of the reciprocity of power concentration and digitized performance management with regard to transparency, accountability, and legitimacy on multiple levels of analysis, that is, the organizational level, the team level, and the individual level. From my

analysis, I derive nine theory-based propositions, summarized in Table 10. The design principles can be found in the conclusion section at the end of the paper.

4.2. Theory

Performance management refers to a central practice of the human resources (HR) function that helps it achieve its organizational goals of attracting, selecting, motivating, and retaining talented employees (Stone et al., 2015). Without valid and reliable information on how individuals perform, managers lack critical input for decision-making, even beyond the scope of human resource management and supplier management (e.g. in planning durations in projects). Furthermore, performance management has profound consequences for individual employees: Hiring decisions, promotions, personnel development, and terminations all rely on performance management in one way or the other (Stone et al., 2015). For several decades, performance management encompassed one or two performance reviews per employee per year, dedicated to the discussion of past performance, including formal ratings, and to the definition of goals (Capelli & Tavis, 2016; Levy et al., 2017). This traditional approach (Levy et al., 2017) can still be considered the standard, although many companies have started to change their performance management processes (Capelli & Tavis, 2016). From a power perspective, traditional performance management is designed to keep power concentration in check by including discussions of performance goals and past performance with employees, as well as documenting these discussions and formal ratings (Capelli & Tavis, 2016). Traditional performance management is tailored mostly towards maintaining formal hierarchies, that is, legitimizing and stabilizing hierarchical differentiation (Magee & Galinsky, 2008).

In response to digitization (Loebbecke & Picot, 2015), shortages of talented employees (Capelli & Tavis, 2016), and a stronger focus on talent management (McCord, 2014), new approaches to digitized performance management emerge, among which two

extremes can be identified. On the one hand, algorithmic management (Lee et al., 2015) relies on the collection of large amounts of data and algorithms taking on managerial functions, for instance, evaluating the performance of very large numbers of workers and employees or independent contractors (Lee et al., 2015). Algorithmic performance management is in use in the low-skill sector (Chu, 2016; O'Connor, 2016) and in the high-skill sector (e.g., Fuller, 2017; Kantor & Streitfeld, 2015; Winerip, 2011). On the other hand, the digitization of performance management is one of the triggers that instigates more frequent informal performance reviews, conversations about performance, sometimes including ratings supported by or enabled by IT solutions (Capelli & Tavis, 2016; Stone et al., 2015). Some companies dispensed with formal performance management altogether (Capelli & Tavis, 2016; McCord, 2014). In this paper, I refer to this second approach as informal performance management.

In line with Loebbecke and Picot (2015) I use the term digitization to describe a larger societal trend, that is, the "changes of established patterns caused by the digital transformation and complementary innovations in our economy and society" (Loebbecke & Picot, 2015, p. 149). With regard to power and performance management, two features of digitization matter in particular: First, digitization tends to move markets towards winner-takes-all markets (Brynjolfsson & McAfee, 2014; Lanier, 2014). Second, it becomes increasingly difficult to establish accountability for what algorithms use as input and how they make decisions, a problem also referred to as lack of algorithmic accountability (Couldry, 2016; Pasquale, 2016).

I rely on a sociological concept, the evaluation society (Dahler-Larsen, 2012), to characterize the social context of digitized performance management. In *the evaluation society*, Dahler-Larsen (2012) describes how too little reflexivity in navigating the two trends towards greater complexity in societal and organizational sense-making on the one hand and

towards greater systemization in sanctioned evaluation practices on the other hand resulted in a disconnect between initial hopes that evaluation would contribute to the public interest and the realization that evaluation was effectively used to push self-serving agendas (Dahler-Larsen, 2012; Julnes, 2015). In that sense, the evaluation society is a main driver as well as a central consequence of digitization with close ties to power concentration, for evaluation is a form of exercising power that can contribute to power concentration (Townley, 1993).

Definitions of power revolve around power as the interest of social actors being realized and differ depending on the level of analysis (Magee & Galinsky, 2008; Stirling, 2008; Tost, 2015). Magee and Galinsky (2008) argue that power forms self-maintaining hierarchies, invoking a relatively stable distribution of power. In this article, I extend their argument by positing that the relative distribution of power changes over time. That is, active or passive accumulation causes power to concentrate, in that, over time, individuals end up wielding greater power and power concentrates in less individuals. The concentration of power is mediated by a multitude of factors and processes, such as, the individual desire for power, selection procedures of team leaders, and the accumulation of resources in winner-takes-all markets. Power can also be (re-)distributed, but as opposed to its tendency to concentrate, this requires a dedicated effort by social actors or external events affecting the distribution of the bases power (French Jr & Raven, 1959). I do not think of the distribution of power as a zero-sum game and note that a greater distribution of power on one level of analysis does not preclude a simultaneous increase in concentration of power on another level of analysis. For instance, on the individual level, Uber's decision to let drivers rate passengers re-distributes the power to evaluate from the customer to both parties involved in a transaction. On the organizational level, it leads to a concentration of power, since Uber gains influence over a larger number of individuals. Due to inherent qualities of digitization, winner-takes-all markets and a lack of algorithmic accountability, digitized performance

management involves a stronger power concentration than traditional forms of performance management, unless active design choices aim at increasing the distribution of power.

4.3. Propositions

Table 10

Overview over the Propositions

Level of Analysis	Proposition
Organizational	Proposition 1: Digitized performance management allows for the pursuit of multiple organizational goals and flexible organizations.
	Proposition 2: Digitized performance management reveals the organizational models that organizations operate on.
	Proposition 3: Digitized performance management contributes to defining the boundaries of organizations.
Team	Proposition 4: Digitized performance management can legitimize hierarchical differentiation and lead to steeper as well as flatter hierarchies.
	Proposition 5: Digitized performance management leads to the co-existence of multiple hierarchies.
	Proposition 6: Digitized performance management provides new opportunities for self-interested uses of social influence.
Individual	Proposition 7: Digitized performance management is subject to bias and depending on its design it may increase or decrease bias.
	Proposition 8: Digitized performance management influences employee engagement.
	Proposition 9: Digitized performance management represents the model of the human that organizations operate on.

4.3.1. Organizational Level

In this section, I discuss the reciprocity of power concentration and digitized performance management on the organizational level. In my analysis, I draw on literatures from organizational science, management science, and sociology. I derive propositions regarding (1) the link between digitized performance management, participation, and organizational goals, (2) the representation of organizational models in digitized performance

management and (3) the function of digitized performance management for the definition of organizational boundaries. Figure 8 presents the central argument on the organizational level.

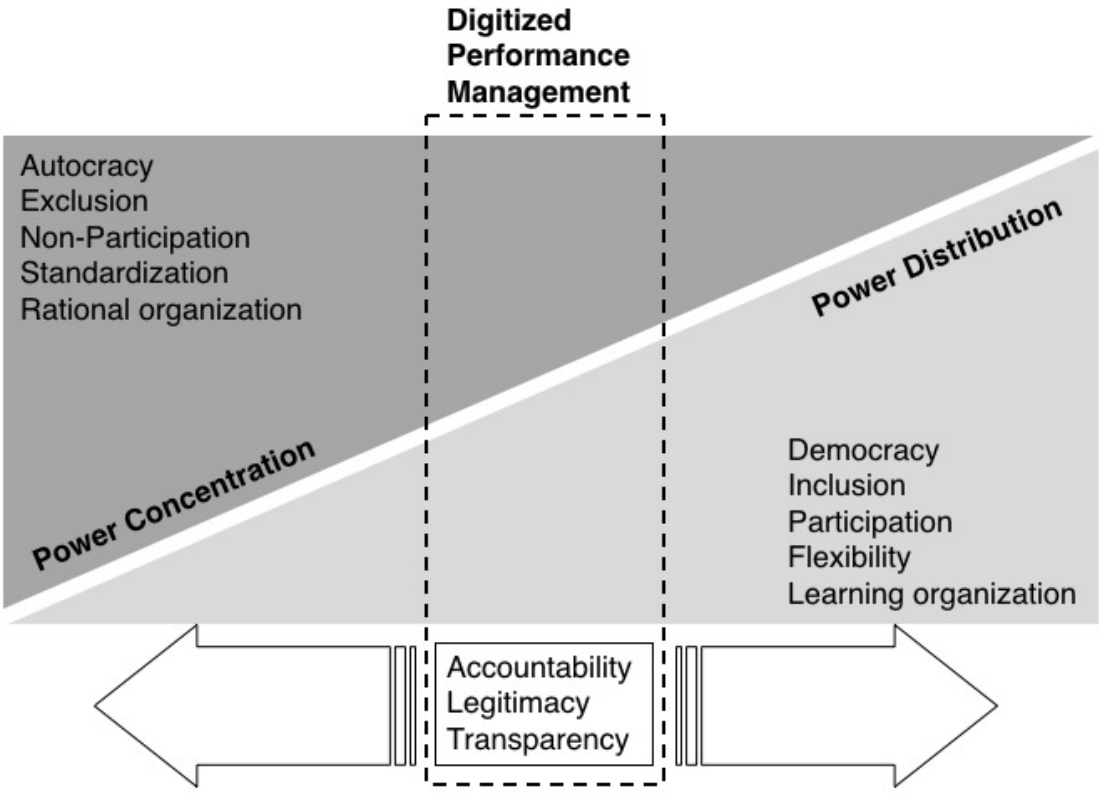

Figure 8. Organizational level consequences of the reciprocity of power concentration and digitized performance management. Realizing transparency, accountability, and legitimacy in digitized performance management leads to a shift towards power distribution and favorable consequences.

Participation in the design of digitized performance management makes organizations flexible. Digitization increases the complexity of managing performance in changing contexts, while also offering new ways to address this challenge. The difficulty lies in striking a balance between standardization, non-participation, and closing down on the one hand, as well as flexibility, participation, and opening up on the other hand (Dahler-Larsen, 2012; Julnes, 2015; Stirling, 2008, 2010). The concentration of power is intricately linked to

the question of whether the design of digitized performance management contributes to maintaining the balance between the two trends and thus can enable flexible, participative and engaging organizations to realize the benefits of digitization (Gruber et al., 2015). If these positive consequences are realized depends on how well the design of digitized performance management realizes transparency, accountability, and legitimacy.

Dahler-Larsen (2012) argues that too little reflection in the shift towards greater systematization in evaluation resulted "in a gap between idealistic hopes for evaluation as advancing the public interest and the reality of evaluation in service of self-serving agendas" (Julnes, 2015, p. 585). By analogy, I argue that too little attention being paid to transparency, accountability, and legitimacy concerns in the design of digitized performance management systems will result in a gap between hopes for digitized performance management and the reality of digitized performance management as a tool for the realization of the interests of (a small group of) power holders.

