Year : 2020

Interpersonal Power in Organizational Context : Empirical and Methodological Advances

Khademi Mahshid

Khademi Mahshid, 2020, Interpersonal Power in Organizational Context : Empirical and Methodological Advances

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_18DD954DBF0A8

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT DE COMPORTEMENT ORGANISATIONNEL

Interpersonal Power in Organizational Context: Empirical and Methodological Advances

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteure ès Sciences Économiques, mention « Management »

par

Mahshid KHADEMI

Directrice de thèse
Prof. Marianne Schmid Mast

Jury

Prof. Felicitas Morhart, Présidente
Prof. Christian Zehnder, expert interne
Prof. Petra Schmid, experte externe

LAUSANNE
2020

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT DE COMPORTEMENT ORGANISATIONNEL

Interpersonal Power in Organizational Context: Empirical and Methodological Advances

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Docteure ès Sciences Économiques, mention « Management »

par

Mahshid KHADEMI

Directrice de thèse
Prof. Marianne Schmid Mast

Jury

Prof. Felicitas Morhart, Présidente
Prof. Christian Zehnder, expert interne
Prof. Petra Schmid, experte externe

LAUSANNE
2020

UNIL | Université de Lausanne
HEC Lausanne
Le Décanat
Bâtiment Internef
CH-1015 Lausanne

IMPRIMATUR

Sans se prononcer sur les opinions de l'autrice, la Faculté des Hautes Etudes Commerciales de l'Université de Lausanne autorise l'impression de la thèse de Madame Mahshid Khademi, titulaire d'un bachelor in Industrial Engineering et d'un master in Business Administration de Sharif University of Technology, en vue de l'obtention du grade de docteure ès Sciences économiques, mention management.

La thèse est intitulée :

INTERPERSONAL POWER IN ORGANIZATIONAL CONTEXT: EMPIRICAL AND METHODOLOGICAL ADVANCES

Lausanne, le 11.08.2020

Le doyen

Jean-Philippe Bonardi

Members of the thesis committee

Prof. Marianne Schmid Mast

HEC Lausanne, Department of Organizational Behavior

Thesis supervisor

Prof. Christian Zehnder

HEC Lausanne, Department of Organizational Behavior

Internal member of the thesis committee

Prof. Petra Schmid

ETH Zurich, Department of Management, Technology, and Economics

External member of the thesis committe

University of Lausanne
Faculty of Business and Economics

Ph.D. in Economics
Subject area "Management"

I hereby certify that I have examined the doctoral thesis of

Mahshid KHADEMI

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: 15.7.2020

Prof. Marianne SCHMID MAST
Thesis supervisor

University of Lausanne
Faculty of Business and Economics

Ph.D. in Economics
Subject area “Management”

I hereby certify that I have examined the doctoral thesis of

Mahshid KHADEMI

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: 09.07.2020

Prof. Christian ZEHNDER
Internal member of the doctoral committee

University of Lausanne
Faculty of Business and Economics

Ph.D. in Economics
Subject area "Management"

I hereby certify that I have examined the doctoral thesis of

Mahshid KHADEMI

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: 10.7.2020

Prof. Petra SCHMID
External member of the doctoral committee

ACKNOWLEDGMENTS

I would like to first express my deep gratitude to my supervisor, an academic role model Professor Schmid Mast for giving me the opportunity to do research in her team. She gave me great autonomy while being supportive all the way during my doctoral studies. It was a great privilege and honor to do research under her supervision. Being a strong leader in academia, she taught me how to be humble, yet to believe in what you have worked hard and accomplished. Her tireless guidance through this journey helped me professionally and personally.

I would like to thank Prof. Christian Zehnder, my internal advisor, for his support, invaluable discussions, and the scrutiny in defining the elements of a clean research design. Many of the research ideas in this thesis were developed during the CUSO summer school that he established. I would also like to thank him for his mentorship, dedication, and his great sense of humor which was very helpful in the face of obstacles.

Besides my advisors, I thank the external expert of my thesis committee, Prof. Petra Schmid for her insightful comments and great suggestions that significantly inspired the outcome of this dissertation. I am also very grateful to my colleagues and co-authors on various research projects of my PhD: Oriana De Saint Priest, Denise Frauendorfer, Tristan Palese, Benjamin Tur, Victoria Valle Lara. I also thank Anina Eggenberger and Katja Schwab-Weis for their administrative support and great assistance.

A good support system is important for surviving and staying sane in the doctoral school. Thus, I am so grateful for my fellow lab and office mates, present and past, for the inspiring discussions, helpful feedback, and for all the fun and happy moments during the coffee breaks, Christmas, and summer parties. I was also lucky to be part of the large community of Iranian students in Switzerland that has provided distractions to reset our mind outside of work and

research; I thank all and every one of them especially my close friends in “sans-alcohol” who claim that they can boost your mood even better than a drink!

I am extremely grateful to my mother, father and my two brothers for their love, caring and prayers whom without their support I would not have made it this far. My mother, a strong independent woman who sacrificed all for her children always pushed us to achieve more and be restless. My two brothers have been my guardians all my life and I love them both dearly and thank them for all their advice and support. I know I always have my family to count on.

My deepest appreciation goes to my best friend, my one and only Amin. I am extremely grateful for his love, understanding, and care during the past thirteen years. He gave me strength in my darkest days and had faith in me when I did not have in myself. I am constantly learning from him and beside him how to live and how to be. There are no words to convey how much I love him. Thus, I would get help from a poem by Rumi:

"In your light I learn how to love.
In your beauty, how to make poems.
You dance inside my chest,
where no one sees you,
but sometimes I do, and that
sight becomes this art."

Thesis overview

Power as the influence that one can exert over others (Schmid Mast, Jonas, & Hall, 2009) is the fundamental force in our social interactions. Power changes thoughts, emotions, and behaviors of the powerholder (Galinsky, Rucker, & Magee, 2015; Guinote, 2017; Kipnis, 1972) and prepares the powerholder in the path to leadership. For instance, individuals in power tend to be more goal- and reward-oriented (Galinsky et al., 2003; Magee, Galinsky, & Gruenfeld, 2007; Slabu & Guinote, 2010), feel more distant (Magee & Smith, 2013), become abstract thinkers (Smith & Trope, 2006), and prioritize their goals better (Guinote, 2007b; Schmid, Schmid Mast, & Mast, 2015).

In this dissertation, I conducted experiments to better understand the behavioral outcomes of power in organizational settings. However, current methods that examine power in experimental studies have raised many issues such as generating demand effects and lack of external validity (Flynn, Gruenfeld, Molm, & Polzer, 2011; Sturm & Antonakis, 2015). Therefore, I was required to find a robust method that enables me to study power in experimental designs. In my first paper, I suggested a method that induces power in a robust and consequential manner in the experimental settings and I tested it against a prevalent method vastly used in the power domain. In my second paper, I used a version of the proposed method in Paper 1 to investigate the effect of power on unethical conducts. Finally, in my last paper I looked at the emergent hierarchy and how it can be influenced in the context of interdependent groups where all members need to contribute in order to reach a collective outcome. This thesis takes a very first step to achieve methodological clarity in the field of social power and it answered important questions about risk taking, unethical conducts, and emergent hierarchy in the organizational context.

In the next chapter, I will first review the definitions of power and state what definition I use through this dissertation. Then, I argue why power is important and discuss

the relevance of power in organizational context. I will also briefly review how power is studied in the experiments. Finally, I discuss the methodological and empirical contributions of this thesis and how it advances our understanding of power in social and organizational context.

Definitions of power

Power is an indispensable element of our social life. Weber (1947) defined power as the ability to carry out one's own will. Later on, Emerson (1962) introduced the notion of independence/freedom and formed a relational context in which "the influencer is independent and the target is dependent for resources". With development of Leadership theories, power was formulated as a means of influence over ideas, emotions, and behaviors (Siu, 1979) or having impact on others (McClelland, 1975). However, the conceptualization of power changed through the course of time. In the recent definitions of power there exists a convergence on the notion of control over *valued resources* (Galinsky, Gruenfeld, & Magee, 2003; Keltner, Gruenfeld, Anderson., 2003; Malhotra & Gino, 2011). Some scholars also highlighted the asymmetric or unshared aspect of such control (Dubois, Rucker, & Galinsky, 2011; Galinsky, Magee, Gruenfeld, Whitson, & Liljenquist, 2008; Magee & Galinsky, 2008; Magee & Smith, 2013) and in some definitions specific behavioral examples were made to add more clarity such as "… providing or withholding resources or administering punishments" (Keltner, et al., 2003, p. 265). The main common aspect of these definitions is the idea of power as a means of control over resources. Therefore, in this dissertation I will adopt this definition that commonly used in power research as well: *power is asymmetric control over valued resources* (Depret & Fiske, 1993; Fiske, 2010; Magee & Galinsky, 2008; Keltner et al., 2003).

Note that in this dissertation I distinguish between social power and similar concepts such as sense of power which refers to the perception of one's own influence in a certain

situation (Anderson, John, Keltner, 2012) or personal power that is generally viewed as the extent one can act and make things happen according to his/her will (Overbeck & Park, 2001; Van Dijke & Poppe, 2006). Social power rather roots in the dependencies and influences that individuals can enforce over others (Schmid Mast, 2010) and thus is embedded in the social and relational context.

Power in Organizational Settings

Power in organizational settings is commonplace. A substantial body of evidence shows that execution of power can predict organizational outcomes such as manager's level of influence and effectiveness (Raven, Schwarzwald, & Koslowsky, 1998; Yukl & Falbe, 1991), employees' commitment (Koslowsky, Schwarzwald, & Ashuri, 2001; Pierro, Raven, Amato, & Bélanger, 2013), innovative behavior (Krause, 2004), performance (Pierro, Kruglanski, & Raven, 2012), job stress (Erkutlu & Chafra, 2006), and job satisfaction (Raven, et al., 1998).

But power in organizational settings has different bases. It is important to understand the effect of power bases because with each power base, different influence processes will emerge that can lead to employee's commitment (Pierro, et al., 2013) and compliance (Koslowsky, et al., 2001). According to French & Raven's (1959) framework, power can have different bases such as reward, coercive, legitimate, expert, and referent power. *Reward* power can be attributed to a supervisor using different forms of positive stimulus from monetary compensations and promotions to any signs of an approval. In the same but opposite vein, influencing the targets by threats and punishments or withholding valued resources could be attributed to the use of *coercive* power. *Legitimate* power is mostly found in formal structures and among individuals with different levels of authority in which an obligation forces individuals to obey the powerholder. What is less discussed in management and organizational behavior research is expert and referent power (Elias, 2008). Experts have

power because we assume that they know best even if we do not understand the reasons behind their argumentations; but sometimes we listen to a person because of a feeling of admiration, a unity of values, a common identity (Raven, 1992; Tuk, Verlegh, Smidts, & Wigboldus, 2019); this person holds *referent* power. I distinguish referent power from status or prestige as status demonstrates the respect one receives from others (Magee & Galinsky, 2008) but referent power is based on individuals identification with the powerholder (Raven, et al., 1998) that might consequently lead to respect and trust. Later a sixth base of power was added to this classification as *information* power which refers to the use of critical information or persuasive argumentation in order to exert influence (Raven, 1965).