Power influences the selection and design processes of technology and the use of participation as a justification strategy (Stirling, 2008). Stirling (2008) argues that processes of commitment to new technologies can be improved, if participation is used as a way to establish transparency and accountability in the appraisal stage (identifying options) already, rather than in the commitment stage (deciding which of the pre-selected options to choose) only. Participation in the appraisal stage would open up the room for deliberation by allowing for the generation of diverse decision options (Stirling, 2008).

This reasoning can be applied to my analysis of power concentration and digitized performance management. Inviting participation in the design of digitized performance management systems is a way to distribute power by means of increasing accountability, transparency, and legitimacy. Since performance management links organizational goals to individual performance (Neely et al., 2005), this kind of participation will allow

organizations to pursue multiple goals at the same time and to stay flexible in times of digitization. From this insight, I derive the first proposition.

Proposition 1: Digitized performance management allows for the pursuit of multiple organizational goals and flexible organizations.

Digitized performance management and organizational models. Considering the role of transparency, accountability, and legitimacy reveals organizational models that underlie the design of digitized performance management. Julnes (2015) notes that greater attention has to be paid to the analysis of organizational models underlying the design of evaluation plans and the training of evaluators with the goal to apply the most appropriate models for specific purposes and contexts (Julnes, 2015). He refers to a theory challenge by Dahler-Larsen (2012), who points out that evaluation based on rather simple organizational models, as rational or learning models, is of limited use for understanding more complex organizational and social dynamics (Dahler-Larsen, 2012; Julnes, 2015). According to Julnes (2015), greater awareness of underlying organizational models could, for instance, induce evaluators to take institutional systems theory into account, identify more complex social aspects of evaluation, such as ritual and guild aspects, and thus to make better cost benefit decisions regarding evaluation (Julnes, 2015). Ultimately, greater awareness of organizational models would enable them to perform beneficial forms of evaluation rather than to just appear evaluative (Julnes, 2015).

One way to raise awareness to underlying organizational models is to consider the role of participation (Stirling, 2008). Participation and the surrounding discourse allow to infer whether participation is seen as normatively correct and necessary (process focus), whether it is used as a means of justification for the powerholder's interest (outcome focus),

or whether it is seen as substantive (outcome focus), that is, good in itself (Stirling, 2008). In essence then, Stirling (2008) makes the same distinction between being participatory and appearing participatory, that (Julnes, 2015) made between evaluating and appearing evaluative. With regard to digitized performance management this means that an analysis of participation and the awareness of transparency, accountability, and legitimacy can reveal the underlying organizational models and reveal whether organizations are actually evaluating performance or are just appearing to do so. We thus put forward our second proposition.

Proposition 2: Digitized performance management reveals the organizational models that organizations operate on.

Power relations draw organizational boundaries. Establishing transparency, accountability, and legitimacy, as well as participation, in the design of digitized performance management raises the important issue of who becomes part of the organization and what this implies in terms of power concentration. The boundaries of organizations are currently more unclear than ever (Phan et al., 2017) and digitized performance management can contribute to defining the boundaries of organizations.

Today it is customary that individuals from outside of the organization, such as customers, contribute to performance management in the form of feedback sheets, surveys or instant in-app ratings (e.g. Lee et al., 2015). Customers and recipients of services are asked about their satisfaction with products and services and this information is then used to manage employee or contractor performance (Lee et al., 2015; O'Connor, 2016). Interestingly, through digitization service providers get to rate the recipients of services as well, see Uber or Airbnb (Lee et al., 2015).

When considering this from a power perspective, one could get the impression that each of the groups has power over the other group, which effectively would make them part of the organization. A similar position can be found in the constructivist argument that evaluation is a form of assisted sense making that leads employees to construct the organization they work in (Dahler-Larsen, 2012; Julnes, 2015). In fact, I suggest all parties that take part in performance management be understood as members of a de-facto organization, for they exert power over the organization or conversely find themselves under the power of the organization.

The degree to which transparency, accountability, and legitimacy are realized in the design of digitized performance management determines the degree to which it contributes to power concentration by wider definitions of organizational boundaries. In other words, the more individuals find their performance being managed by an organization, the greater is the concentration of power in that organization. This is unless, of course, high levels of transparency, strong mechanisms of accountability, and an emphasis on legitimacy lead to a re-distribution of power in that very organization. This argument leads to the third proposition.

Proposition 3: Digitized performance management contributes to defining the boundaries of organizations.

4.3.2. Team Level

In this section, I discuss the reciprocity of power concentration and digitized performance management on the team level. I focus my analysis on two concepts that are critical to team functioning and organizational success: Social comparison and social influence. Attending to perceptions of fairness, accuracy, and purposefulness in designing

digitized performance management enables social comparison, one of the most basic processes in interactions between humans (Festinger, 1954; Suls & Wheeler, 2013). Another very important process in the interaction between humans, social influence (Cialdini & Goldstein, 2004) or power (Pratto, 2016; Raven, 1992), should be considered as well in the design of digitized performance management.

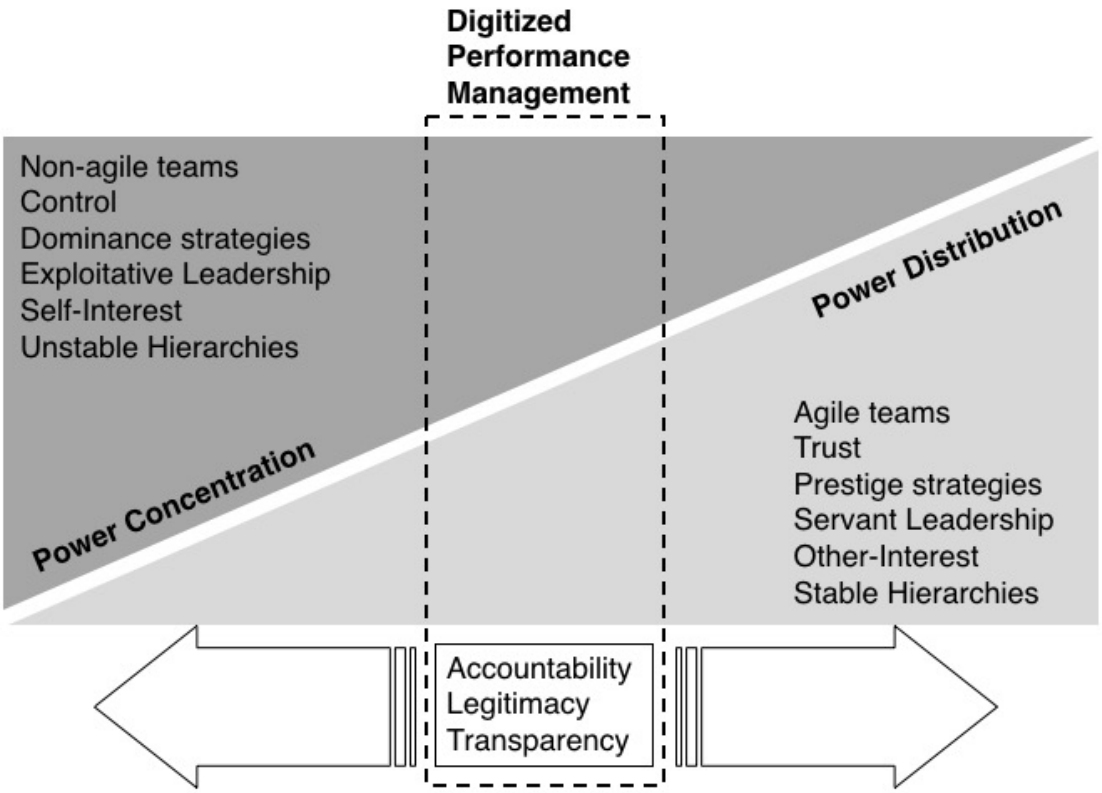

Figure 9. Team level consequences of the reciprocity of power concentration and digitized performance management. Realizing transparency, accountability, and legitimacy in digitized performance management leads to a shift towards power distribution and favorable consequences.

In my analysis, I draw on literatures from management science, organizational psychology, and social psychology. I decided to discuss social comparison and social influence on the team level rather than the organizational level of analysis, because the

research I draw on primarily focuses on the team level. I derive propositions regarding (4) the influence of digitized performance management on the justification of hierarchical differentiation, (5) hierarchy proliferation, and (6) the way social influence is exercised through digitized performance management. Figure 9 presents the central argument on the team level.

Digitized performance management and the shape of hierarchies. Hierarchical differentiation as such responds to a human need for social comparison (Festinger, 1954; Suls & Wheeler, 2013) and digitized performance management is uniquely situated to respond to that need in a highly individualized way. Magee and Galinsky (2008) state that: "Social hierarchy exists as long as there is differentiation across individuals or groups on any valued dimension." (p. 364). Applying this to digitized performance management, I note that by accepting discrete inputs only (Loebbecke & Picot, 2015), digital technologies inherently introduce differentiation. Since performance management purposefully differentiates on valued dimensions, digitized performance management contributes to the maintenance (Magee & Galinsky, 2008) and to the proliferation of hierarchies. I use the term hierarchy proliferation to describe that digital technologies create many hierarchies by differentiating on many valued dimensions, potentially based on any digitally represented information.

With regard to maintenance, I argue that digitized performance management is not only maintaining hierarchies, as traditional performance management did, but actively changes their form due to concentration of power. Digitized performance management can lead to flatter hierarchies when the design establishes transparency, accountability, and legitimacy with regard to hierarchical differentiation, and to steeper hierarchies due to power concentration if transparency, accountability, and legitimacy are not present. Under the condition that performance management is perceived as fair, accurate and purposeful (Iqbal, Akbar, & Budhwar, 2015) the steepening and flattening of hierarchies in digitized

performance management can be legitimized. Legitimizing hierarchical differentiation (Lammers et al., 2008) is an important means to avoid unstable hierarchies. Unstable hierarchies harm teams and organizations due to the ensuing fights that arise when team members try to improve their position in the hierarchy (Case & Maner, 2014; Magee & Galinsky, 2008; Maner & Case, 2016).

An early empirical study that analyzed employee reactions to algorithmic performance management (Lee et al., 2015) points to problems with ratee perceptions of current approaches to digitized performance management. The lack of transparency of Uber's and Lyft's, two ride-sharing companies, algorithms caused negative feelings and questions about the purposefulness of the technology (Lee et al., 2015). The algorithmic performance evaluation was seen as unfair and ineffective (Lee et al., 2015). In the long run, these perceptions could lead to protests and maybe unionization, thereby flattening hierarchies, but in a costly way to both parties involved.

Proposition 4: Digitized performance management can legitimize hierarchical differentiation and lead to steeper as well as flatter hierarchies.

Digitized performance management and the proliferation of hierarchies. With regard to hierarchy proliferation, I argue that in digitized performance management any digitally represented information can become the base of hierarchical differentiation. This could lead to more complex representations of performance or digitized performance management based on a hierarchy other than the performance hierarchy, which can become a serious problem for an organization. Think of patients rating their satisfaction with their treatment in hospital (Robbins, 2015). If satisfaction becomes a part of performance management and doctors adapt their behavior this can easily become a quality problem with

fatal consequences. The opinions of lay persons could gain weight in a domain where expert judgement is required, in short, patients need to be healthy not satisfied (Robbins, 2015). This example illustrates consequences of a case where performance management is based on a hierarchy other than the performance hierarchy.

In companies that changed from the traditional to the informal approach to performance management and suspended performance ratings, soon other proxies were found and used by the employees to compare themselves (Capelli & Tavis, 2016). Salary, for instance, although determined by many factors unrelated to performance was used to infer performance by employees (Capelli & Tavis, 2016). In this case as well, performance management was based on a hierarchy other than the performance hierarchy and may hinder the best to rise to the top. In this context it is interesting to consider the role of social media that adds yet another layer of hierarchies to the organizations of today. In the last elections for the German parliament, not only accredited journalists got to interview the candidates running to be German chancellor, but also a handful of teenagers with a huge followership on social media channels (Schmitt & Meinberger, 2017). Transparency, accountability, and legitimacy in digitized performance management are ultimately going to decide whether hierarchy proliferation leads to a concentration of power or to a distribution of power due to the reasonable broadening of the performance concept.

Proposition 5: Digitized performance management leads to the co-existence of multiple hierarchies.

Digitized performance management and social influence. Depending on power concentration, digitized performance management can reinforce the exercise of self-serving uses of social influence by managers and leaders. With regard to social influence, traditional

performance management aimed at reducing self-serving behavior on part of the leader, for example, by discussing performance goals and ratings with employees (Capelli & Tavis, 2016) or by taking different sources of data into account. When thoroughly implemented, these measures established certain levels of transparency, accountability, and legitimacy and thus kept power concentration in check. The same mechanisms need to be realized in digitized performance management to constrain self-serving uses of social influence.

Research on social influence documented self-serving leader behavior, for instance in the literature on power (Maner & Case, 2016) and on leadership (Schmid, Pircher Verdorfer, & Peus, 2017). Both fields of research distinguish between self-serving and other-serving ways to use social influence (Maner & Case, 2016; Schmid et al., 2017). In the literature on power motivation, researchers distinguish between personalized and socialized power motivation (Magee & Langner, 2008), or dominance and prestige motivation (Maner & Case, 2016) to highlight the expectation of self-serving or other-serving outcomes of the exercise of power. Empirical work corroborates the notion of favorable outcomes of socialized power motivation (Magee & Langner, 2008), prestige motivation (Case & Maner, 2014; Maner & Mead, 2010), or a communally oriented use of power (Chen, Lee-Chai, & Bargh, 2001), as well as the notion of unfavorable outcomes of personalized power motivation (Magee & Langner, 2008), dominance motivation (Case & Maner, 2014; Maner & Mead, 2010), or an exchange-oriented use of power (Chen et al., 2001). Leadership research distinguishes between leadership, defined as a process of influencing others (Yukl, 2010) with the expectation of favorable outcomes (Schyns & Schilling, 2013), and destructive leadership (Schyns & Schilling, 2013), such as, when leaders exploit their followers to their own benefit (Schmid et al., 2017).

In digitized performance management, a lack of transparency, accountability, and legitimacy can encourage self-serving behavior. In particular process accountability, as

opposed to outcome accountability, has been found to moderate self-serving leader behaviors (Pitesa & Thau, 2013; Rus et al., 2012). Digital technologies that enforce transparency will grant a certain level of protection against exploitation and also contribute to perceptions of legitimacy. On the other hand, it will be important for future research to observe what types of expressions different leadership concepts, such as exploitative leadership (Schmid et al., 2017), and basic motivations, such as power motivation, will find in digitized performance management or more generally in digitization. It is conceivable that digitization provides new means of exercising self-serving, dominant, and exploitative behavior by leaders, for instance, when the speed of food deliveries is compared to an algorithmic value that is not specified further (O'Connor, 2016) or the manager gets feedback on his employees anonymously provided by their colleagues (Kantor & Streitfeld, 2015).

In designing performance management systems, it seems advisable to opt for transparency and to build in mechanisms that foster process accountability (Pitesa & Thau, 2013). Indeed, I argue that both, transparency and process accountability, not only attenuate unfavorable effects that may arise from processes of social comparison and social influence between individuals, but also act against individual biases and to reinforce engagement.

Proposition 6: Digitized performance management provides new opportunities for self-interested uses of social influence.

4.3.3. Individual Level

In this section, I discuss the reciprocity of power concentration and digitized performance management on the individual level. I focus my analysis on biases, engagement, and the model of the human that organizations operate on. In this analysis, I draw on literatures from management science as well as organizational psychology and sociology. I

derive propositions regarding (7) the relation of digitized performance management and cognitive bias, (8) the influence of digitized performance management on employee engagement as well as (9) the representation of the model of the human in digitized performance management. Figure 10 presents the central argument on the individual level.

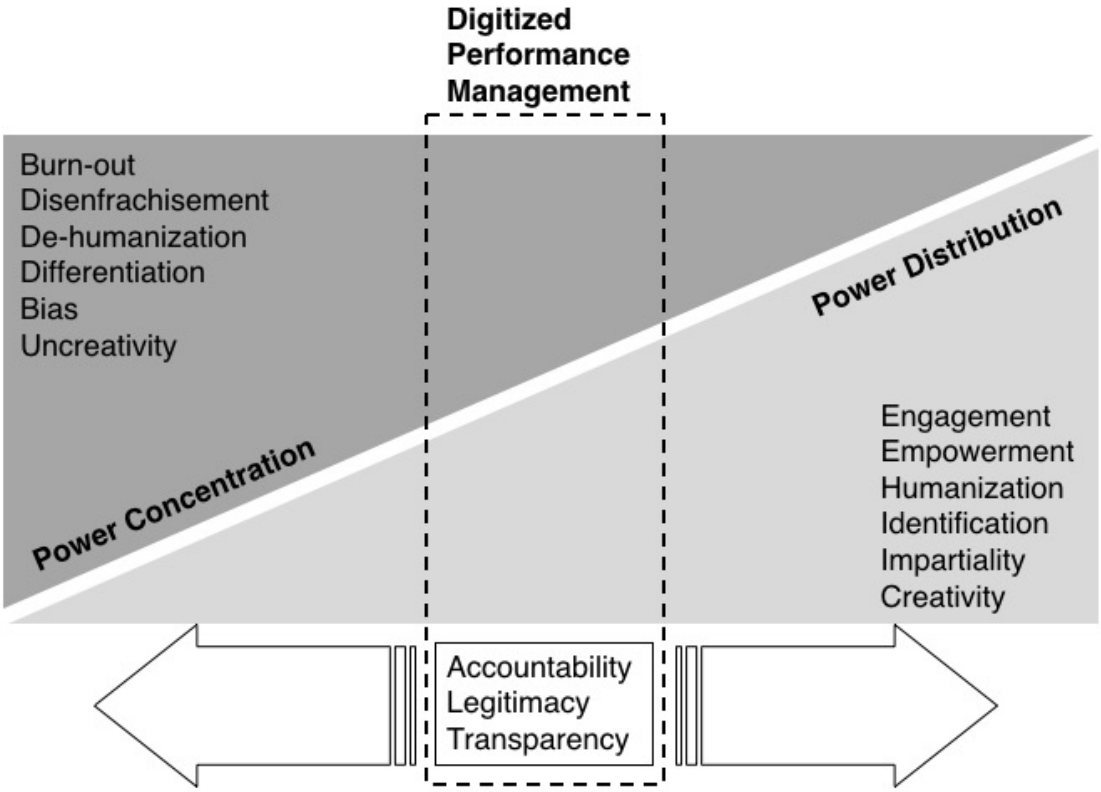

Figure 10. Individual level consequences of the reciprocity of power concentration and digitized performance management. Realizing transparency, accountability, and legitimacy in digitized performance management leads to a shift towards power distribution and favorable consequences.

Digitized performance management and cognitive bias. Digitized performance management, just as traditional forms of performance management, is subject to cognitive bias, even though it may appear to be more objective and actually increase rather than reduce bias (towards, e.g., gender, ethnicity, etc.). Traditional performance management has been

designed to increase standardization, reduce bias, and avoid errors, but a large number of empirical studies attests to its failure in this domain (for an overview, see Iqbal et al., 2015). In terms of consequences, biases and errors in performance management lead to incorrect results and have the potential to lower employee satisfaction and organizational commitment (Iqbal et al., 2015). Digitized performance management often appears more objective in this respect, even when it is not, and may reinforce bias as well as power concentration.

From the outside, it seems that algorithms are just combining data according to rules and thus will produce more objective, unbiased results than traditional, formal performance management (Bilić, 2016; Couldry, 2016; Pasquale, 2016). This would be hard to judge, however, because in order to identify bias in algorithmic performance management, the creation of algorithms and the human involvement in this process (Bilić, 2016; Couldry, 2016; Pasquale, 2016), need to be taken into account. At this stage, political agendas, opinions, individual cognitive biases, and heuristics may all take effect and influence downstream performance appraisals. If no information on the programming and training of algorithms is disclosed or if this information remains proprietary, that is, if no algorithmic accountability is established (Couldry, 2016), it will be very difficult to identify biases and errors in performance ratings.

In designing digitized performance management systems, it will be essential to establish accountability, transparency, and legitimacy to reduce the impact of errors and attenuate bias, as I have argued in relation to social influence.

Proposition 7: Digitized performance management is subject to bias and depending on its design it may increase or decrease bias.

Digitized performance management influences employee engagement. Digitized performance management can be designed to positively influence employee engagement and the actual experience of work (Gruber et al., 2015). Employee engagement is assumed to be an important mediator of the relationship of performance management and actual increases in performance (Gruman & Saks, 2011). In their model, Gruman and Saks (2011) assume levels of employee engagement to change depending on how much the elements of a performance management system are designed to foster engagement. Thus, in digitized performance management, designs that lead to a relative distribution of power by attending to transparency, accountability, and legitimacy concerns will be more likely to foster engagement than designs that reinforce power concentration.

In their study on algorithmic management of individual contractors, Lee et al. (2015) identified algorithmic performance management as a source of negative affect and perceptions of unfairness, both of which would negatively impact on engagement (Gruman & Saks, 2011). Informal performance management can have positive and negative effects on employee engagement, depending, for example, on similarity (Avery, McKay, & Wilson, 2007) or the correspondence of moral perceptions between supervisor and subordinate (Fehr, Yam, & Dang, 2014). From these considerations I derive a proposition regarding employee engagement.