How Power Affects Powerholder's Behavior

The notion of power is tied with freedom and agency (Rucker, Galinsky, & Dubois, 2012). Agency refers to an act of self-expression and proactive inclination towards behavioral response as an independent entity (Bakan, 1966). Theoretical and empirical pieces on different manifestations of power demonstrated a well-established link between power and agency. For instance, approach/inhibition theory (Keltner et al., 2003) posits that high-power activates behavioral approach system that increases the focus on the rewards and opportunities as opposed to low-power that increases a focus on threats and punishments. Therefore, powerholder are more action and reward oriented compared to low-power individuals (Galinsky et al., 2003) and consequently display more goal-oriented, risk-taking behavior (Anderson & Galinsky, 2006; Guinote, 2007b).

Social distance theory of power (Magee & Smith, 2013) proposes that powerholders feel psychologically distant from low-power people and hence more behave in accordance with one's own values and experiences (Tost, Gino, & Larrick, 2012). Moreover, they will engage in a higher level of construal that is shown to facilitate risk-taking in gain-context (Lermer, Streicher, Sachs, Raue, & Frey, 2015; Raue, Streicher, Lermer, & Frey, 2015).

According to the situated focus theory, powerholders are more attuned to situational goals, needs, and feelings (Guinote, 2007a). Therefore, powerful individuals are better at focusing on the relevant information (Slabu & Guinote, 2010), prioritization (Schmid, Schmid Mast, & Mast, 2015), and adapting their behavior across situations (Guinote, 2007b).

However, recent theoretical advancement suggests an alternative pathway for the effect of power on behavior. Tost (2019) argues that while psychological power can activate the cognitive network of power based on which agentic behavior displays, structural power can pursue a different path through a sense of responsibility to the dependent counterparts and therefore, lead to rather communal outcome behaviors. Tost claims that a focus of empirical findings on the outcome of psychological power (vs. structural power) produced a possible bias in our understanding towards power outcomes. It seems than although power frees individual in and of itself, it also bound powerholder's behavior through a sense of responsibility.

In sum, the approach/inhibition theory and social distance theory both provide prediction for an agentic and at time asocial view points towards powerholders (i.e., seeing their interaction partners in an instrumental way and feeling psychologically distant from powerless individuals). Situated focus theory proposed a goal-orientation viewpoint in the sense that powerholders focus on achieving their goal and would adapt their behavior accordingly –whether it concerns personal gains or communal outcomes. However, Tost propose that power can elevate powerholder through the path of dependency awareness.

How Power Is Studied in Experiments

Despite various definitions, conceptualizations, and theories of power most of the experimental studies in the domain of Social Psychology and Management used very few operationalizations of power (Schaerer, Lee, Galinsky, Thau, 2018). For instance, a meta-analysis showed that more than 50% of power experiments manipulated felt power (e.g., by

using power priming) and less than a third actually manipulated power in a relational situation (e.g., by structural context) (Schaerer, du Plessis, Yap, & Thau, 2018).

Manipulating felt power in experimental settings is done via various methods such as role playing and explicit power priming. In role playing method, participants are typically assigned to either the role of a boss/leader (high-power) or the role of a subordinate/follower (low-power) (e.g., Anderson & Berdahl, 2002; Briñol, Petty, Stavraki, 2012). The powerholder is in charge of evaluating the performance of the powerless or selecting a specific task for them. In power priming method, participants are mostly asked to write an essay about a memory. In high-power condition, they are asked to write about a time when they had power over somebody. In the low-power condition though, they must write about a time that someone else had power over them. Typically, a power-neutral condition is added in which participants should write about an event happened to them a day before their participations or about the last time they went grocery shopping (Galinsky et al., 2003).The idea behind power priming is that by writing an essay about a high-power situation, participant are primed with a high-power mindset and therefore, feel, act, and decide accordingly. Another version of this manipulation asks participants to imagine oneself in a predefined role (for the high-power manipulation imaging a boss and for the low-power manipulation imagining a subordinate) and then to write an essay about the relevant feelings and actions in their role (Dubois, Rucker, & Galinsky, 2010; Schmid, 2018).

The use of aforementioned methods raised several potential methodological concerns. For example, one possible concern is about the lack of realism and ecological validity that had a long history in psychological studies (Campbell, 1957). Ecological validity refers to a lack of correspondence between the state of the individuals in the experiment and the real world settings (Wegener & Blankenship, 2007). In priming methods, often no objective power or control is given to the participants in high-power roles; besides, playing the role of a

powerful person or recalling a powerful memory could be different from experiencing objective power (Flynn et al., 2011). Consequently, the emotional and behavioral outcomes of solely recalling a power-related memory or playing a powerful role could be different from experiencing power real time.

The second concern is about the potential for demand effects. Demand effects refer to the totality of the cues that make participants aware of the implicit motives of the study and change their behavior according to their guesses about the aim of the study (Orne, 1969). In priming method, as power is manipulated subjectively, participants are deliberately told to recall an incident about a power situation or play the role of a boss. Such salient cues about the purpose of the study can raise participants' awareness about the expected behavior (Sturm & Antonakis, 2015).

In my thesis I tried to take the very first step to address these concerns. In Paper 1, I thus suggest an alternative method that enables us to study power in a more robust manner. Then in paper 2, I used a version of this method to understand the link between power and unethical behavior. I also tested whether individual characteristics can play a role in the relation of power and unethical behavior. Finally, in my third paper I looked at the notion of power in the context of working groups. I looked at how hierarchy emerged inside leaderless groups and how it can be influenced in different task groups.

Paper 1- The Problem of Demand Effects in Power Studies: Moving Beyond Power Priming

Power priming method in spite of its popularity, raised many issues concerning demand effects (Sturm & Antonakis, 2015; Schaerer, Lee, Galinsky, Thau, 2018). Demand effects refer to the proper responses to experimental stimuli according to the given cues by the frame or instructions of the study (Zizzo, 2010). In this paper we intend to see if these concerns are valid. Power priming as it induces subjective power, requires the experimenter

to ensure that the participants are affected by the manipulation. Thus, the experimenters must guarantee the effect firstly through clear instructions to the participants and secondly by verifying the effect via a manipulation check. In this paper we argued that both of these practices could increase the chance of revealing critical information about the study and therefore engender demand effects.

Such demand characteristics will be more problematic when it occurs in non-consequential experimental designs such as hypothetical self-report designs (Lonati, Quiroga, Zehnder, & Antonakis, 2018; Podsakoff & Podsakoff, 2019); because, in such designs it is not costly for participants to comply with what is expected from them. Therefore, non-consequential designs in presence of demand prone conditions could lead to demand induced responses. In sum, in this paper, we set out to see whether the power priming method is prone to demand effects, if using manipulation checks play a role in producing such demand, and if using consequential designs can mitigate the effect of demand-prone methods.

In an incentivized study, we manipulated power level (high-power vs. control), power manipulation method (power priming vs. resource allocation), and whether a manipulation check is used or not after the power manipulation. We measured participant's risk-taking in a non-consequential and consequential way as the outcome behavior. Our results showed that power priming is more prone to demand effects when compared to resource allocation method. Moreover, adding an explicit manipulation check to the design exacerbated the thread of demand effects especially when the dependent variable was measured non-consequentially. These results show how demand prone methods could be dangerous especially when they are mixed with explicit manipulation checks and self-report measures of the behavioral outcome. Despite the widespread use of the power priming methods and its low implementation cost, we invite power researchers to be extra prudent in choosing methods and techniques that are prone to demand characteristics.

Paper 2- Leader's Unethical Behavior Depends on the Level of Power, Machiavellianism, and Narcissism

Power as the influence that one can exert over others (Schmid Mast, 2010) is present in our every-day interactions. Power can engender responsibility, coordination, and sacrifice; but it can also lead to dominance, opportunistic behavior, and corruption. Research also shows that power typically leads to immoral behavior (Bendahan, Zehdner, Pralong, & Antonakis, 2015; Dubois, Rucker, & Galinsky, 2014; Lammers, Stapel, & Galinsky, 2010); but that it can also elevate moral behavior depending on the individual and contextual characteristics (Chen, Lee-Chai, & Bargh, 2001; DeCelles, DeRue, Margolis, & Ceranic, 2012). Looking at these controversies, one can ask when powerholders show the dark side rather than the elevating, bright side of their power. We often say "Power tends to corrupt and absolute power corrupts absolutely"; but is it entirely true? Do individual differences play a role? In this paper, we intended to answer these questions.

Using a version of the method developed in Paper 1, we investigated the effect of power on (un)ethical behavior of the powerholders. In two studies we manipulated power level (high- vs. low-power) once without and once with self-serving opportunities (Study 1 and Study 2 respectively). In Study 1, power was manipulated by the number of followers and the amount of control powerholders had over the followers' outcomes. We observed that it is not power but an interaction between power and Machiavellianism that led to unethical behavior; Machiavellian powerholders were more likely to steal from their followers when they are appointed to the high-power positions compared to when they were appointed to the low-power positions. In Study 2 we manipulated power by the control that powerholders enforced over their followers' outcomes; besides we provided self-serving opportunities for powerholders by letting them use their power for self-gain. We observed that power with self-serving opportunities facilitated unethical behavior more than the power with no such

opportunities. The moderating effect of Machiavellianism on the link between power and unethical behavior in Study 1 transformed into a strong main effect in Study 2. Seemingly, power with self-serving opportunities provoked Machiavellians, regardless of their power position, to practice self-indulgence. Moreover, Narcissists in power positions with self-serving opportunities showed more unethical behavior when they are appointed to high- versus low-power positions (probably due to the boost of their sense of grandiosity). These results imply that power has a revealing effect; those individuals who are prone to manipulate and exploit others for personal gain and self-serving behavior, when appointed to high-power positions, are more likely to engage in unethical behavior.

Paper 3- From Hierarchical to Egalitarian: Hierarchy Steepness Depends on Speaking Time Feedback and Task Interdependence

In the last sections power was studied in individual level and dyads. In this paper, we looked at power dynamics manifested through emergent hierarchy in the context of leaderless groups. When individuals gather together in leaderless groups a hierarchy emerges rapidly within the early stages of group interactions (Bales, Strodtbeck, Mills, & Roseborough, 1951; Kalma, 1991). The steepness of the group hierarchy can be understood by the distribution of the floor time among the group members (Schmid Mast, 2001). The more the floor time is equally distributed among the members, the more egalitarian the group hierarchy is (Schmid Mast, 2002).

It has been shown that speaking time feedback (giving members information on how much each member is speaking as the discussion unfolds) would help the members to balance their contributions and flatten their group hierarchy (Bachour, Kaplan, & Dillenbourg, 2010; DiMicco, Hollenbach, Pandolfo, & Bender, 2007; Kim, Chang, Holland, & Pentland, 2008; Terken & Sturm, 2010). In this paper we aimed to see whether and how speaking time feedback and group task affect hierarchical formation in problem-solving groups. We also

observed the behavior of under- and over- participators (respectively members who spoke less than and who spoke more than an average group member) in presence of feedback on their speaking time behavior during the group discussion.

In a 2 (speaking time feedback vs. control) by 2 (high interdependent task vs. low interdependent task) design, we recruited 3-person groups for solving a selection task. Results showed that speaking time feedback helped over-participators to decrease their speaking time behavior; but it did not encourage the under-participators to contribute more. Moreover, providing speaking time feedback helped groups to build flatter hierarchies, mostly due to over-participators decreasing their speaking time, and that in groups faced with a rather interdependent task, steeper hierarchies emerged, mostly due to under-participators decreasing their speaking time. These findings show that a task that needs more sharing of information and thus coordination and collaboration, drives groups to form steeper hierarchies compared to a task that needs lower levels of coordination and collaboration.