Proposition 8: Digitized performance management influences employee engagement.

Digitized performance management and the model of the human being. An analysis of how transparency, accountability, and legitimacy are realized in digitized performance management can reveal what models and beliefs organizations have of the human being and how this is linked to power concentration. This argument reflects my argumentation from the

organizational level, that digitized performance management reveals organizational models, on the individual level. It is long known that traditional performance management is an expression of the model and beliefs of the human (MacGregor, 1960). MacGregor (1960), for instance, distinguished between a management approach that emphasized control to counterbalance the purported unwillingness of employees, referred to as Theory X, and a management approach aimed at contexts that elicit and align the employees' willingness to work towards organizational goals, referred to as Theory Y (MacGregor, 1960). I argue that power concentration determines whether individuals and organization are more inclined to embrace Theory X or Theory Y (MacGregor, 1960). That is, power concentration leads individuals and organizations to progressively embrace beliefs of the human as a lazy, unwilling, change-resistant, and self-centered being that needs managerial control to perform at all.

In digitized performance management designs, realizing transparency, accountability, and legitimacy allows individuals and organizations to move towards a Theory Y conception of the human (MacGregor, 1960), as an active, motivated, and self-directed being that needs managers mostly to create contexts enabling performance. Considering the model of the human that organizations have is important because it may influence the perception of the organization as an ethical organization (Peus, Kerschreiter, Traut-Mattausch, & Frey, 2010), justice perceptions, and employee engagement (Gruman & Saks, 2011). Currently, we see a disconnect in the models of the human that approaches to digitized performance management in the low-skill and the high-skill sector seem to be based on. Algorithmic management clearly emphasizes a Theory X model of humans in food delivery services (O'Connor, 2016), transportation (Lee et al., 2015), and warehouses (Chu, 2016) to name a few. The informal approach to performance management, on the other hand, emphasizes a Theory Y model of the human notably in technology companies (McCord, 2014), the financial sector, and

consultancies (Capelli & Tavis, 2016). To be sure, digitized performance management offers ways to enforce Theory X based control in high-skilled jobs as well (Fuller, 2017; Kantor & Streitfeld, 2015). In my view, however, organizations should not go down that road in either the high-skill or the low-skill sector and instead design digitized performance management to allow highly individualized, self-directed, and engaging assessments of performance that direct a highly capable, motivated, and self-organized workforce in flexible organizations. This would make organizations not only act in ethically responsible ways, but will reinforce organizational success (Peus et al., 2010) in times of digitization – consider that Netflix' stock performed better even than Amazon's over ten years' time (Carter, 2018). Netflix relies on informal performance management (McCord, 2014) and Amazon relies on algorithmic performance management in high-skilled and low-skilled jobs (Kantor & Streitfeld, 2015). Thus, whether organizations will be able to leverage the full potential of digitized performance management, depends on how they deal with power concentration and how they realize transparency, accountability, and legitimacy in the design of digitized performance management.

Proposition 9: Digitized performance management represents the model of the human that organizations operate on.

4.4. Conclusion

In this article, I analyzed the reciprocity between power concentration and digitized performance management. I focused on the influence of transparency, accountability, and legitimacy on this relationship and the consequences of digitized performance management on multiple levels of analysis. From my analysis I derived nine propositions to guide the development of and research on digitized performance management. In order to guide the

development of and research on digitized performance management, where little theory is available yet (Levy et al., 2017; Phan et al., 2017), I decided to draw on diverse literatures and to privilege breadth over depth. Performance management has always been an interdisciplinary endeavor (Neely, 2005) and I aim to stimulate research in diverse fields that addresses digitized performance management from different angles. Analyzing the reciprocal relationship of power concentration and digitized performance management allowed me to contribute to research and practice in several ways. First, I identify power concentration as a mechanism that explains when and why digitized performance management increases the likelihood that organizations miss out on fostering flexibility, participation, and engagement (Gruber et al., 2015; Gruman & Saks, 2011). Second, I highlight the influence of technology and the social context on performance management, both of which have been identified as important and lacking theory (Levy & Williams, 2004). Third, my interdisciplinary analysis allows me to derive theory-based propositions on multiple levels of analysis. Fourth, I explain the theoretical foundations of seemingly disparate trends in digitized performance management in practice. Fifth, the present research contributes to practice as I set forth a series of design principles for digitized performance management.

Through my analysis, it became clear that organizations need to design digitized performance management with transparency, accountability, and legitimacy concerns in mind (Gruber et al., 2015) in order to be successful in digitization. Not only does power concentration have negative consequences, but it also increases the likelihood to miss out on realizing the full potential of digitization. With properly designed digitized performance management, organizations will be able to (re-)distribute power, to stay flexible, to allow individuals to grow based on performance appraisals, and to offer engaging workplaces to an active and motivated workforce (Gruber et al., 2015). Current approaches to digitized performance management, however, appear not to fare well with regard to these goals.

Neither current approaches to algorithmic performance management, nor current approaches to informal performance management seem to implement transparency, accountability, and legitimacy in ways that would allow them to fully realize the benefits of digitization. Algorithmic management emphasizes a Theory X (MacGregor, 1960) based model of the human as lazy, unwilling, and in need of managerial control, as well as a rational organizational model that does not capture the complexities of human interaction in social contexts (Dahler-Larsen, 2012). A lack of transparency and accountability (Couldry, 2016; Pasquale, 2016) makes it impossible to say what kind of information algorithms use to differentiate employees in hierarchies (Lee et al., 2015) and allows for the self-serving use of social influence by power holders. For the same reason, it would be difficult to detect errors and bias in these algorithms. Taken together, these factors contribute to perceptions of illegitimacy and negatively impact on engagement (Gruman & Saks, 2011). Informal performance management has similar consequences, albeit for different reasons. Informal performance management emphasizes a Theory Y (MacGregor, 1960) model of the human as active, willing, and self-directed, as well as allowing for more complex organizational models, such as the learning model (Dahler-Larsen, 2012). With regard to accountability, transparency, and the possibilities to identify biases, however, informal performance management de-emphasizes control too much. Leaving it all to the manager is likely to reinforce the influence of bias (Avery et al., 2007) and thus to negatively impact on legitimacy and engagement, where properly designed digitized performance management would be able to establish transparency, accountability, and legitimacy and thus to foster engagement (Gruman & Saks, 2011).

I conclude my analysis outlining a few principles for the design of digitized performance management based on our theoretical propositions:

a. To achieve flexibility and the pursuit of multiple organizational goals, encourage participation in the project definition and selection stages of the implementation of digital technologies in performance management.
b. Identify and clearly communicate the organizational model you operate on to allow for discussions and the participatory development of appropriate digitized performance management systems.
c. Be clear on who belongs to your organization for taking part in performance management and how including each party contributes to reaching your organization's goals.
d. Make transparent what aspects are considered in (algorithmic) performance management to allow individuals in your organization to compare themselves to others and direct themselves towards better performance.
e. Account for the fact that in digitization multiple hierarchies outside of your organization, for instance, the size of one's network on social media, play a role in determining contextual and sometimes even task performance.
f. Establish transparency and accountability in order to constrain self-serving uses of social influence, unbiased appraisals of performance, and perceptions of legitimacy that foster engagement.
g. Encourage participation in the design of all aspects of digitized performance management to (a) establish transparency, accountability, and legitimacy and (b) to foster engagement.
h. Critically reflect upon and adapt the model of the human that underlies your design decisions and, conversely, can be inferred from your approach to digitized performance management.

Respecting these principles in the design of digitized performance management increases the likelihood that organizations realize the full the potential of digitization to foster flexibility, participation, inclusion, other-serving uses of power, and engagement.

REFERENCES

Aime, F., Humphrey, S., Derue, D. S., & Paul, J. B. (2014). The riddle of heterarchy: Power transitions in cross-functional teams. *Academy of Management Journal, 57*(2), 327-352. doi:10.5465/amj.2011.0756

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science, 10*(4), 327-333.

Anderson, C., & Berdahl, J. L. (2002). The experience of power: Examining the effects of power on approach and inhibition tendencies. *Journal of Personality and Social Psychology, 83*(6), 1362-1377. doi:10.1037/0022-3514.83.6.1362

Anderson, C., & Brion, S. (2014). Perspectives on Power in Organizations. *Annual Review of Organizational Psychology and Organizational Behavior, 1*(1), 67-97. doi:10.1146/annurev-orgpsych-031413-091259

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. *European Journal of Social Psychology, 36*(4), 511-536. doi:10.1002/ejsp.324

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

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5-21.

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

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596-612. doi:10.1037/0022-3514.63.4.596

Avery, D. R., McKay, P. F., & Wilson, D. C. (2007). Engaging the aging workforce: the relationship between perceived age similarity, satisfaction with coworkers, and employee engagement. *Journal of Applied Psychology, 92*(6), 1542.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language, 59*(4), 390-412.

Banchefsky, S., Westfall, J., Park, B., & Judd, C. M. (2016). But you don't look like a scientist!: Women scientists with feminine appearance are deemed less likely to be scientists. *Sex Roles, 75*(3-4), 95-109.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language, 68*(3), 255-278. doi:http://dx.doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *2015, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology, 5*(4), 323.

Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. *The Leadership Quarterly, 26*(2), 101-122. doi:https://doi.org/10.1016/j.leaqua.2014.07.010

Benzion, U., Cohen, Y., Peled, R., & Shavit, T. (2008). Decision-making and the newsvendor problem: an experimental study. *Journal of the Operational Research Society, 59*(9), 1281-1287.

Berson, Y., Halevy, N., Shamir, B., & Erez, M. (2015). Leading from different psychological distances: A construal-level perspective on vision communication, goal setting, and

follower motivation. *The Leadership Quarterly, 26*(2), 143-155. doi:http://dx.doi.org/10.1016/j.leaqua.2014.07.011
Bilić, P. (2016). Search algorithms, hidden labour and information control. *Big Data & Society, 3*(1), 1-9. doi:10.1177/2053951716652159
Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben [Presentation and validation of a German version of the Need for Cognition Scale]. *Zeitschrift für Sozialpsychologie*(25), 147-154.
Bolte, A., Goschke, T., & Kuhl, J. (2003). Emotion and intuition. *Psychological Science (Wiley-Blackwell), 14*(5), 416. doi:10.1111/1467-9280.01456
Bolton, G. E., & Katok, E. (2008). Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing & Service Operations Management, 10*(3), 519-538.
Bowers, K. S., Regehr, G., Balthazard, C., & Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology, 22*(1), 72-110. doi:http://dx.doi.org/10.1016/0010-0285(90)90004-N
Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York, N.Y.: W.W. Norton & Company, Inc.
Bugental, D. B., & Lewis, J. C. (1999). The Paradoxical Misuse of Power by Those Who See Themselves as Powerless: How Does it Happen? *Journal of Social Issues, 55*(1), 51-64.
Capelli, P., & Tavis, A. (2016). The Future of Performance Reviews. Retrieved from https://hbr.org/2016/10/the-performance-management-revolution
Carter, S. M. (2018). Amazon, Apple and Google stock rose over the last 10 years, but this company outperformed them all. Retrieved from https://www.cnbc.com/2018/01/26/netflix-stock-performed-better-than-amazon-apple-and-google.html
Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319.
Case, C. R., & Maner, J. K. (2014). Divide and conquer: When and why leaders undermine the cohesive fabric of their group. *Journal of Personality and Social Psychology, 107*(6), 1033.
Chen, S., Lee-Chai, A. Y., & Bargh, J. A. (2001). Relationship orientation as a moderator of the effects of social power. *Journal of Personality and Social Psychology, 80*(2), 173-187. doi:10.1037/0022-3514.80.2.173
Chu, B. (2016). Nobel prize winners' research worked out a theory on worker productivity – then Amazon and Deliveroo proved it wrong. Retrieved from http://www.independent.co.uk/voices/amazon-deliveroo-workers-nobel-economics-prize-winners-workers-productivity-measures-uber-eats-a7356201.html
Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annu Rev Psychol, 55*(1), 591-621. doi:10.1146/annurev.psych.55.090902.142015
Colbert, A., Yee, N., & George, G. (2016). The Digital Workforce and the Workplace of the Future, Editorial. *Academy of Management Journal,* pp. 731-739. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=116369835&site=ehost-live
Couldry, N. (2016). What's at Stake in Algorithmic Accountability. *Media Policy Blog.* Retrieved from http://blogs.lse.ac.uk/mediapolicyproject/2016/04/28/whats-at-stake-in-algorithmic-accountability/

Crusius, J., van Horen, F., & Mussweiler, T. (2012). Why process matters: A social cognition perspective on economic behavior. *Journal of Economic Psychology, 33*(3), 677-685.
Dahler-Larsen, P. (2012). *The Evaluation Society*. Stanford, California: Stanford University Press.
Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods, 12*(2), 121.
Executive Summary of Final Class Settlement Program (2.0 Liter Engine Vehicles). (2016). Retrieved from http://www.cand.uscourts.gov/crb/vwmdl/final-settlement-2-Liter
Fast, N. J., Sivanathan, N., Mayer, N. D., & Galinsky, A. D. (2012). Power and overconfident decision-making. *Organizational Behavior and Human Decision Processes, 117*(2), 249-260. doi:10.1016/j.obhdp.2011.11.009
Fehr, R., Yam, K. C., & Dang, C. (2014). Moralized Leadership: The Construction and Consequences of Ethical Leader Perceptions. *Academy of Management Review, 40*(2), 182.
Fellner, B., Holler, M., Kirchler, E., & Schabmann, A. (2007). Regulatory focus scale (RFS): Development of a scale to record dispositional regulatory focus. *Swiss Journal of Psychology, 66*(2), 109-116.
Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117-140.
Fiedler, K. (2007). Construal level theory as an integrative framework for behavioral decision-making research and consumer psychology. *Journal of Consumer Psychology, 17*(2), 101-106.
Fiedler, K., Jung, J., Wänke, M., Alexopoulos, T., & de Molière, L. (2015). Toward a deeper understanding of the ecological origins of distance construal. *Journal of Experimental Social Psychology, 57*, 78-86.
Fischer, J., Fischer, P., Englich, B., Aydin, N., & Frey, D. (2011). Empower my decisions: The effects of power gestures on confirmatory information processing. *Journal of Experimental Social Psychology, 47*(6), 1146-1154. doi:10.1016/j.jesp.2011.06.008
Fleming, P., & Spicer, A. (2014). Power in Management and Organization Science. *The academy of management annals, 8*(1), 237-298. doi:10.1080/19416520.2014.875671
Flynn, F. J., Gruenfeld, D. H., Molm, L. D., & Polzer, J. T. (2011). Social Psychological Perspectives on Power in Organizations. *Administrative science quarterly, 56*(4), 495-500. doi:10.1177/0001839212440969
Frees, E. W. (2004). *Longitudinal and panel data: analysis and applications in the social sciences*. Cambridge, UK: Cambridge University Press.
French Jr, J. R., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power.* (pp. 150-167). Oxford, England: Univer. Michigan, ix, 225 pp.
Fügener, A., Schiffels, S., & Kolisch, R. (2015). Overutilization and underutilization of operating rooms - insights from behavioral health care operations management. *Health Care Management Science*, 1-14. doi:10.1007/s10729-015-9343-1
Fujita, K., Eyal, T., Chaiken, S., Trope, Y., & Liberman, N. (2008). Influencing attitudes toward near and distant objects. *Journal of Experimental Social Psychology, 44*(3), 562-572.
Fuller, R. (2017). Transform your organization with Microsoft Workplace Analytics. Retrieved from https://blogs.office.com/en-us/2017/07/05/transform-your-organization-with-microsoft-workplace-analytics/?eu=true
Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology, 85*(3), 453-466. doi:10.1037/0022-3514.85.3.453

Galinsky, A. D., Rucker, D. D., & Magee, J. C. (2015). Power: Past findings, present considerations, and future directions. In M. Mikulincer, P. R. Shaver, J. A. Simpson, & J. F. Dovidio (Eds.), *APA handbook of personality and social psychology, Vol. 3: Interpersonal relations. APA handbooks in psychology.* (pp. 421-460). Washington, DC: American Psychological Association.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association, 1*(1), 114-125.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics, Vol. 3, Speech Acts* (pp. 41-58). New York: Academic Press.

Gruber, M., de Leon, N., George, G., & Thompson, P. (2015). Managing by Design. *Academy of Management Journal, 58*(1), 1-7. doi:10.5465/amj.2015.4001

Gruenfeld, D. H., Inesi, M. E., Magee, J. C., & Galinsky, A. D. (2008). Power and the objectification of social targets. *Journal of Personality and Social Psychology, 95*(1), 111-127. doi:10.1037/0022-3514.95.1.111

Gruman, J. A., & Saks, A. M. (2011). Performance management and employee engagement. *Human Resource Management Review, 21*(2), 123-136. doi:https://doi.org/10.1016/j.hrmr.2010.09.004

Guinote, A. (2007). Power and goal pursuit. *Personality and Social Psychology Bulletin, 33*(8), 1076-1087.

Halkjelsvik, T., & Jørgensen, M. (2012). From origami to software development: A review of studies on judgment-based predictions of performance time. *Psychological Bulletin, 138*(2), 238-271. doi:10.1037/a0025996

Heller, S., Borsay, F., & Ullrich, J. (2017). Social power and dimensions of self-control: Does power benefit initiatory self-control but impair inhibitory self-control? *Cogent Psychology, 4*(1), 1-16. doi:10.1080/23311908.2017.1288351

Hiemer, J., & Abele, A. E. (2012). High power = Motivation? Low power = Situation? The impact of power, power stability and power motivation on risk-taking. *Personality and Individual Differences, 53*(4), 486-490. doi:http://dx.doi.org/10.1016/j.paid.2012.04.008

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

Hobday, M. (2000). The project-based organisation: an ideal form for managing complex products and systems? *Research policy, 29*(7-8), 871-893.

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology, 27*(1), 99-114.

Inesi, M. E. (2010). Power and loss aversion. *Organizational Behavior and Human Decision Processes, 112*(1), 58-69.

Iqbal, M. Z., Akbar, S., & Budhwar, P. (2015). Effectiveness of Performance Appraisal: An Integrated Framework. *International Journal of Management Reviews, 17*(4), 510-533. doi:10.1111/ijmr.12050

Isaac, M. (2017a, 22 February). Inside Uber's Aggressive, Unrestrained Workplace Culture. *New York Times*. Retrieved from https://www.nytimes.com/2017/02/22/technology/uber-workplace-culture.html?action=click&contentCollection=Technology&module=RelatedCoverage%C2%AEion=Marginalia&pgtype=article

Isaac, M. (2017b). Uber Founder Travis Kalanick Resigns as C.E.O. *New York Times*. Retrieved from https://www.nytimes.com/2017/06/21/technology/uber-ceo-travis-kalanick.html

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: a new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*(1), 54-69.

Julnes, G. (2015). Managing Evaluation Theories, Practices, and Communities:Reflections on Dahler-Larsen's The Evaluation Society. *American Journal of Evaluation, 36*(4), 584-600. doi:10.1177/1098214015599636

Kantor, J., & Streitfeld, D. (2015, 15 August). Inside Amazon: Wrestling Big Ideas in a Bruising Workplace. *New York Times*. Retrieved from https://www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html

Katsikopoulos, K. V., & Gigerenzer, G. (2013). Behavioral operations management: A blind spot and a research program. *Journal of Supply Chain Management, 49*(1), 3-7.

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

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review, 2*(3), 196-217.

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der" Positive and Negative Affect Schedule"(PANAS). *Diagnostica*(42), 139-156.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *2017, 82*(13), 26. doi:10.18637/jss.v082.i13

Lammers, J., Dubois, D., Rucker, D. D., & Galinsky, A. D. (2013). Power gets the job: Priming power improves interview outcomes. *Journal of Experimental Social Psychology, 49*(4), 776-779. doi:10.1016/j.jesp.2013.02.008

Lammers, J., Dubois, D., Rucker, D. D., & Galinsky, A. D. (2017). Ease of retrieval moderates the effects of power: Implications for the replicability of power recall effects. *Social cognition, 35*(1), 1-17.

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2008). Illegitimacy moderates the effects of power on approach. *Psychological Science, 19*(6), 558-564. doi:10.1111/j.1467-9280.2008.02123.x

Lammers, J., & Stapel, D. A. (2009). How power influences moral thinking. *Journal of Personality and Social Psychology, 97*(2), 279.

Lammers, J., Stoker, J. I., Rink, F., & Galinsky, A. D. (2016). To Have Control Over or to Be Free From Others? The Desire for Power Reflects a Need for Autonomy. *Personality and Social Psychology Bulletin, 42*(4), 498-512. doi:10.1177/0146167216634064

Lanier, J. (2014). *Who owns the future?* London, England: Penguin Group Ltd.

Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). *Working with machines: The impact of algorithmic and data-driven management on human workers.* Paper presented at the Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea.