Contributions

Methodological contributions. My thesis investigates power outcomes by focusing on actual behavior of the powerholders in consequential designs. In Paper 1 and 2, I gave participants real control over valued resource and assessed their behavior using incentivized designs. In Paper 3, I measured individual's actual dominance behavior (speaking time behavior) during group discussion and again in an incentivized design. My emphasis is on the use of consequential designs in conducting research despite of a widespread use of self-report and hypothetical designs in behavioral science (Baumeister, Vohs, & Funder, 2007; Patterson, Giles, & Teske, 2011; Petterson, 2008). The use of such designs imposes many issues when they are interpreted as actual behavior. Firstly, the outcome of the self-report measures does not necessarily correspond to real behavior (Furr, 2009); what is measured through such designs is rather what individuals think they would do and not what they would

actually do in the real situation. Secondly, we showed in Paper 1 that non-consequential designs are much more vulnerable than incentivized behavioral designs in presence of demand characteristics. Incentivized designs produce higher experimental realism and decrease the threat of demand characteristics by making it costly for participants to comply with social desirability bias or demand effects (de Quidt, Vesterlund, & Wilson, 2019). Therefore, I suggest that the focus of research designs in behavioral science should shift from the self-report measures and hypothetical designs to performance based measures.

Moreover, I tried to take the very first step to methodological clarity in power research. Power priming methods used to induce power in experimental designs raised several concerns regarding demand effects (Strum & Antonakis, 2015). In Paper 1, we set out to address these concerns. We tested an alternative method (resource allocation) against power priming and showed that while producing similar effects on felt power (if not stronger), resource allocation method harbors minimal potential for demand effects when compared to power priming.

Power research also suffers from the lack of proper control conditions (Schaerer, du Plessis, Yap, & Thau, 2018). But what is a good control condition for a power treatment? To achieve a baseline level in power, many studies put participants in a context unrelated to power such as writing about their last grocery shopping. The problem of an unrelated control condition is two-fold. Firstly, power is a relational concept and it is not meaningful to compare it to a non-relational notion. Secondly, participants in the treatment conditions receive explicit cues about the aim of the study (power) while in the control condition participants are clueless about power; this way the possible demand effects correlate with the treatment and produce a confound problem (Zizzo, 2010). Although control conditions by design demonstrate the effect of the treatment, if not designed properly they can be a source of unfair comparison or demand effects (Lonati et al., 2018). In this dissertation, I used

control conditions that were implemented in the same relational context as the treatments. Moreover, the design of the control conditions used in this dissertation ensures similar information sharing to treatment and control conditions.

Paper 1 also showed that running explicit manipulation checks can be a source of demand effects. Thus all manipulation checks done in this dissertation were conducted by using an independent but comparable sample to the original pool of participants in the experiment to decrease the chance of generating demand effects.

Empirical contributions. We assessed the link between power and risk-taking in Paper 1. However, we failed to fully replicate the main effect of power on risk-taking that was originally demonstrated by Anderson and Galisnky (2006). Powerholders in our study were not different from the control group in their risky choices. Previously, similar failed replications for a main effect of power on risk-taking has been reported in several studies (Jordan, Sivanathan, & Galinsky, 2011; Hiemer & Abele, 2012; Maner, Gailliot, Butz, & Peruche, 2007; Ronay & Von Hippel, 2010). In our study, despite of incorporating two methods of power manipulations (power priming and resource allocation) and three different measures of risk-taking (self-report, Bart, & lottery task) we still did not find the evidence for an effect of power on risk-taking. We therefore became skeptical about the existence of power/risk-taking link. Future studies might want to see whether the effect could be replicated with powerholders' risk-taking in different context; for instance when risky decisions influence not only the powerholder but also another counterpart as well.

Paper 2 showed that individuals in high-power positions will not necessarily become unethical and self-centered. But it is rather their individual characteristics or better said their pre-existing traits that help or hinder unethical behavior. We showed that Machiavellianism and Narcissism in general engender unethical behavior. However, they can also interact with power levels. For instance, power interacted with Machiavellianism explaining unethical

behavior in the context where power brought no self-serving opportunities to the powerholder; Machiavellians in high-power positions more often engaged in unethical decisions compared to Machiavellians in low-power situations. Narcissism on the other hand moderated the link between power and unethical behavior when power produced self-serving opportunities for the powerholders. Individuals high on Narcissistic orientations, when appointed to high-power positions with self-serving opportunities, displayed unethical behavior more so than the Narcissist in low-power positions. Altogether, the two studies showed the importance of individual differences in power related context and highlighted the importance of "choosing the right person for the right position".

The findings of Paper 2 are also not commonplace in power field. Power is more often associated with negative behavioral outcomes such as lying, cheating (Dubois et al., 2015), corrupted, and self-serving behavior (Bendahan, Zehnder, Pralong, & Antonakis, 2015). However, our results showed that power does not engender unethical behavior. There could be many reasons for our findings, but one plausible explanation argued by Tost (2015); different operationalizations of power can lead to different conceptual associations (Gawronski & Brannon, 2020; Tost & Johnson, 2019). Many of the studies that showed the negative and dark side of power induced psychological power (e.g., (e.g., Galinsky, Magee, Inesi, & Gruenfeld, 2006; Gruenfeld, Inesi, Magee, & Galisnky, 2008; Lammers, Galinsky, Gordijin, & Otten, 2012). However, such manifestation of power is rather associated with competitive attitude and strategic decision making; whereas social power induced by resource allocation method in this paper rather evoked a sense of dependency and responsibility and therefor may lead to different behavioral outcomes (Tost, 2015; Tost & Johnson, 2019). Our results are in line with the findings of Gawronski and Brannon (2020) that showed power induced via power priming method reduces the sensitivity to moral norms; but inducing power by assigning participants to social roles led to a higher sensitivity to moral norms.

In Paper 3 we successfully replicated the effect of speaking time feedback on group hierarchy. We added to that by showing that groups facing high interdependent tasks – a task that needs all members' collective contribution to achieve to the final performance- tend to build steeper hierarchies compared to the groups that faced a low interdependent task. These findings provided evidence in line with propositions of the theoretical piece by Halevy, Chou, and Galinsky (2011) that suggest interdependent groups need steeper hierarchy in order to achieve better coordination among team members. Indeed, the emergence of a steep hierarchy can be a result of the group task.

Practical implications. Believing in "power corrupts" sets powerholders unaccountable for their unethical conducts. Maybe it is time to give a chance to the belief that says "power does not corrupt, but people corrupt power". Our results show that power lured those who already had tendencies toward unethical behavior. Machiavellians and Narcissists more often engaged in unethical behavior in high-power positions. Therefore, I would like to pinpoint these toxic person/situation combinations (Machiavellian/high-power without and Narcissists/high-power with self-serving opportunity). Institutions thus might want to be extra prudent on their governance by building a strong ethical climate and more importantly pay attention to dark personality traits in personnel selections for high-stake jobs.

Paper 3 showed that giving dominant members of the group (over-participators) feedback on their speaking time behavior can help them adapting their speaking time behavior and consequently lead to a better distribution of the floor time in group discussions. Our results showed that the opposite effect is not true; the under-speakers do not increase their contributions when faced with such speaking time feedback. So it is fruitful to find other ways to encourage under-speakers to voice their opinions.

Implications for Future Research. In the three papers of this dissertation, we showed how power affects individuals, dyads, and groups and at times challenged the existing findings in power literature. For example the results of Paper 1 did not demonstrate the link between power and risk-taking and Paper 2 disproved the belief about corruptive nature of power. Now the question is that should we be doubtful about the findings emerged from other methods from now on? The robustness and the strength of the resource allocation method that used to induce power in this dissertation have been assessed several times; the method was also successfully tested against demand characteristics. This being said, if we decide to incorporate this method and approve its findings, then what shall we do with contradictory findings of the former studies? One answer to this question can be that we should be careful about the interpretation of the former findings with such high levels of demand effects. More importantly, we shall be skeptical about the findings of the study designs where power priming and explicit manipulation checks were mixed together with non-consequential designs (e.g., self-reported measures). Replication studies are needed to ensure the researcher about the reliability and robustness of these findings.

However, I would like to highlight that power is a complex notion. Therefore, its outcomes can also be seen as not straightforward and sometimes contradicting. For instance, it has been already shown that experienced power as opposed to social power can lead to different outcomes (Rucker, Hu, & Galinsky, 2014; Tost & Johnson, 2019). The different findings emerged from the resource allocation manipulation can be due to its relational and social manifestation of power which is far from psychological power that activates agency and goal-pursuit behavior. Moreover, how power is perceived by the powerholder might play a role on the behavioral outcomes of such a multifaceted concept (Schmid Mast, Khademi, Palese, 2020). Therefore, different outcomes of various manipulation methods should not necessarily mean the falsification of the former methods and findings. But, rather it should

encourage researcher to seek for a better understanding about the distinctions between the forms and perceptions of social power.

We should also not forget that power have different manifestations and can be fed by various sources such as knowledge, information, and referent power. In order to advance our understanding of power outcomes we also need to consider the different sources of power. In organizational settings, power is not always enforced through hierarchical structure, but it can manifest as expert power, information power, or referent power (e.g., the influence of a role-model) (French & Raven, 1959). Although research has shown that the possession of power can change the powerholder; the full extent of power facets on individual's cognitive, behavioral and emotional responses are not fully understood. Questions like how different types of power determine powerholders' understanding of power role, and eventually powerholders behavior remain unanswered. These gaps call for a plan of research to better understand the link between power types and their impact on powerholders.

Although the effect of power on powerholder is much discussed in the literature, most of the theories and operationalizations of power revolved around solely one notion of power as control over valued resources (most highlighted in coercive/reward power of French and Raven framework) (Keltner et al., 2003; Magee & Galinsky, 2008; Smith & Bargh, 2008). However, a person's power may stem from other sources than the administration of rewards and punishments. More importantly, execution of power does not always occur through organizational structures. Power as influence sometimes occurs in the opposite expected directions (Elias, 2008). For instance in upward influence attempts, it is the subordinate who influences the supervisor (Yukl & Falbe, 1991; Yukl & Tracey, 1992). Therefore, one potential concern is that the conducted studies do not effectively capture all the features and complexities of power because most of the findings from the experiments in this domain are

triggered by limited facets of power (mostly administering coercive power through psychological and structural power).

Moreover, Power is embedded in the social and relational context. Therefore, we need to pay more attention to these two fundamental notions when we study power in organizational context. We cannot (and probably should not) conclude an organizational behavior outcome by solely relying on felt power in a non-relational context. The manipulation of power (or any independent variable) should be considered within the specificities of the context. Therefore, I suggest researchers to return to relational operationalization of power in their experimental designs. Another interesting but understudied aspect of power is how it evolves through the course of time. To understand how powerholders change with power or how they maintain their power, more longitudinal studies are needed (e.g., Fiol, O'Connor, & Aguinis, 2001; Heller, 2019; Martorana, Galinsky, & Rao, 2005).

References

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.

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

Anderson, C., John, O. P., & Keltner, D. (2012). The personal sense of power. *Journal of Personality*, *80*(2), 313-344.

Bachour, K., Kaplan, F., & Dillenbourg, P. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies*, *3*(3), 203-213.

Bakan, D. (1966). *The duality of human existence*. Reading: Addison-Wesley.

Bales, R. F., Strodtbeck, F. L., Mills, T. M., & Roseborough, M. E. (1951). Channels of communication in small groups. *American Sociological Review*, *16*(4), 461-468.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?. *Perspectives on Psychological Science*, *2*(4), 396-403.

Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. *The Leadership Quarterly*, *26*(2), 101-122.

Briñol, P., Petty, R. E., & Stavraki, M. (2012). Power increases the reliance on first-impression thoughts. *Revista de Psicología Social*, *27*(3), 293-303.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, *54*(4), 297-312.