Levy, P. E., Tseng, S. T., Rosen, C. C., & Lueke, S. B. (2017). Performance Management: A Marriage between Practice and Science – Just Say "I do" *Research in personnel and human resources management* (pp. 155-213). Published online 25 Jul 2017: Emerald Publishing Limited.

Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management, 30*(6), 881-905.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75*(1), 5-18. doi:10.1037/0022-3514.75.1.5

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems, 24*(3), 149-157. doi:https://doi.org/10.1016/j.jsis.2015.08.002

Lossie, H., & Hadem, M. (2016). Warum Volkswagen den Schuldigen nie finden wird. https://www.welt.de/wirtschaft/article154412154/Warum-Volkswagen-den-Schuldigen-nie-finden-wird.html.

Lu, J., Xie, X., & Xu, J. (2013). Desirability or Feasibility: Self–Other Decision-Making Differences. *Personality and Social Psychology Bulletin, 39*(2), 144-155. doi:10.1177/0146167212470146

Macenczak, L. A., Campbell, S., Henley, A. B., & Campbell, W. K. (2016). Direct and interactive effects of narcissism and power on overconfidence. *Personality and Individual Differences, 91*, 113-122. doi:http://dx.doi.org/10.1016/j.paid.2015.11.053

MacGregor, D. (1960). *The human side of enterprise*. New York: McGraw-Hill.

Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: the self-reinforcing nature of power and status. *The academy of management annals, 2*(1), 351-398.

Magee, J. C., & Langner, C. A. (2008). How personalized and socialized power motivation facilitate antisocial and prosocial decision-making. *Journal of Research in Personality, 42*(6), 1547-1559.

Magee, J. C., & Smith, P. K. (2013). The Social Distance Theory of Power. *Personality and Social Psychology Review, 17*(2), 158-186. doi:10.1177/1088868312472732

Maner, J. K., & Case, C. R. (2016). Chapter Three - Dominance and Prestige: Dual Strategies for Navigating Social Hierarchies. In M. O. James & P. Z. Mark (Eds.), *Advances in experimental social psychology* (Vol. Volume 54, pp. 129-180): Academic Press.

Maner, J. K., & Mead, N. L. (2010). The essential tension between leadership and power: When leaders sacrifice group goals for the sake of self-interest. *Journal of Personality and Social Psychology, 99*(3), 482-497. doi:10.1037/a0018559

McCord, P. (2014). How Netflix reinvented HR. Retrieved from https://hbr.org/2014/01/how-netflix-reinvented-hr

McShane, B. B., & Böckenholt, U. (2017). Single-Paper Meta-Analysis: Benefits for Study Summary, Theory Testing, and Replicability. *Journal of Consumer Research, 43*(6), 1048-1063. doi:10.1093/jcr/ucw085

Moon, A., & Chen, S. (2014). The power to control time: Power influences how much time (you think) you have. *Journal of Experimental Social Psychology, 54*, 97-101. doi:http://dx.doi.org/10.1016/j.jesp.2014.04.011

Moritz, B. B., Hill, A. V., & Donohue, K. L. (2013). Individual differences in the newsvendor problem: Behavior and cognitive reflection. *Journal of Operations Management, 31*(1), 72-85.

Moskowitz, D. S. (2004). Does Elevated Power Lead to Approach and Reduced Power to Inhibition? Comment on Keltner, Gruenfeld, and Anderson (2003). *Psychological Review, 111*(3), 808-811. doi:10.1037/0033-295X.111.3.808

Mumford, M. D. (2011). A hale farewell: The state of leadership research. *The Leadership Quarterly, 22*(1), 1-7. doi:https://doi.org/10.1016/j.leaqua.2010.12.001

Neely, A. (2005). The evolution of performance measurement research: developments in the last decade and a research agenda for the next. *International journal of operations & production management, 25*(12), 1264-1277.

Neely, A., Gregory, M., & Platts, K. (2005). Performance measurement system design: a literature review and research agenda. *International journal of operations & production management, 25*(12), 1228-1263.

Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *The Journal of Strategic Information Systems, 24*(1), 3-14. doi:https://doi.org/10.1016/j.jsis.2015.02.001

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology, 2*(2), 175.

O'Connor, S. (2016, September 8). When your boss is an algorithm. *Financial Times*. Retrieved from https://www.ft.com/content/88fdc58e-754f-11e6-b60a-de4532d5ea35

Olivares, M., Terwiesch, C., & Cassorla, L. (2008). Structural Estimation of the Newsvendor Model: An Application to Reserving Operating Room Time. *Management Science, 54*(1), 41-55. doi:10.1287/mnsc.1070.0756

Open-Science-Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). doi:10.1126/science.aac4716

Overbeck, J. R., & Park, B. (2001). When power does not corrupt: Superior individuation processes among powerful perceivers. *Journal of Personality and Social Psychology, 81*(4), 549-565. doi:10.1037/0022-3514.81.4.549

Papke-Shields, K. E., Beise, C., & Quan, J. (2010). Do project managers practice what they preach, and does it matter to project success? *International Journal of Project Management, 28*(7), 650-662.

Pasquale, F. (2016). Bittersweet Mysteries of Machine Learning (A Provocation). *Media Policy Blog.* Retrieved from http://blogs.lse.ac.uk/mediapolicyproject/2016/02/05/bittersweet-mysteries-of-machine-learning-a-provocation/

Peus, C., Kerschreiter, R., Traut-Mattausch, E., & Frey, D. (2010). Ethics and economic success. *Zeitschrift für Psychologie/Journal of Psychology, 218*(4), 195.

Phan, P., Wright, M., & Lee, S.-H. (2017). Of Robots, Artificial Intelligence, and Work. *The academy of management perspectives, 31*(4), 253-255. doi:10.5465/amp.2017.0199

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team, R. (2015). _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-120, <URL: http://CRAN.R-project.org/package=nlme>.

Pitesa, M., & Thau, S. (2013). Masters of the universe: How power and accountability influence self-serving decisions under moral hazard. *Journal of Applied Psychology, 98*(3), 550-558. doi:10.1037/a0031697

PMI. (2013). *A Guide to the Project Management Body of Knowledge (PMBOK Guide), Fifth Edition*. Pennsylvania, USA: Project Management Institute, Inc.

Pratto, F. (2016). On power and empowerment. *British Journal of Social Psychology, 55*(1), 1-20. doi:10.1111/bjso.12135

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*(4), 741-763. doi:10.1037/0022-3514.67.4.741

R Core Team, R. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/.

Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S., & Weber, R. A. (2015). Assessing the Robustness of Power Posing: No Effect on Hormones and Risk Tolerance in a Large Sample of Men and Women. *Psychological Science, 26*(5), 653-656. doi:10.1177/0956797614553946

Raven, B. H. (1992). A power/interaction model of interpersonal influence: French and Raven thirty years later. *Journal of Social Behavior & Personality, 7*(2), 217.

Ren, Y., & Croson, R. (2013). Overconfidence in Newsvendor Orders: An Experimental Study. *Management Science, 59*(11), 2502-2517. doi:10.1287/mnsc.2013.1715

Robbins, A. (2015). The problem with satisfied patients. Retrieved from http://www.theatlantic.com/health/archive/2015/04/the-problem-with-satisfied-patients/390684/?utm_source=atlfb
Rus, D., van Knippenberg, D., & Wisse, B. (2012). Leader power and self-serving behavior: The Moderating Role of Accountability. *The Leadership Quarterly, 23*(1), 13-26. doi:http://dx.doi.org/10.1016/j.leaqua.2011.11.002
Russell, B. (1957). *Power*. London, Great Britain: George Allen & Unwin Ltd.
Sagristano, M. D., Trope, Y., & Liberman, N. (2002). Time-dependent gambling: odds now, money later. *Journal of Experimental Psychology: General, 131*(3), 364-376.
Schaerer, M., Galinsky, A. D., Lee, A. J., & Thau, S. (2017). Contextualizing Social Power Research Within Organizational Behavior. In D. L. Ferris, R. E. Johnson, & C. Sedikides (Eds.), *The Self at Work* (pp. 214-242). New York, NY: Routledge.
Schiffels, S., Fügener, A., Kolisch, R., & Brunner, O. J. (2014). On the assessment of costs in a newsvendor environment: Insights from an experimental study. *Omega, 43*, 1-8.
Schmid, E. A., Pircher Verdorfer, A., & Peus, C. (2017). Shedding Light on Leaders' Self-Interest: Theory and Measurement of Exploitative Leadership. *Journal of Management*. doi:https://doi-org.eaccess.ub.tum.de/10.1177/0149206317707810
Schmitt, M., & Meinberger, C. (2017). #DeineWahl - YouTuber fragen Angela Merkel | Mit Ischtar Isik, AlexiBexi, MrWissen2go, ItsColeslaw. Retrieved from https://www.youtube.com/watch?v=Uq2zIzscPgY
Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*(2), 195-202.
Schweitzer, M. E., & Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science, 46*(3), 404-420.
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. doi:https://doi.org/10.1016/j.leaqua.2012.09.001
See, K. E., Morrison, E. W., Rothman, N. B., & Soll, J. B. (2011). The detrimental effects of power on confidence, advice taking, and accuracy. *Organizational Behavior and Human Decision Processes, 116*(2), 272-285. doi:10.1016/j.obhdp.2011.07.006
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science, 22*(11), 1359-1366. doi:10.1177/0956797611417632
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (October 14, 2012). A 21 Word Solution. Retrieved from https://ssrn.com/abstract=2160588
Smith, K. M., & Apicella, C. L. (2016). Winners, losers, and posers: The effect of power poses on testosterone and risk-taking following competition. *Hormones and Behavior*. doi:http://dx.doi.org/10.1016/j.yhbeh.2016.11.003
Smith, P. K., & Hofmann, W. (2016). Power in everyday life. *Proceedings of the National Academy of Sciences, 113*(36), 10043-10048. doi:10.1073/pnas.1604820113
Smith, P. K., & Magee, J. C. (2015). The interpersonal nature of power and status. *Current Opinion in Behavioral Sciences, 3*, 152-156.
Smith, P. K., & Trope, Y. (2006). You focus on the forest when you're in charge of the trees: Power priming and abstract information processing. *Journal of Personality and Social Psychology, 90*(4), 578-596. doi:10.1037/0022-3514.90.4.578
Smith, P. K., Wigboldus, D. H., & Dijksterhuis, A. (2008). Abstract thinking increases one's sense of power. *Journal of Experimental Social Psychology, 44*(2), 378-385.

Steinberg, U., Knipfer, K., & Peus, C. (2019). *Power Influences Information Processing: Testing Predictions of the Approach/Inhibition Theory and the Social Distance Theory of Power* Working Paper. TUM School of Management. Technical University of Munich.