Dépret, E., & Fiske, S. T. (1993). Social cognition and power: Some cognitive consequences of social structure as a source of control deprivation. In *Control motivation and social cognition* (pp. 176-202). Springer, New York, NY.

De Quidt, J., Vesterlund, L., & Wilson, A. (2019). Experimenter demand effects. In A. Schram & A. Ule (Eds.), *Handbook of research methods and applications in experimental economics*. Cheltenham: Edward Elgar Publishing.

DiMicco, J. M., Hollenbach, K. J., Pandolfo, A., & Bender, W. (2007). The impact of increased awareness while face-to-face. *Human–Computer Interaction*, *22*(1-2), 47-96.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2010). The accentuation bias: Money literally looms larger (and sometimes smaller) to the powerless. *Social Psychological and Personality Science*, *1*(3), 199-205.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2011). Super size me: Product size as a signal of status. *Journal of Consumer Research, 38*(6), 1047-1062.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2015). Social class, power, and selfishness: When and why upper and lower class individuals behave unethically. *Journal of Personality and Social Psychology*, *108*(3), 436 -449.

Elias, S. (2008). Fifty years of influence in the workplace: The evolution of the French and Raven power taxonomy. *Journal of Management History, 14*(3), 267-283.

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review,27* (1), 31-41.

Erkutlu, H. V., & Chafra, J. (2006). Relationship between leadership power bases and job stress of subordinates: example from boutique hotels. *Management Research News, 29*(5), 285-297.

Fiol, C. M., O'Connor, E. J., & Aguinis, H. (2001). All for one and one for all? The development and transfer of power across organizational levels. *Academy of Management Review*, *26*(2), 224-242.

Fiske, S. T. (2010). Interpersonal stratification: Status, power, and subordination. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (p. 941–982). John Wiley & Sons Inc.

Flynn, F. J., Gruenfeld, D., Molm, L. D., & Polzer, J. T. (2011). Social psychological perspectives on power in organizations. *Administrative Science Quarterly*, *56*(4), 495-500.

French, J., & Raven, B. (Eds.). (1959). *The bases of social power* (Vol. 7). Michigan: Institute for Social Research.

Furr, R. M. (2009). Personality psychology as a truly behavioral science. *European Journal of Personality: Published for the European Association of Personality Psychology*, *23*(5), 369-401.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453-466.

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., Whitson, J. A., & Liljenquist, K. A. (2008). Power reduces the press of the situation: implications for creativity, conformity, and dissonance. *Journal of Personality and Social Psychology*, *95*(6), 1450-1466.

Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, *17*(12), 1068-1074.

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 handbooks in psychology. APA handbook of personality*

and social psychology, Vol. 3. Interpersonal relations (p. 421–460). American Psychological Association.

Gawronski, B., & Brannon, S. M. (2020). Power and moral dilemma judgments: Distinct effects of memory recall versus social roles. *Journal of Experimental Social Psychology*, *86*, 103908.

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.

Guinote, A. (2007a). Behavior variability and the situated focus theory of power. *European Review of Social Psychology*, *18*(1), 256-295.

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

Guinote, A. (2017). How power affects people: Activating, wanting, and goal seeking. *Annual Review of Psychology*, *68*, 353-381.

Halevy, N., Y. Chou, E., & D. Galinsky, A. (2011). A functional model of hierarchy: Why, how, and when vertical differentiation enhances group performance. *Organizational Psychology Review*, *1*(1), 32-52.

Heller, S. (2019). *Manipulating social power within-person*. Unpublished Manuscript.

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.

Jordan, J., Sivanathan, N., & Galinsky, A. D. (2011). Something to lose and nothing to gain: The role of stress in the interactive effect of power and stability on risk taking. *Administrative Science Quarterly*, *56*(4), 530-558.

Kalma, A. (1991). Hierarchisation and dominance assessment at first glance. *European Journal of Social Psychology*, *21*(2), 165-181.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265–284.

Kim, T., Chang, A., Holland, L., & Pentland, A. S. (2008). Meeting mediator: enhancing group collaborationusing sociometric feedback. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, USA,* 457-466.

Kipnis, D. (1972). Does power corrupt? *Journal of Personality and Social Psychology, 24*(1), 33–41.

Koslowsky, M., Schwarzwald, J., & Ashuri, S. (2001). On the relationship between subordinates' compliance to power sources and organisational attitudes. *Applied Psychology, 50*(3), 455-476.

Krause, D. E. (2004). Influence-based leadership as a determinant of the inclination to innovate and of innovation-related behaviors: An empirical investigation. *The Leadership Quarterly, 15*(1), 79-102.

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2012). Power increases social distance. *Social Psychological and Personality Science, 3*(3), 282-290.

Lermer, E., Streicher, B., Sachs, R., Raue, M., & Frey, D. (2015). The effect of construal level on risk-taking. *European Journal of Social Psychology, 45*(1), 99-109.

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

Lunenburg, F. C. (2012). Power and leadership: An influence process. *International Journal of Management, Business, and Administration, 15*(1), 1-9.

Magee, J. C., & Galinsky, A. D. (2008). 8 social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals, 2*(1), 351-398.

Magee, J. C., Galinsky, A. D., & Gruenfeld, D. H. (2007). Power, propensity to negotiate, and moving first in competitive interactions. *Personality and Social Psychology Bulletin, 33*(2), 200-212.

Magee, J. C., & Smith, P. K. (2013). The social distance theory of power. *Personality and Social Psychology Review, 17*(2), 158-186.

Malhotra, D., & Gino, F. (2011). The pursuit of power corrupts: How investing in outside options motivates opportunism in relationships. *Administrative Science Quarterly, 56*(4), 559-592.

Maner, J. K., Gailliot, M. T., Butz, D. A., & Peruche, B. M. (2007). Power, risk, and the status quo: Does power promote riskier or more conservative decision making?. *Personality and Social Psychology Bulletin, 33*(4), 451-462.

Martorana, P. V., Galinsky, A. D., & Rao, H. (2005). From system justification to system condemnation: Antecedents of attempts to change power hierarchies. *Research on Managing Groups and Teams, 7*, 283-313.

McClelland, D. (1975). *Power: the inner experience*. New York: Halsted Press.

Orne, M. T. (1969). Demand characteristics and the concept of quasi-controls. In R.

Rosenthal, & R. L. Rosnow (Eds.). *Artifact in behavioral research* (pp. 147–179). New York, NY: Academic Press.

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.

Patterson, M. L., Giles, H., & Teske, M. (2011). The decline of behavioral research? Examining language and communication journals. *Journal of Language and Social Psychology, 30*(3), 326-340.

Patterson, M. L. (2008). Back to social behavior: Mining the mundane. *Basic and Applied Social Psychology*, *30*(2), 93-101.

Pierro, A., Kruglanski, A. W., & Raven, B. H. (2012). Motivational underpinnings of social influence in work settings: Bases of social power and the need for cognitive closure. *European Journal of Social Psychology*, *42*(1), 41-52.

Pierro, A., Raven, B. H., Amato, C., & Bélanger, J. J. (2013). Bases of social power, leadership styles, and organizational commitment. *International Journal of Psychology, 48*(6), 1122-1134.

Podsakoff, P. M., & Podsakoff, N. P. (2019). Experimental designs in management and leadership research: Strengths, limitations, and recommendations for improving publishability. *The Leadership Quarterly*, *30*(1), 11-33.

Raue, M., Streicher, B., Lermer, E., & Frey, D. (2015). How far does it feel? Construal level and decisions under risk. *Journal of Applied Research in Memory and Cognition*, *4*(3), 256-264.

Raven, B. H. (1965). Social power and influence. In D. a. F. Steiner, M. (Ed.), *Current studies in social psychology* (pp. 371-382). New York, NY: Holt, Rinehart and Winston.

Raven, B. H. (1992). A power/interaction model of interpersonal influence: French and Raven thirty years later. *Journal of Social Behavior & Personality, 7*(2), 217-244.

Raven, B. H., Schwarzwald, J., & Koslowsky, M. (1998). Conceptualizing and measuring a power/interaction model of interpersonal influence 1. *Journal of Applied Social Psychology, 28*(4), 307-332.

Ronay, R., & Von Hippel, W. (2010). Power, testosterone, and risk-taking. *Journal of Behavioral Decision Making*, *23*(5), 473-482.

Rucker, D. D., Galinsky, A. D., & Dubois, D. (2012). Power and consumer behavior: How power shapes who and what consumers value. *Journal of Consumer Psychology*, *22*(3), 352-368.

Rucker, D. D., Hu, M., & Galinsky, A. D. (2014). The experience versus the expectations of power: A recipe for altering the effects of power on behavior. *Journal of Consumer Research*, *41*(2), 381-396.

Schaerer, M., Lee, A. J., Galinsky, A. D., & Thau, S. (2018). Contextualizing social power research within organizational behavior. In D. L. Ferris, R. E. Johnson, & C. Sedikides (Eds.), SIOP organizational frontiers series. *The self at work: Fundamental theory and research* (p. 194–221). Routledge-Taylor & Francis Group.

Schaerer, M., du Plessis, C., Yap, A. J., & Thau, S. (2018). Low power individuals in social power research: A quantitative review, theoretical framework, and empirical test. *Organizational Behavior and Human Decision Processes*, *149*, 73-96.

Schmid, P. C. (2018). Less power, greater conflict: Low power increases the experience of conflict in multiple goal settings. *Social Psychology, 49*(1), 47–62.

Schmid, P. C., & Schmid Mast, M. (2010). Mood effects on emotion recognition. *Motivation and Emotion*, *34*(3), 288-292.

Schmid, P. C., Schmid Mast, M., & Mast, F. W. (2015). Prioritizing—The task strategy of the powerful?. *The Quarterly Journal of Experimental Psychology*, *68*(10), 2097-2105.

Schmid Mast, M. (2001). Gender differences and similarities in dominance hierarchies in same-gender groups based on speaking time. *Sex Roles*, *44*(9-10), 537-556.

Schmid Mast, M. (2002). Dominance as expressed and inferred through speaking time: A meta-analysis. *Human Communication Research*, *28*(3), 420-450.

Schmid Mast, M. (2010). Interpersonal behavior and social perception in a hierarchy: The interpersonal power and behavior model. *European Review of Social Psychology*, *21*(1), 1-33.

Schmid Mast, M., Jonas, K., & Hall, J. A. (2009). Give a person power and he or she will show interpersonal sensitivity: The phenomenon and its why and when. *Journal of Personality and Social Psychology, 97*(5), 835–850.

Schmid Mast, M., Khademi, M., & Palese, T. (2019). Power and social information processing. *Current Opinion in Psychology, 33*, 42-46.

Sharpe, D., & Whelton, W. J. (2016). Frightened by an old scarecrow: The remarkable resilience of demand characteristics. *Review of General Psychology, 20*(4), 349-368.

Siu, R. G. H. (1979). *The craft of power*. New York: John Wiley & Sons.

Slabu, L., & Guinote, A. (2010). Getting what you want: Power increases the accessibility of active goals. *Journal of Experimental Social Psychology, 46*(2), 344-349.

Smith, P. K., & Bargh, J. A. (2008). Nonconscious effects of power on basic approach and avoidance tendencies. *Social Cognition*, *26*(1), 1-24.

Smith, P. K., & Magee, J. C. (2015). The interpersonal nature of power and status. *Current Opinion in Behavioral Sciences*, *3*, 152-156.

Sturm, R. E., & Antonakis, J. (2015). Interpersonal power: A review, critique, and research agenda. *Journal of Management*, *41*(1), 136-163.

Terken, J., & Sturm, J. (2010). Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing*, *14*(8), 703-714.

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.

Tost, L. P., Gino, F., & Larrick, R. P. (2013). When power makes others speechless: The negative impact of leader power on team performance. *Academy of Management Journal, 56*(5), 1465-1486.

Tost, L. P., & Johnson, H. H. (2019). The prosocial side of power: How structural power over subordinates can promote social responsibility. *Organizational Behavior and Human Decision Processes, 152*, 25-46.

Tuk, M. A., Verlegh, P. W., Smidts, A., & Wigboldus, D. H. (2019). You and I have nothing in common: The role of dissimilarity in interpersonal influence. *Organizational Behavior and Human Decision Processes, 151,* 49-60.

Van Dijke, M., & Poppe, M. (2006). Striving for personal power as a basis for social power dynamics. *European Journal of Social Psychology*, *36*(4), 537-556.

Weber, M. (1947). *The theory of economic and social organization*. New York: Oxford University Press.

Wegener, D. T., & Blankenship, K. L. (2007). Ecological validity. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 275–276). Thousand Oaks, CA: Sage.

Yukl, G., & Falbe, C. M. (1991). Importance of different power sources in downward and lateral relations. *Journal of Applied Psychology, 76*(3), 416-423.

Yukl, G., & Tracey, J. B. (1992). Consequences of influence tactics used with subordinates, peers, and the boss. *Journal of Applied Psychology, 77*(4), 525–535.

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

The Problem of Demand Effects in Power Studies: Moving Beyond Power Priming

Mahshid Khademi, Marianne Schmid Mast, Christian Zehnder, and Oriana De Saint Priest

Faculty of Business and Economics, University of Lausanne, Switzerland

Author Note

Mahshid Khademi, Marianne Schmid Mast, Christian Zehnder, and Oriana De Saint Priest, Department of Organizational Behavior, Faculty of Business and Economics, University of Lausanne, Switzerland.

Correspondence concerning this article should be addressed to Mahshid Khademi, Department of Organizational Behavior, University of Lausanne UNIL- Chamberonne, Internef, CH-1015 Lausanne, Switzerland. Phone number: +41 21 692 34 50 E-mail: Mahshid.Khademi@unil.ch

This work was supported by HEC Research Fund for PhD and Postdocs.

Abstract

Power in experimental research has been commonly induced by methods that raise concerns regarding demand effects. In this paper, we investigate the empirical relevance of these concerns. In an incentivized online study (N = 1632), we manipulated the method of power manipulation (power priming vs. resource allocation), the level of power (high-power vs. control), and the presence of a manipulation check after the power manipulation. We then assessed risk-taking as an outcome variable in two ways, once as a non-consequential measure (self-report measure) and twice as a consequential measure (incentivized behavioral choices). Results showed that using power priming (vs. resource allocation) and implementing a manipulation check increased the potential for demand effects as measured by the share of participants who were aware of the study hypothesis. Also, when a consequential measure was used to capture participants' risk-taking, the effect of power on risk-taking was shown to be non-significant. But, this effect became significant when risk-taking was measured non-consequentially using a self-report measure. Our pattern of results shows that concerns about demand effects in power studies cannot be dismissed. To avoid future discussions about demand-driven confounds, we advise researchers to steer away from demand-prone manipulations and to measure behavior through consequential choices.

Keywords: demand effects, power priming, resource allocation, manipulation check, consequential design

The Problem of Demand Effects in Power Studies: Moving Beyond Power Priming

Although leadership is often naturally associated with power, little research to date has explicitly addressed this link and its implications (e.g., Doldor, 2017; Bendahan, Zehnder, Pralong & Antonakis, 2015; Gordon, 2002). We argue that it is important for future research on leadership to consider the impact of power. Investigating how power affects individuals resembles how a person being promoted to a leadership position reacts. However, such parallels are only fruitful for the leadership literature if the methods used in the power field are sound. If there are methodological weaknesses in how the effect of power is investigated, the lessons learned for leadership are hampered. In this paper, we tackle the methodological issues in traditional power studies. The aim of our research is to provide evidence for the empirical relevance of these methodological concerns and to suggest more robust, alternatives methods to do research on power – and thus leadership—in the future.

Current methods used to test the effect of power in experimental settings are mainly constituted of power priming (Schaerer, Lee, Galinsky, Thau, 2018). Power priming aims at activating a high or low-power mindset in participants. For this purpose, participants are asked to put themselves in the shoes of either a high-power or a low-power person. Subsequently, they either act out this role in a role-play or imagine being that person. For instance, participants are assigned to either the role of a manager (high-power) or the role of a subordinate (low-power; e.g., Anderson & Berdahl, 2002; Briñol, Petty, & Stavraki, 2012) or participants imagine themselves in a predefined role (boss or employee) and write about the feelings and actions related to that role (Dubois, Rucker, & Galinsky, 2010; Schmid, 2018). The probably most widely used technique consists of asking participants to write about a situation in their past in which they had power over others (for high-power priming) or in which someone else had power over them (for low-power priming). In the power-neutral condition, participants write about an event that happened to them the day before (Galinsky,

Gruenfeld, & Magee, 2003). The idea behind this manipulation is that recollection of memories activates a power-related mindset that then affects behavior and decision-making.

Despite their popularity, using such power priming methods[1] might be problematic in that they potentially create *demand effects* (Sturm & Antonakis, 2015; Schaerer et al., 2018). Demand effects refer to experimenter-induced cues and expectations in the context of research experiments which may influence the behavior of participants (Orne, 1962). Such cues can, for example, be embedded in the description of the experiment (if instructions to participants reveal the research question) (Charness, Gneezy, & Kuhn, 2012; Zizzo, 2010) or may emerge if a manipulation check is performed before the dependent variable is measured (Chester & Lasko, 2019; Hauser, Ellsworth, & Gonzalez, 2018; Lonati, Quiroga, Zehnder, & Antonakis, 2018).

Power priming methods require the explicit mentioning of the manipulated variable (power) in the treatment conditions. Such manipulations may therefore reveal critical information about the hypothesis of the study and may induce participants to adapt their behavior to what they infer being the researcher's expectations. Demand-prone methods may yield confounded results because it remains unclear whether the ultimate responses are the result of what powerful or powerless people *actually* do or if they reflect what people think powerful or powerless people are *expected* to do (i.e., demand effects).

Another drawback of existing power priming methods is that, most of the time, being in a powerful or powerless position does not entail real consequences (Flynn, Gruenfeld, Molm, & Polzer, 2011). For instance, if the powerholder is in charge of evaluating the performance of the powerless or selecting a specific task for them, these choices typically neither have consequences for the powerholder nor for the followers. This lack of consequential outcomes implies that power priming does not manipulate "actual power" but rather "felt power". Such an indirect approach of manipulating power leads to two potential

complications. First, it is not clear whether actual power and felt power affect behavior in the same way. The latter might, in fact, operate through the perceived requirements of the power role reinforced by the potential for demand effects (Anderson & Berdahl, 2002). Second, the lack of an objective change in actual power requires testing whether the manipulation subjectively affected felt power in participants (participants in the high-power condition should report feeling more powerful than those in the low-power or neutral condition). In many studies using power priming (Galinsky et al. 2003; Guinote, 2007a; Guinote, 2007b; Smith, Jostmann, Galinsky, & van Dijk, 2008), the manipulation check is executed right after the power manipulation and before the dependent variable is measured. Such a procedure reinforces the potential for demand effects, because the manipulation check may provide strong hints on the manipulated variable.

Finally, there is an additional reason—unrelated to the use of power priming techniques—for why power studies may be particularly prone to demand effects. Power research is one of the fields that often uses self-report measures or hypothetical scenarios to measure outcomes of interest (e.g., participants are asked to describe what they would do in a particular situation). Such non-consequential operationalizations of the dependent variable are especially likely to be affected by what people think is expected of them in a given scenario, because deviations from true behavioral intentions have no costs for participants.

In a nutshell, we argue that using power priming and manipulation checks of this priming can create a large potential for demand effects. Moreover, non-consequential outcome measures make it likely that the demand-effect potential has an effect on results, because participants possibly respond according to demand. To the best of our knowledge, there is not yet any empirical work that explicitly measures the extent to which power priming and the corresponding manipulation checks affect the potential for demand effects in power studies and how more or less consequential measures of outcomes are affected by

different power manipulations. In this paper, we make a first attempt at filling this gap in the literature.

For this purpose, we compare an established and often used power priming method with an alternative power manipulation that directly affects actual power in a consequential way in that it provides powerholders with asymmetric control over valued resources, according to the definition widely used in research (e.g., Emerson, 1962; French & Raven, 1959; Goldstein & Hays, 2011; Gruenfeld, Inesi, Magee, & Galinsky, 2008; Jordan, Sivanathan, & Galinsky, 2011; Keltner, Gruenfeld, & Anderson, 2003; Magee & Galinsky, 2008; Rucker, Dubois, & Galinsky, 2011). Giving actual control over others' resources (Resource Allocation) as the power manipulation has two immediate advantages. First, we do not need to explicitly mention the manipulated variable. Second, because there is an objective difference in the degree of participant's control (i.e., an objective difference in actual power), our manipulation does not necessitate a subsequent manipulation check.

Moreover, we disentangle the effects of the power manipulation and the manipulation check using a 2 by 2 design in which we cross two different power manipulations (Power Priming and Resource Allocation) with the presence/absence of a manipulation check. In a first step, we investigate the potential for demand effects under each of the four conditions. To this end, we measure how many participants correctly guessed the hypothesis of the study. In a second step, we investigate the impact of the type of power manipulation and the presence/absence of a manipulation check on the link between power and a dependent variable that has been repeatedly studied in the power literature: risk-taking. In order to gain insight into the difference between non-consequential and consequential outcome variables, we use versions of both consequential and non-consequential measures of risk-taking.

Power Priming and Demand Effects

As mentioned before, power is typically defined as asymmetric control over valued resources (Magee & Galinsky, 2008). However, prevalent methods of power priming often give no actual control to powerholders and by manipulating felt instead of actual power, the chances of generating demand effects are increased (Antonakis, 2017; Sturm & Antonakis, 2015).

This is not the first time that power priming methods are under scrutiny. Certain power priming methods have been shown to produce spurious results. A well-known example is power posing (Carney, Cuddy, & Yap, 2010). Replication studies found no effect of striking a power pose on either cortisol or risk-taking, or any other dependent variable, except for self-reported felt power (Jonas et al., 2017; Ranehill et al., 2015). Moreover, doubts concerning the robustness of priming procedures in general have been voiced by many scholars (Chivers, 2019; Kahneman, 2012) as a result of failed attempts at replicating well-established findings in the social priming literature (e.g., Doyen, Klein, Pichon, & Cleeremans, 2012; Gilder & Heerey, 2018; Harris, Coburn, Rohrer, & Pashler, 2013; Pashler, Rohrer, & Harris, 2013; Shanks et al., 2013).

In power priming (e.g., role play of high- and low-power roles, imagine being a high- or low-power person, or think about a situation in the past in which a person had power or not), the goal is to put the person in a high- or low-power mindset, meaning having them *feel* powerful or powerless (or neutral with respect to power). The effect of the priming technic is rather unconscious and works through a mental representation of the primed concept (Bargh & Chartrand, 1999). In other words, the prime activates a cognitive network in the powerholder and thus several learnt associations in relation to the prime become available and affect an individual's feelings and behavior. For example, in power priming, the recalled memory about a power incident or the role play activates a cognitive network of power that brings about the emotional, cognitive, and behavioral consequences of being in power,

despite of the person primed not actually being granted objective power or control over valued resources (Tost, 2015).