Steinberg, U., Peus, C., & de Rijcke, S. (2019). *Missing out by Design: How Power Explains Consequences of Digitized Performance Management*. Working Paper. TUM School of Management. Technical University of Munich.

Steinberg, U., Schiffels, S., Fügener, A., & Peus, C. (2019). *Mind the Bottom Line: How Psychological Power and Costs Influence Time Planning in Projects*. Working Paper. TUM School of Management. Technical University of Munich.

Stingl, V., & Geraldi, J. (2017). Errors, lies and misunderstandings: Systematic review on behavioural decision making in projects. *International Journal of Project Management, 35*(2), 121-135.

Stirling, A. (2008). "Opening up" and "closing down" power, participation, and pluralism in the social appraisal of technology. *Science, Technology, & Human Values, 33*(2), 262-294.

Stirling, A. (2010). Keep it complex. *Nature, 468*(7327), 1029-1031.

Stone, D. L., & Deadrick, D. L. (2015). Challenges and opportunities affecting the future of human resource management. *Human Resource Management Review, 25*(2), 139-145. doi:https://doi.org/10.1016/j.hrmr.2015.01.003

Stone, D. L., Deadrick, D. L., Lukaszewski, K. M., & Johnson, R. (2015). The influence of technology on the future of human resource management. *Human Resource Management Review, 25*(2), 216-231. doi:https://doi.org/10.1016/j.hrmr.2015.01.002

Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). A German version of carver and white's BIS/BAS scales. *Zeitschrift für Differentielle und Diagnostische Psychologie, 22*(3), 216-227.

Sturm, R. E., & Antonakis, J. (2015). Interpersonal Power: A Review, Critique, and Research Agenda. *Journal of Management, 41*(1), 136-163. doi:10.1177/0149206314555769

Sturm, R. E., Herz, H., & Antonakis, J. (2018). Leader Power: Rigorous Insights on its Causes and Consequences. *The Leadership Quarterly, I-II*.

Suls, J., & Wheeler, L. (2013). *Handbook of social comparison: Theory and research*: Springer Science & Business Media.

Svejvig, P., & Andersen, P. (2015). Rethinking project management: A structured literature review with a critical look at the brave new world. *International Journal of Project Management, 33*(2), 278-290.

Todorov, A., Goren, A., & Trope, Y. (2007). Probability as a psychological distance: Construal and preferences. *Journal of Experimental Social Psychology, 43*(3), 473-482.

Tost, L. P. (2015). When, why, and how do powerholders "feel the power"? Examining the links between structural and psychological power and reviving the connection between power and responsibility. *Research in Organizational Behavior, 35*, 29-56. doi:http://dx.doi.org/10.1016/j.riob.2015.10.004

Tost, L. P., Gino, F., & Larrick, R. P. (2012a). Power, competitiveness, and advice taking: Why the powerful don't listen. *Organizational Behavior and Human Decision Processes, 117*(1), 53-65.

Tost, L. P., Gino, F., & Larrick, R. P. (2012b). When power makes others speechless: The negative impact of leader power on team performance. *Academy of Management Journal, 56*(5), 1465-1486. doi:http://dx.doi.org/10.5465/amj.2011.0180

Townley, B. (1993). Foucault, power/knowledge, and its relevance for human resource management. *Academy of Management Review, 18*(3), 518-545.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440-463. doi:10.1037/a0018963

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: the density hypothesis. *Journal of Personality and Social Psychology, 95*(1), 36.

Weick, M., & Guinote, A. (2010). How long will it take? Power biases time predictions. *Journal of Experimental Social Psychology, 46*(4), 595-604. doi:http://dx.doi.org/10.1016/j.jesp.2010.03.005

Whitson, J. A., Liljenquist, K. A., Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., & Cadena, B. (2013). The blind leading: Power reduces awareness of constraints. *Journal of Experimental Social Psychology, 49*(3), 579-582. doi:10.1016/j.jesp.2012.10.009

Winerip, M. (2011, March 21). Grading New York Teachers: When the Formulas Lie. *The New York Times*. Retrieved from http://www.nytimes.com/2011/03/07/education/07winerip.html

Yukl, G. (2010). *Leadership in Organizations (7th Ed.)*. Upper East Saddle, New Jersey: Pearson.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*(1), 75-98.

Appendix A: Preregistration – Proposal Stage (Chapter 2, Study 1)

Short Introduction Time in Planning Project Proposals

The present experiment addresses the influence of decision-maker power on time-planning in projects that other people execute. Specifically, the experiment analyzes the influence of power on time planning in projects during the proposal stage. The uncertainty lies within the probability of winning the project proposal. The shorter a project is planned, the higher is the probability that the project proposal is won. However, monetary resources have to be invested to decrease the planned project duration. We consider both a high and a low-cost setting, where long and short planned durations are beneficial, respectively. The realization of winning the project proposal is based on a random draw and the probability of winning due to the planned project duration.

We rely on the social distance theory of power (Magee & Smith, 2013) to derive hypotheses for the present experiment.

Hypotheses

1. High-power individuals plan shorter times than low power individuals, due to a stronger focus on the central/abstract goal of winning the project (by offering the shortest time to completion)
2. In the high-cost setting, longer times are planned than in the low-cost setting.

Method

Design. 3 power (high vs. control vs. low) x 2 cost setting (high vs. low)

Participants, Sample Size, Exclusions. 300 undergraduate students will be randomly assigned to one experimental condition each. We aim for 50 participants per condition. The distribution of participants to experimental conditions may vary slightly due to the

randomization algorithm used in our experimental software and the exclusion of participants who fail the attention check after the experimental manipulation.

The sample size has been determined based on recommendations from power research for cases when effect sizes are unknown (Lammers et al., 2016).

If less than 40 participants remain in any one condition due to exclusions, we will organize an additional session with 30 participants to achieve a minimum of 40 participants per condition.

Procedure. The study will be conducted in a computer laboratory, all instructions, manipulations and dependent variables will be presented on the computer. A show-up fee and performance-based incentive will be paid upon departure.

Participants will be randomly assigned to computers and experimental conditions. At the end of the study we ask for age, gender, and mother tongue of the participants.

Materials (IV, DV, Moderators, Mediators)

Power Manipulation. We use the same power manipulation and attention check as in Moon and Chen (2014)

Manipulation Check. We use a manipulation check based on P. K. Smith et al. (2008)

Dependent Variable. The planned durations of ten project proposals – Participants are asked to plan the duration of project proposals with stochastic duration thresholds for acceptance. In total, the planning task includes 10 independent rounds, where one project proposal has to be planned within each round. The duration thresholds for acceptance are equal for all participants and follow a discrete uniform distribution. All the information necessary to calculate the expected profit maximizing planned project duration is provided. After each round, feedback about the duration threshold and the costs resulting from the participant's decision is provided. The participants' compensation for the experiment depends on their performance in the time planning task.

Control questions. After the power manipulation, we will ask participants whether they found the task to be difficult (yes / no, see, Schwarz et al., 1991). We will also ask for age, gender, and the participant's mother tongue.

Data Analysis Plan

Main Analysis. We will calculate a linear mixed model to analyze the main effects of power and cost setting. The linear mixed model allows us to control for participants and the thresholds for acceptance as random factors.

Exploratory. (1) We will conduct the same analyses for the first round (project proposal) and the first three rounds only, to account for potential effects of experience with the task on the planned durations and on the effects of the power manipulation itself. As Schwarz et al. (1991) demonstrated, tasks perceived as difficult may weaken the effects of experiential power manipulations. Thus, the effect of our manipulation is expected to weaken over the 10 rounds. (2) To explore the possibility that participants who had difficulties imagining the experimental manipulation responded differently, we will use the control question to exclude those participants according to the control question and repeat the analysis.

Appendix B: Instructions Time Planning Task – Proposal Stage (Chapter 2, Study 1)

Proposal stage high cost

Your task is to plan the duration of a project that a customer will consider purchasing. The project is performed by someone else independently of you.

As a default, the project requires 300 days. However, it is possible to speed up the project for additional costs of 9 Experimental Currency Units (ECU) per day and to reduce the duration by up to 100 days. Thus, the minimum duration is 200.

The customer will only accept your project plan if the duration is less than or equal to the customer's maximum acceptable duration. The maximum acceptable project duration ranges from 200 to 299 days, with each value between 200 and 299 (including 299) having the same probability.

If 300 days are planned, the probability that the project plan is accepted is 0% (i.e., 0 out of 100 project plans with a duration of 300 days are accepted), and the likelihood of acceptance increases by 1% for each day the planned project duration is reduced (i.e., 1 out of 100 project plans is accepted additionally for each day planned shorter). Thus, the project plan acceptance rate is 100% (i.e., 100 out of 100 project plans are accepted) if 200 days are planned.

If your project plan is accepted, you earn a fixed revenue of 450 ECU. This revenue minus the costs to speed up the project defines your profit. If your project plan is not accepted, your profit is 0 ECU.

On the next screens, you will be asked to plan the duration for 10 projects (in days). Each project is independent of all of the other projects. The cumulative profit of the 10 projects (in ECU) will be divided by 60 to determine your final compensation (in Euro €).

Proposal stage low cost

Your task is to plan the duration of a project that a customer will consider purchasing. The project is performed by someone else independently of you.

As a default, the project requires 300 days. However, it is possible to speed up the project for additional costs of 9 Experimental Currency Units (ECU) per day and to reduce the duration by up to 100 days. Thus, the minimum duration is 200.

The customer will only accept your project plan if the duration is less than or equal to the customer's maximum acceptable duration. The maximum acceptable project duration ranges from 200 to 299 days, with each value between 200 and 299 (including 299) having the same probability.

If 300 days are planned, the probability that the project plan is accepted is 0% (i.e., 0 out of 100 project plans with a duration of 300 days are accepted), and the likelihood of acceptance increases by 1% for each day the planned project duration is reduced (i.e., 1 out of 100 project plans is accepted additionally for each day planned shorter). Thus, the project plan acceptance rate is 100% (i.e., 100 out of 100 project plans are accepted) if 200 days are planned.

If your project plan is accepted, you earn a fixed revenue of 450 ECU. This revenue minus the costs to speed up the project defines your profit. If your project plan is not accepted, your profit is 0 ECU.

On the next screens, you will be asked to plan the duration for 10 projects (in days). Each project is independent of all of the other projects. The cumulative profit of the 10 projects (in ECU) will be divided by 60 to determine your final compensation (in Euro €).