Power priming is suspected to increase the potential for demand effects (Sturm & Antonakis, 2015; Schaerer et al., 2018). Power priming methods require that the experimenter mentions a power role or a power-related incident in the instructions. Participants are therefore confronted with salient cues about the manipulated variable. To the extent that participants have expectations about how a powerful or a powerless person ought to behave in a particular situation, these beliefs might guide how participants will react, behave, or decide in the subsequent task. This mechanism can be problematic because the observed effect of power on the behavior of interest is not driven by a change in actual objective power, but simply reflects what people think the effect should be (demand effect). We do not know whether people who are really given power would react in the same way.

We used an alternative power manipulation method, called the resource allocation (RA) manipulation. In this method, actual power is given to a person through granting control over valued resources in a structural setting. Individuals in positions of power are free to decide about valued resources for others. The freedom of choice as well as the control over the resource allocation for others are regarded as dimensions of power (Keltner et al., 2003; Magee & Smith, 2013).

Concretely, participants in our RA manipulation find themselves in groups and are randomly assigned the role of either an allocator, an observer, or a receiver. Each group has one allocator (high-power role), one observer (neutral role), and five receivers (low-power role). The allocator distributes a number of lottery tickets among the receivers and has absolute control over the distribution of the tickets (power) but cannot take any tickets for themselves. The observer is not involved in the lottery ticket distribution and only observes the distribution of the allocator. The observer has no power to disagree or change the

distribution that the allocator makes and has no gain or loss from the distribution. Both the allocator and the observer receive the same fixed payment for participation. The receivers also are compensated with a similar amount, but their expected payment also depends on how many tickets they receive.

The resource allocation method has several advantages. First, we use actual power as in the definition of power (asymmetric control over valued resources). Power is induced using a real endowment (i.e., no deception is used) which not only raises the ecological validity of the design, it also increases internal validity of the manipulation (correspondence between definition and operationalization of power) and it reduces demand effects because participants do not have to think about what might be expected from them in their power role, they simply act within the role. Also, there is nothing in the instructions that mentions power or any related concept and participants are thus not made aware of the hypothesis of the study. It is important to use neutral wordings because masking the dependent and independent variable lowers the chance of generating demand effects (de Quidt, Vesterlund, Wilson, 2018).

Note that demand effects cannot be eliminated entirely from an experimental design (Orne, 1962, 1969). Yet researchers are obliged to minimize the cues that convey the experimental hypothesis to their participants (Rosnow & Rosenthal, 1997), because participants tend to behave in accordance to what they believe is the hypothesis of the study, meaning how the independent variable affects or is related to the dependent variable (Nichols & Maner, 2008). If participants' beliefs match the hypothesis, the threat of demand effects is heightened (Zizzo, 2010); a positive correlation between the demand effects and the hypothesis is problematic because it will obscure the interpretation of the results (it is not clear whether the results are caused by the treatment or the demand). In order to understand and measure such potential for demand effects, we followed the traditional "post-

experimental inquiries" by simply asking participants what they thought was the purpose of the study (Orne, 1962).

In the current research, we compare our RA manipulation with one of the most widely used power priming methods (Galinsky, Rucker, & Magee, 2015; Gruenfeld et al., 2008; Guinote, 2007a; Lammers, Galinsky, Gordijn, & Otten, 2012; Schmid, Kleiman, & Amodio, 2015; Smith & Trope, 2006; Tost, Gino, & Larrick, 2013) in which participants are asked to write an essay about a time when they had power over somebody (high-power condition) or to write about a situation when someone else had power over them (low-power condition), or to write about what they did the day before (neutral condition) (Galinsky et al., 2003).

Hypothesis 1: *Fewer participants correctly guess the research hypothesis of the study (i.e., there is a lower potential for demand effects) under resource allocation (RA) than under power priming (PP).*

Manipulation Check and Demand Effects

Manipulation checks aim at ensuring that the manipulation was successful (Perdue & Summers, 1986; Kidd, 1976). Manipulation checks contribute to the construct validity of the study (Sigall & Mills, 1998), i.e., the extent to which the intervention operationalizes what it claims to operationalize (Cook, Campbell, & Shadish, 2002). In a survey, 75% of the scholars in social psychology think that "manipulation checks are necessary" in a well-designed study (Fayant, Sigall, Lemonnier, Retsin, & Alexopoulos, 2017, p. 127).

When actual power is manipulated (e.g., through the RA power manipulation), the need for a manipulation check will logically be removed (Lonati et al., 2018), because the objective changes in the elements of the design do not require a subjective interpretation from the participants. However, when using power priming, the activation of the power mindset can vary from one participant to the other, so that scholars are required to conduct a manipulation check to ensure that the relative subjectivity of the mindset priming goes in the

desired direction. Often, the check is performed right after the manipulation; this is commonly done by explicitly asking participants how "powerful", "in charge", or "in control" they felt after the essay writing task (Guinote, 2007a; Guinote, 2007b; Rucker et al., 2011; Smith et al., 2008).

When the manipulation check is implemented after the manipulation and before measuring the dependent variable, key information about the study may become salient to the participants (Kühnen, 2010) and the study hypothesis may become transparent (Ejelöv & Luke, 2020; Lonati et al., 2018). Participants might have converging ideas about how a person in power ought to behave, react, and respond, thus introducing demand effect. In other words, sometimes "what we call a manipulation check could also be a manipulation" (Hauser et al., 2018, p. 7). Participants' response to the manipulation check might affect the outcome variable, because the manipulation check impacts what participants think to be the expected behavior. Such reactivity to a manipulation check can inflate (or deflate) the true causal link between the manipulation and the outcome variable (Ejelöv & Luke, 2020).

Hypothesis 2: *More participants correctly guess the research hypothesis of the study (i.e., there is a higher potential for demand effects) if a manipulation check is performed (before measuring the dependent variable).*

Non-consequential Designs and Demand Effects

Non-consequential self-report measures are an appropriate tool to assess individual's emotions, perceptions, and beliefs about a phenomenon (Podsakoff & Organ, 1986). However, using non-consequential or hypothetical measures to measure behavior is often problematic for at least two reasons. First, non-consequential designs are generally weak in experimental realism (Colquitt, 2008; Podsakoff & Podsakoff, 2019). Therefore, the results of such designs tend to show the "behavioral intentions" in a situation instead of "real behavior" (Furr, 2009; Nave, Feeney, & Furr, 2018). Moreover, what individuals think they

would do in a situation, sometimes diverges tremendously from what they actually do in real situations (Baumeister, Vohs, & Funder, 2007). For example, in a famous field study, participants reported that they would be fairly generous in helping the victim of an accident (West & Brown, 1975). However, when those participants faced a victim in a seemingly real (but staged) situation, their actual helping behavior decreased significantly. It seems that when nothing or little is at stake (hypothetical situations), individuals often tend to report that they would engage in socially desirable behavior—behavior that they might display to a lesser degree or not at all in reality.

Second, when the treatment entails indications or cues (explicitly or implicitly) about the appropriate/expected behavior, non-consequential designs can ultimately lead participants to behave in accordance with experimental demand (Antonakis, 2017). In non-consequential designs, taking action to please the experimenter is not costly (Podsakoff, MacKenzie, & Podsakoff, 2012), entailing considerable room for demand effects (e.g., Durgin et al., 2009). In consequential settings, in contrast, it is costly for participants to not pick their preferred choice in order to comply with the demand (de Quidt, Vesterlund, & Wilson, 2019; Lonati et al., 2018). In this sense, the use of non-consequential outcome variables makes it more likely that a demand effect results in a confound of the actually observed effect.

In our study, we used one non-consequential and two consequential assessments of risk-taking as outcome measures. We selected "risk-taking" as our variable of interest, because previous research has found a strong relation between power and risk-taking (Anderson & Galinsky, 2006; Galinsky et al., 2003; Jordan et al., 2011; Lewellyn & Muller-Kahle, 2012; Maner, Gailliot, Butz, & Peruche, 2007; Ronay & Von Hippel, 2010). When choosing the risk-taking tasks, we pursued two goals. First, we wanted to use what is commonly used in the literature and second, we wanted to use measures of which we thought they were differentially susceptible to demand. This is why we chose a self-report risk taking

task—highly affectable by demand—and two incentivized tasks: the Balloon Analog Risk Task (BART) (Lejuez et al. 2002; Lejuez et al. 2003), and the lottery task (Dohmen et al., 2011). The BART task is used rather in psychology and in the neurosciences (Lewellyn & Muller-Kahle, 2012; Jordan et al., 2011) whereas the lottery task is rather used in economics. In both of the consequential measures of risk-taking, participants were compensated based on their choices in risky settings. This broad range of tasks should allow us to detect a link between power and risk-taking if there is one because the tasks are different and may tap into different aspects of risk-taking.

Hypothesis 3: *Using methods that are hypothesized to generate a high potential for demand effects (power priming and manipulation checks) strengthens the observed positive relation between power and risk-taking, in particular if risk-taking is measured non-consequentially.*

Method

Pre-test

Although we detailed that when using a manipulation of actual power instead of a power priming procedure, there is no need for checking whether subjectively, power is experienced differently, we performed a pre-test to measure the impact of both power manipulations, PP and RA, on participants' felt power. The reason for doing so is related to our setting: we need both manipulations to be comparable in strength so that if we find effects, they cannot be interpreted as stemming from a difference in the strength of manipulating felt power. In other words, if we find differences between the effects of the two power manipulations on risk-taking and one manipulation is substantially stronger with respect to felt power, we do not know whether this difference is driven by the manipulation type—which is what we are interested in—or simply by the strength of the manipulation. We

therefore needed to ensure that the PP and the RA methods resulted in similar effect sizes on felt power.

Phase one. The goal of phase one of the pretest was to investigate how different levels (high, neutral, low) of each manipulation affect reports of felt power by participants. We used a 2 (power manipulation method: PP vs. RA) by 3 (power level: high-power, neutral, and low-power) design with random assignment to one of the 6 experimental conditions. We recruited 323 participants from MTurk (https://www.mturk.com) and excluded 31 observations due to attention check failure,[2] resulting in a sample of 292 individuals, 47.95% female; $M_{age} = 33.48$, $SD_{age} = 9.16$.

Power priming (PP). Following Galinsky et al. (2003), participants in the high-power condition were asked to recall an incident in which they had power over another individual or individuals and write about it. In the low-power condition, participants were asked to write about an incident in which someone else had power over them and in the neutral condition, participants were invited to write about the last time they went to the grocery store. We followed the procedure described in Galinsky et al. (p. 458).

Resource allocation manipulation (RA). Participants were assigned to groups of seven individuals. There were three roles in each group: allocator (high-power), observer (neutral), and receiver (low-power). There were 1 allocator, 1 observer, and 5 receivers in each group. The roles were distributed randomly among participants. The allocator was given five lottery tickets to distribute among the receivers. The allocator had full control over how to distribute the 5 tickets among the receivers. The only restrictions were that the allocator was not allowed to keep any lottery tickets for him/herself and that all the tickets needed to be distributed. Before the distribution was made, the allocator was informed about the gender and age of the receivers (e.g., receiver A: female, age 25-30; receiver B: male, age 20-25 etc.). The participant in the role of the observer (neutral power) was not involved in the

distribution phase and only received the information about the allocator's distribution. The observer could not take any action to change the allocator's decision and had no gain or loss from the distribution. Both the allocator and observer received the same amount of money ($3) for their participation in the study. The receivers also gained the same exact fixed payment ($3). In addition, receivers were paid based on the lottery ticket distribution. A receiver with one lottery ticket had a 1 in 10 chance to win 1 additional dollar. A random device selected 25 out of the 250 distributed lottery tickets and the final gains were paid to participants accordingly. The participants in the different roles were not informed about the final payments of participants in the other roles to avoid salient comparisons.

Felt power. After the power manipulation, participants answered one question about how powerful they felt using a Likert scale from 1 (not at all) to 7 (very much) (M = 3.77, SD = 1.94). In the PP condition, high-power participants felt more powerful (M = 5.16, SD = 1.68) than participants in the neutral condition (M = 3.94, SD = 1.59), $F(1,286) = 13.54$, $p < .001$, Cohen's d = 0.75, and than participants in the low-power condition (M = 2.56, SD = 1.49), $F(1,286) = 65.73$, $p < .001$, Cohen's d = 1.63. In the RA condition, high-power participants felt more powerful (M = 5.38, SD = 1.27) than participants in the neutral condition (M = 2.71, SD = 1.79), $F(1,286) = 71.63$, $p < .001$, Cohen's d = 1.71, and than participants in the low-power condition (M = 2.88, SD = 1.60), $F(1,286) = 73.66$, $p < .001$, Cohen's d = 1.73 (see Figure 1).[3]

To assure that the two types of power manipulation have comparable strength, we selected for the PP the high-power versus low-power condition (d = 1.63) and for the RA the high-power versus neutral condition (d = 1.71) for further pretesting and for the main study. There were two reasons why we selected the neutral condition over the low-power condition for the RA method: 1) to have a relatively similar effect size compared to the PP manipulation method and thus to avoid the possibility that the PP manipulation leads to a

weaker effect on risk-taking because of a smaller effect size; 2) to avoid income effects between the conditions chosen in the RA method. The allocators and observers obtained the exact same payment by design, whereas the receivers' income depends partially on the number of lottery tickets received from the allocators and partially on chance.

In sum, to conduct the next phase of the pre-test and the main experiment, for the PP condition we kept the high and low-power conditions and for the RA method, we used the high-power and the neutral condition. Note that from now on, we simply call the group to which the high-power individuals are compared, the *control group* because being in the control group elicits a comparable amount of feeling relatively less powerful than being in the high-power group for both types of power manipulations.

Phase two. In this second phase, our main objective was to confirm that the two manipulations selected in phase one indeed have comparable impacts on felt power using a larger sample of participants. In addition, we also measured how transparent each of the two manipulations was to participants. The transparency of the manipulation is an important pre-requisite for the emergence of a potential for demand effects. If the participants understand what variable the study manipulates, they may be able to correctly guess the hypothesis that the researcher investigates.

On the same online platform, 746 participants were recruited. We excluded 140 participants due to lack of correct responses to the attention checks (same as in phase one). The final sample consisted of 606 participants, 41.42% female; M_{age} = 36.13, SD_{age} = 10.86.

We randomly assigned participants to one of the experimental conditions in a 2 (method of power manipulation: PP vs. RA) by 2 (level of power manipulation: high-power vs. control) between-subjects design. We used the same item as in phase one to measure for how powerful they felt. In the PP condition, high-power participants felt more powerful (M = 5.08, SD = 1.54) than participants in the control condition (M = 2.97, SD = 1.73), $F(1,602)$ =

125.21, $p < .001$, Cohen's d = 1.29. In the RA condition, high-power participants (allocators) felt more powerful ($M = 5.24$, $SD = 1.38$) than participants in the control condition (observers) ($M = 2.56$, $SD = 1.69$), $F(1,602) = 228.10$, $p < .001$, Cohen's d = 1.73 (Figure 2). Observed effect sizes were similar to those identified in phase one. The effect size in the RA condition was slightly larger than in the PP condition. However, if anything, having a stronger RA manipulation works against our predictions, because we hypothesized that the RA power manipulation would have a *smaller* effect on risk-taking than the PP manipulation, (see Hypothesis 3).

To measure the transparency of the manipulation for participants we asked whether they could guess what the study was about in a free-form question. We coded their answers as 1 when it mentioned power or control (e.g., being in control, being in charge, having any saying in a situation) and 0 otherwise. Participants in the PP correctly guessed the manipulated variable more often ($M = 0.44$, $SD = 0.50$) compared to participants in the RA condition ($M = 0.15$, $SD = 0.36$), t (604) = 8.09, $p < 0.001$. The fact that the manipulation is more transparent does not necessarily imply a larger potential for demand effects because participants may still fail to correctly guess the research hypothesis. However, a transparent manipulation is a necessary predisposition for the potential presence of a demand effect.

Main Experiment

Participants. We recruited 1752 individuals from the same online platform (MTurk). We used the same attention check questions and exclusion criteria. Additionally, we checked the content of the stories written in the PP condition and excluded stories that were either 1) non-relevant to power situations, 2) included pieces of text copy-pasted from the internet, or 3) nonsense responses (Rinderknecht, 2019). This resulted in excluding 119 observations with a final sample consisting of 1633 individuals (43.11% female; $M_{age} = 35.80$, $SD_{age} = 10.75$). Participants received a fixed amount of $3. They could gain an additional $10 based

on some random elements in the protocol of the risk-taking tasks, described in more detail below.[4]

Procedure and measures. Participants were randomly assigned to one condition of a 2 (power manipulation method: PP vs. RA) by 2 (level of power: high-power vs. control) by 2 (manipulation check: with vs. without) experimental design. In the first step, we manipulated power using either the PP or the RA power manipulation method described above. We then randomly assigned participants to either the condition with or without manipulation check and participants in the "with manipulation check" condition reported how powerful they felt. We then assessed participants' level of risk-taking with different measures. There was a non-consequential way of measuring risk-taking by using a self-report questionnaire and there were two consequential ways of measuring risk-taking. Participants faced always the same sequence of the risk-taking tasks. After this, we measured the potential for demand effects by asking participants if they could guess the hypothesis of the study.

Power manipulation methods. We used the PP and the RA methods manipulating two levels of power: high-power and control. It is important to recall the main differences between the two methods. PP does not change actual power, but aims at affecting participants' power mindset (i.e., felt power) by inducing them to remember and describe a situation in which they found themselves in a position of power or in a powerless state (depending on the condition). RA directly manipulates actual power by changing participants' impact on incomes of others. Participants in the high-power condition can choose how many lottery tickets other participants get and participants in the control condition observe the allocation choices of the high-power people.

Manipulation check. We either used a manipulation check right after the power manipulation or not. Those who were assigned to the "with manipulation check" condition were asked how powerful they felt with the following question: **"**After doing this task, please

indicate how powerful you feel now on a 7-point Likert scale ranging from 1 (not at all) to 7 (very much)? “ In the condition without manipulation check, the participants simply advanced to the next steps of the study.

Self-reported risk-taking. As the non-consequential measurement of risk-taking, we used a self-report measure. Self-reported risk-taking has shown to be a good predictor of risk-taking tendencies in various domains such as job choices and portfolio selections (Dohmen et al., 2011). We asked participants “How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?” (Pannenberg, 2010). They had to rate their risk-taking from 1 (unwilling to take risks) to 7 (fully prepared to take risks). Higher scores indicate higher risk-taking (M = 4.36; SD = 1.63).

BART risk-taking. We also measured risk-taking behavior in a consequential way, using a behavioral measure developed by Lejuez et al. (2002). In the BART task, participants are faced with several rounds of risky decision-making. Participants had to “inflate” 10 (digital) balloons. More specifically, they saw a balloon that they needed to inflate with a pump. Each air blast that they added improved their score by 1 point, but it also increased the risk that the balloon would burst. Participants could inflate the balloon as much as they liked and could collect the points at any time. However, when the balloon burst, all the points accumulated in that specific round were lost. Bursting the balloon did not have an effect on the collected points in former rounds. There were 10 rounds of balloon inflating and the threshold for bursting was different for each balloon. The final payoff was based on the overall collected points (100 points = $1) and participants could gain up to $3.68 in this task. Past research has shown that the points collected via this measure have a significant and reliable relation with risk-taking in various domains including minor crimes, substance use, risky sexual behavior, and unhealthy behavior (Lejuez et al., 2007). People who inflate

balloons more, show more risk-taking behavior. We used the average number of air blasts added across all rounds as the BART risk-taking behavior measure (M = 14.76; SD = 7.29).

Lottery risk-taking. As another consequential way of measuring risk-taking, we used a paid lottery task (Dohmen et al., 2011). In this task, participants were presented with a table of 20 rows. In each row participants could either choose a safe option (a fixed amount of money) or play a lottery (win $0 or $7 with 0.5 probability). The lottery remained unchanged in all rows (both the probability and the amounts stayed the same) but the safe option increased by $0.25 from row to row. So, for example, in the first row the participant had to select between a safe option of $0 and a lottery in which they could win $0 or $7 with 0.5 probability. In the second row, the participants had to choose between a safe option of $0.25 and the lottery (winning $0 or $7 with 0.5 probability). This continued up to the last (twentieth) row, where the participant had to select between a safe option of $4.75 and the lottery (winning $0 or $7 with 0.5 probability). In the first rows the safe option is very unattractive relative to the lottery. Accordingly, most participants decided to play the lottery. At some point, however, when the safe option increases and approaches the expected value of the lottery ($3.5), participants start to switch to the safe option. We use this switching point (i.e., the lowest amount at which participants prefer the safe option) as the measure of risk-taking in this task.

Participants were informed that for 1 out of 7 participants, one row would be randomly selected and the participant's decision in the selected row would be implemented and paid out. Participants were therefore able to gain up to $7. Some participants exhibited choice patterns with multiple switching points. These choices are most likely caused by confusion and are hard to interpret. We therefore decided to exclude these 235 participants from the analysis for this task (but they remained in our sample for all other analyses).

***Potential for demand effects*.** We define the potential for demand effects as the share of participants who correctly guess the research hypothesis of the study (i.e., identify the correct independent and dependent variables and correctly state the direction of their relation). We measured the potential for demand effects in two ways.

First, we asked the participants a very open, free-form question: "*What is your guess about the aim of this study? We are interested in your personal opinion*". The variable was scored 1 if participants mentioned anything related to power or risk-taking and 0 otherwise. We did not incentivize the answers. This unstructured measure (UM) of the potential for demand effects is very conservative in the sense that it puts no pressure on participants to think hard about their answer and people were free to respond whatever they wanted (including evasive or neutral answers such as "I don't know"). We therefore see the data collected with this first measure as a lower bound for the true potential for demand effects, because it is likely that not all participants who (at least intuitively) understood the hypothesis actually also reported it.

Second, participants were faced with a multiple-choice question. They were asked to identify the relation investigated in the study. They had to select one single relation among 18 possible options presented to them. Examples of suggested relations were: "The relation between stereotype and self-control", "The relation between power and corruption", or "The relation between power and risk-taking". Participants were informed that if they selected the correct relation, they would gain a bonus of $0.1. Then, participants were asked: "Please tell us how you think the relation is? What could be the hypothesis (e.g., I think X increases/decreases Y*)."* Their answers were coded as 1 if participants mentioned the correct hypothesis, "*power increases risk-taking*" and 0 otherwise). This structured measure (SM) of the potential for demand effects is more "invasive" in the sense that the monetary incentives might induce participants to think much harder about the study hypothesis than they would

have done otherwise. Moreover, the structured question also forces the participants to think about the study in terms of a hypothesis. We therefore see the data collected with this second method as an upper bound for the true potential for demand effects, because the method might not only have induced correct answers from those who had actively thought about the hypothesis of the study before, but also from those whose understanding was rather unconscious.