Table 11. Results of the Preregistered Exploratory Analyses at the Proposal Stage

	Round 1 only			Round 1-3 only				After exclusions			
Parameter				Variance		Chi^2		Variance		Chi^2	
Random effects											
Participants											
Intercept				160.04		175.68		128.44		903.01	
Stimuli											
Intercept				0		0					
Low costs				0.27				0.02			
High costs				3.20		0		5.29		11.57	
Low-power				1.52				3.77			
Control				0.11				1.15			
High-power				1.80		0.19		0.41		0	
Residual				184.63				185.44			
	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*df*	*t*	*p*	Estimate (*SE*)	*df*	*t*	*p*
Fixed effects											
Intercept	237.71			240.11				238.76			
	(2.41)	98.55	$p < .0001$	(1.93)	22.38	124.44	$p < .0001$	(1.62)	116.89	147.41	$p < .0001$
Low costs x	19.78			19.90				20.21			
High costs	(2.36)	8.40	$p < .0001$	(1.88)	23.07	10.61	$p < .0001$	(1.61)	84.00	12.52	$p < .0001$
Low power x	0.39			-1.56				-3.58			
Control	(2.90)	0.13	$p = .8940$	(2.19)	48.61	-0.71	$p = .4790$	(1.81)	169.18	-1.98	$p = .0490$
Low-power x	1.22			0.98				-0.13			
High-power	(2.90)	0.42	$p = .6740$	(2.13	273.56	0.46	$p = .6440$	(1.81)	252.67	-0.07	$p = .9410$
High power x	0.83			2.56				3.45			
Control[a]	(2.85)	0.29	$p = .8940$	(2.19)	34.57	1.17	$p = .2520$	(1.74)	250.78	1.984	$p = .0483$
-2log-likelihood							7589.0				23513.4

Note. Table 1 shows the results of the preregistered exploratory analyses of Experiment 1. [a]Taken from an identical model with recoded power group

Appendix D: Preregistration – Execution Stage (Chapter 2, Study 2)

Short Introduction Project Execution Stage

The present experiment addresses the influence of decision-maker power on time-planning in projects that other people execute. Specifically, the experiment analyzes the influence of power on time planning in projects during the execution stage. The uncertainty lies within the realized project duration. If the realized duration exceeds the planned project duration, higher overtime costs occur. For the planned duration, regular costs apply. Here, we consider both a high and a low-cost setting, where short and long planned durations are beneficial, respectively. The realized durations are based on a random draw.

We rely on the social distance theory of power (Magee & Smith, 2013) to derive hypotheses for the present experiment.

Hypotheses

3. High-power individuals plan longer times than low power individuals, due to a focus on the central/abstract goal of completing the project on time (by planning more time to completion)
4. In the high-cost setting, shorter times are planned than in the low-cost setting.

Method

Design. 3 power (high vs. control vs. low) x 2 cost setting (high vs. low)

Participants, Sample Size, Exclusions. 300 undergraduate students will be randomly assigned to one experimental condition each. We aim for 50 participants per condition. The distribution of participants to experimental conditions may vary slightly due to the randomization algorithm used in our experimental software and the exclusion of participants who fail the attention check after the experimental manipulation.

The sample size has been determined based on recommendations from power research for cases when effect sizes are unknown (Lammers, Stoker, Rink, & Galinsky, 2016).
If less than 40 participants remain in any one condition due to exclusions, we will organize an additional session with 30 participants to achieve a minimum of 40 participants per condition.

Procedure. The study will be conducted in a computer laboratory, all instructions, manipulations and dependent variables will be presented on the computer. A show-up fee and performance-based incentive will be paid upon departure.

Participants will be randomly assigned to computers and then experimental conditions will be randomly as-signed by the experimental software when starting the experiment. At the end of the study we ask for age, gender, and mother tongue of the participants.

Materials (IV, DV, Moderators, Mediators). Power Manipulation: We use the same power manipulation as in Moon and Chen (2014)

Manipulation Check: We use a manipulation check based on Smith, Wigboldus, and Dijksterhuis (2008)

Dependent Variable: The planned durations of ten projects – Participants are asked to plan the duration of projects with stochastic realized durations. In total, the planning task includes 10 independent rounds, where one project has to be planned within each round. The realized durations are equal for all participants and follow a discrete uniform distribution. All the information necessary to calculate the expected profit maximizing planned project duration is provided. After each round, feedback about the realized duration and the costs is provided. The participants' compensation for the experiment depends on their performance in the time planning task.

Control questions: After the power manipulation, we will ask participants whether they found the task to be difficult (yes / no, see, Schwarz et al., 1991). We will also ask for age, gender, and the participant's mother tongue.

Data Analysis Plan

Main Analysis. We will calculate a linear mixed model to analyze the main effects of power and cost setting. The linear mixed model allows us to control for participants and the thresholds for acceptance as random factors.

Exploratory. (1) We will conduct the same analyses for the first round (project proposal) and first three rounds only, to account for potential effects of experience with the task on the planned durations and on the effects of the power manipulation itself. As Schwarz et al. (1991) demonstrated, tasks perceived as difficult may weaken the effects of experiential power manipulations. Thus, the effect of our manipulation is expected to weaken over the 10 rounds. (2) To explore the possibility that participants who had difficulties imagining the experimental manipulation responded differently, we will use the control question to exclude those participants according to the control question and repeat the analysis.

Appendix E: Instructions Time Planning Task – Execution Stage (Chapter 2, Study 2)

Execution stage high cost

You have been asked to plan the duration of several projects that somebody else is going to perform. You do not know how many days it will take to perform a given task. However, you know that each duration between 200 days and 300 days has the same probability for any project.

Each day planned costs the project ECU 9 because resources need to be booked.

If a project is realized in less days than planned, the costs accrue for all planned days. If the project takes longer than planned increased costs of ECU 12 are accounted for each additional day. That is, the additional costs for each day not planned in advance are ECU 3.

On the next screens, you will be asked to plan the duration for each of ten projects in days. The duration of each project is independent of the other projects.

The cumulative costs of the ten projects in ECU will be subtracted from the budget of ECU 30 000. The remaining budget will then be divided by 1 000 to determine your compensation in Euro €.

Execution stage low cost

You have been asked to plan the duration of several projects that somebody else is going to perform. You do not know how many days it will take to perform a given task. However, you know that each duration between 200 days and 300 days has the same probability for any project.

Each day planned costs the project ECU 3 because resources need to be booked.

If a project is realized in less days than planned, the costs accrue for all planned days. If the project takes longer than planned increased costs of ECU 12 are accounted for each additional day. That is, the additional costs for each day not planned in advance are ECU 9.

On the next screens, you will be asked to plan the duration for each of ten projects in days. The duration of each project is independent of the other projects.

The cumulative costs of the ten projects in ECU will be subtracted from the budget of ECU 10 000. The remaining budget will then be divided by 200 to determine your compensation in Euro €.

Table 12. Results of the Preregistered Exploratory Analyses at the Execution Stage

	Round 1 only			Round 1-3 only				After exclusions			
Parameter				Variance		Chi^2		Variance		Chi^2	
Random effects											
Participants											
Intercept				178.70		174.19		130.64		751.70	
Stimuli											
Intercept				0				4.22		34.73	
Low costs				0							
High costs				0		0.06					
Low-power				13.14							
Control				12.83							
High-power				24.00		1.48					
Residual				209.60				215.48			
	Estimate (*SE*)	*t*	*p*	Estimate (*SE*)	*df*	*t*	*p*	Estimate (*SE*)	*df*	*t*	*p*
Fixed effects											
Intercept	254.28			258.11				258.16			
	(2.48)	102.70	$p < .0001$	(2.77)	6.91	93.07	$p < .0001$	(1.62)	160.14	159.75	$p < .0001$
Low costs x	-8.63			-9.51				-12.09			
High costs	(2.47)	-3.49	$p < .0001$	(1.82)	300.24	-5.23	$p < .0001$	(1.47)	282.80	-8.24	$p < .0001$
Low power x	-0.04			1.35				2.73			
Control	(3.03)	-0.01	$p = .9885$	(2.23)	298.18	0.61	$p = .5460$	(1.80)	282.80	1.52	$p = .1310$
Low-power x	-2.42			-1.10				0.98			
High-power	(3.02)	-0.80	$p = .4241$	(2.34)	43.17	-0.47	$p = .6430$	(1.83)	282.80	0.54	$p = .5920$
High power x	-2.38			-2.44				-1.74			
Control[a]	(3.03)	-0.79	$p = .4325$	(2.41)	16.29	-1.02	$p = .3247$	(1.84)	282.80	-0.99	$p = .3250$
-2log-likelihood							7779.8				23807.2

Note. Table 2 shows the results of the preregistered exploratory analyses of Experiment 2. [a]Taken from an identical model with recoded power group

Appendix G: Potential Mediators (Chapter 3, Study 3)

Table 13

Control Questions used in Study 3

1. At work, I am used to make decisions involving a desirability and feasibility trade-off.
2. When evaluating the arguments, I imagined that I would have to make the decision.
3. When evaluating the arguments, I imagined that others would have to make the decision.
4. The arguments referred to topics that occur at work frequently.
5. I assume that a report summarized by a positive argument would also shed light on negative aspects.
6. I assume that a report summarized by a negative argument would also shed light on positive aspects.
7. If I suggest my supervisor asks: “Is this feasible?”
8. If I suggest my supervisor asks: “Is this desirable?”
9. I was very well able of evaluating the arguments.
10. I was very sure of my evaluations

Table 14.

Correlations of the Independent Variable with Potential Mediators

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. $Power_{LH}$														
2. Social distance	-.01													
3. Abstract thinking	.06	-.23***												
4. Social Dominance	.13	.08	-.17*											
5. Need for Cognition	.01	.10	.16*	-.12										
6. Positive Affect	.02	-.09	.12	-.14*	.33***									
7. Negative Affect	-.06	.09	-.25***	.13	-.18**	-.19**								
8. Promotion focus	-.01	.00	.10	.03	.10	.28***	.05							
9. Prevention focus	-.07	.06	-.09	.01	-.29***	-.05	.21**	.33***						
10. Behavioral inhibition	-.05	.10	-.03	-.05	-.33***	-.19**	.27***	.09	.46***					
11. Behavioral activation	.02	-.08	.17*	-.06	.22**	.33***	-.03	.63***	.22**	.06				
12. Mean desirability	.05	.06	-.17*	.16*	.00	.15*	.18**	.15*	.15*	.11	.18**			
13. Mean undesirability	.02	.14*	-.04	.04	.02	.06	.06	.03	-.02	.05	.00	.39***		
14. Mean feasibility	-.08	-.06	.08	-.08	.18**	.30***	-.02	.27***	-.02	-.04	.29***	.41***	.12	
15. Mean unfeasibility	-.03	.05	.13	-.16	.12	.12	-.17*	.03	-.01	.00	.03	-.04	.57***	.35***

Note. $Power_{LH}$ = dummy-coded low power group vs. high power group (with low power group = 0 and high-power group = 1). * $p < .05$; ** $p < .01$; *** $p < .001$