Results

We first focus on the potential for demand effects. As expected, overall more participants correctly guessed the study hypothesis when we elicited the potential for demand effects with the structured measure (SM: $M = 0.29$, $SD = 0.46$) than when we used the unstructured measure (UM: $M = 0.07$, $SD = 0.25$). To test how the different power manipulations and the manipulation check influence the potential for demand effects, we conducted 2 (power manipulation method: RA vs. PP) by 2 (manipulation check: with vs. without) ANOVAs with each of our two measures of the potential for demand effects (SM and UM) as the dependent variable (Table 1 and Figure 3). As predicted in Hypothesis 1, the power manipulation method influenced the potential for demand effects: comparing across the two conditions with RA and the two conditions with PP (main effect of power manipulation method), we observe that using the RA method reduced the potential for demand effects (SM: $M = 0.20$, $SD = 0.40$/UM: $M = 0.03$, $SD = 0.17$) in comparison to the PP method (SM: $M = 0.39$, $SD = 0.49$, $F(1,1629) = 86.39$, $p < 0.001$/UM: $M = 0.11$, $SD = 0.32$, $F(1,1629) = 46.75$, $p < 0.001$).

Moreover, as predicted in Hypothesis 2, comparing across the two conditions without and the two conditions with a manipulation check (main effect of manipulation check), we find that using a manipulation check increased the potential for demand effects (SM: $M = 0.39$, $SD = 0.49$/UM: $M = 0.09$, $SD = 0.28$) in comparison to no manipulation check (SM: *M*

= 0.20, *SD* = 0.40), $F(1,1629) = 79.43$, $p < 0.001$/UM: *M* = 0.20, *SD* = 0.40), $F(1,1629) = 7.91$, $p = .005$).[5]

The interaction effect between the use of the PP manipulation and the presence of a manipulation check is positive for both our measures of the potential for demand effects, but it is only significant for the structured measure (SM: $F(1,1629) = 7.30$, $p < . 01$/UM: $F(1,1629) = 0.94$, $p = .33$).

Figure 3 shows the full comparison of all four conditions with regard to the potential for demand effects (both measures are displayed). As expected, participants in the least demand-prone condition (RA without manipulation check) correctly guessed the hypothesis of the study less often (SM: *M* = 0.07, *SD* = 0.26/UM: *M* = 0.005, *SD* = 0.07) than participants in any of the other conditions: Pair-wise contrasts confirm that the potential for demand effects is significantly lower under RA without manipulation check than in any of the three other conditions (this result holds irrespective of the method used to measure the potential for demand effects).[6] Moreover, Figure 3 also reveals how the two suspected sources—a demand-prone power manipulation method and the existence of a manipulation check—contribute to the increase in the potential for demand effects. Both our measures indicate that the potential for demand effects is highest if PP is combined with a manipulation check (SM: *M* = 0.46, *SD* = 0.50/UM: *M* = 0.12, *SD* = 0.33). Concerning the structured measure, the potential for demand effects in this condition is significantly higher than in PP without a manipulation check ($F(1,1629) = 18.55$, $p < .001$) and in RA with a manipulation check ($F(1,1629) = 21.70$, $p < .001$) and there is no significant difference between the two latter conditions ($F(1,1629) = 0.07$, $p = .79$). Concerning the unstructured measure, the combination of PP and manipulation check increases the potential for demand effects significantly, relative to RA with manipulation check ($F(1,1629) = 17.18$, $p < .001$), but not relative to PP without manipulation check ($F(1,1629) = 1.63$, $p = .20$). However, the potential

for demand effects is higher under PP without manipulation check than under RA with manipulation check ($F(1,1629) = 8.17$, $p = .004$).

Hypothesis 3 predicted that demand-prone methods would produce larger effects of power on risk-taking, in particular if risk-taking is measured non-consequentially. We first focused on the non-consequential measure of risk-taking (the self-report measure). We calculated a 2 (power manipulation method: RA vs. PP) by 2 (manipulation check: with vs. without) by 2 (level of power: high vs. low) ANOVA with the self-report measure of risk-taking (non-consequential measure) as the dependent variable (Table 2 and Figure 4). Results showed a significant main effect of power, $F(1, 1625) = 4.47$, $p = .03$,[7] meaning that overall, participants in high-power conditions reported themselves as being more risk-taking ($M = 4.44$, $SD = 1.62$) compared to participants in all control groups together ($M = 4.28$, $SD = 1.63$). None of the other main effects nor the interaction effects were significant.

Nevertheless, to better understand and compare the main effect of power across the different conditions, we conducted contrast analyses. Although non-significant, the results of the simple contrast analysis comparing the high-power condition to control suggests that the effect of power on risk-taking tends to be stronger in the demand-prone conditions (PP with manipulation check, $F(1, 1625) = 1.32$, $p = .25$, Cohen's d = 0.12; PP without manipulation check, $F(1, 1625) = 2.32$, $p = .13$, Cohen's d = 0.15; and RA with manipulation check, $F(1, 1625) = 1.50$, $p = .22$, Cohen's d = 0.12) in comparison to the low demand-prone condition (RA without manipulation check, $F(1, 1625) = 0.09$, $p = .76$, Cohen's d = 0.03).

As shown above, taking all high-power conditions together and comparing them to all low-power conditions showed a significant effect of power on self-reported risk-taking (main effect of power level). Interestingly, excluding the low demand-prone condition (RA without manipulation check) from this contrast made the result of power on risk-taking even stronger, $F(1, 1625) = 5.11$, $p = .02$. Whereas, excluding either of the high demand-prone conditions

from the contrast analysis made the effect of power non-significant at the 5% level. For example, the contrast becomes non-significant if we either exclude the PP with manipulation check, $F(1, 1625) = 3.23$, $p = .07$, or exclude the PP without manipulation check, $F(1, 1625) = 2.42$, $p = .12$, or exclude the RA with manipulation check , $F(1, 1625) = 2.93$, $p = .09$. This pattern of results suggests that the overall effect of power on (non-consequential) risk-taking is mostly driven by the high demand-prone conditions, but the differences in effect sizes across conditions are too small to be significant.[8]

In the next step, we looked at the consequential measures of risk-taking. We used two consequential methods to measure risk-taking behavior: the BART and lottery risk-taking task. We calculated a 2 (power manipulation method: RA vs. PP) by 2 (manipulation check: with vs. without) by 2 (level of power: high vs. low) ANOVA once with the BART risk-taking task as the dependent variable and once with the lottery risk-taking task.[9] We found neither a main nor an interaction effect of power on the consequential measures of risk-taking (Tables 3 and 4, Figures 5 and 6).[10] As implied by Hypothesis 3, when using consequential ways of measuring risk-taking there was no significant effect of power on risk-taking.

Discussion

In this paper, we made a first step in addressing the important concern that using power priming (PP) might lead to demand-driven results. In an incentivized experimental design we showed that using the PP manipulation method and/or adding a manipulation check significantly increased the potential for demand effects compared to using the resource allocation (RA) method. We used two measures to assess the potential for demand effects. Our unstructured measure (UM) indicates that the share of participants who correctly guess the hypothesis of the study increases from 0.5% to 12% when using PP in combination with a manipulation check instead of using RA. One might be tempted to argue that these numbers imply that the potential for demand effects is rather low in general. However, it is important

to keep in mind that these levels need to be regarded as a lower bound, because they most likely underestimate the true relevance of the problem. The UM was not incentivized and provided very little guidance on how to formulate what the aim of the study might have been. It is therefore likely that not all participants who were aware of the study hypothesis were identified by the UM. Our second measure, the structured measure (SM), indicates an increase from 7% to 46% in the potential for demand effects when moving from the RA without a manipulation check to the PP with a manipulation check. These numbers reveal that the magnitude of the problem might be much larger than indicated by the conservative UM. We acknowledge that the levels suggested by the SM should probably be seen as an upper bound, because the financial incentives and the pre-structured answers might have induced some participants to think more actively and possibly in a different manner about the hypothesis of the study than in the absence of such a measure. However, it is important to emphasize that the SM does not inflate the measured potential for demand effects by construction, because it remains at a low level for the not demand-prone manipulation (RA without manipulation check). Taken together our results show that the concern that PP creates a large potential for demand effects cannot be dismissed.

When we investigated the effect of power on risk-taking for non-consequential and consequential ways of assessing risk-taking, there is only weak evidence for an effect of power on risk-taking and it only occurs for the non-consequential risk-taking measure and more so in the demand-prone conditions. When risk-taking is measured in an incentivized manner, we do not find any significant effects of power on risk-taking. These results indicate that demand-prone methods might lead to upward biased effect sizes especially when they are combined with explicit manipulation checks and non-consequential measures of the behavioral outcome.

Despite the low implementation cost of the PP method, the potentially high threat of demand effects suggests that power researchers should use alternative methods to manipulate power, especially when they want to predict behavior of people with actual power. The here suggested RA method strongly reduces the potential for demand effects and has the advantage to induce actual power (i.e., having actual objective control over valued resources). This also increases construct validity of the power manipulation.

Note, however, that the RA method only produced a very low potential for demand effects when it was not followed by a manipulation check. This underlines the importance of paying attention to different elements of the experimental design that can reveal the hypothesis of the study to the participants (e.g., demand-prone method as well as the manipulation checks). In the RA method, the control over valued resources and the decision on the outcomes of others grant objective power to powerholders (as opposed to the subjective power in the priming method) and thus removes the need for a manipulation check. Because power is manipulated subjectively in the PP method, it is recommended to implement manipulation checks to examine the effect of the manipulation (Sigall & Mills, 1998). However, in such cases the researcher should either conduct the manipulation in a separate (pre-)study or add enough filler items to the manipulation check scale in order to conceal the research goal or topic (Ejelöv & Luke, 2020). Ideally, one should use an independent but comparable sample to check the effect of the manipulation (Hauser et al., 2018; Fayant et al., 2017; Kidd, 1976).

Our research also highlights the importance of using consequential designs. Usually what we expect from behavioral studies is that they deal with individuals' behavior. But, how many researchers in the behavioral disciplines really use behavioral designs? Baumeister et al. (2007) believes very few: the percentage of behavioral research has dropped from 80% in 1976 to less than 20% in 2006. What we observe is a decline in behavioral research which

goes hand in hand with an increase in hypothetical and self-report procedures (Patterson, 2008; Patterson, Giles & Teske, 2011). Note that we are not against using self-reports and hypothetical designs. Self-report measures matter in many contexts, for example, when the researchers are interested in the emotion or perception of the individuals. However, they should not replace measuring actual behavior in real situations. Our results added to this discussion by pointing out that self-report measures of behavior are especially vulnerable to the effects of demand. We showed that consequential designs are more immune to the threat that demand-prone conditions pose.

There are a number of possible explanations for why the effect of power on risk-taking was not replicated in the consequential designs using the RA manipulation. Despite the large number of observations in this study and the correlations among the risk-taking measures[11], we still did not find an effect of power on risk-taking behavior. Based on a post-hoc power analysis (Faul, Erdfelder, Lang, & Buchner, 2007), assuming an alpha level of .05, and a small effect size of 0.2, by having 400 observations per cell we had a statistical power of 0.88 to detect the effect of power on risk-taking with our setting. Yet, we still could not replicate the effect. Scrutinizing the studies promoting the power/risk-taking effect, we noticed that many used self-report measures or hypothetical designs to capture risk-taking (e.g., see Studies 1 to 5, Anderson & Galinsky 2006). Moreover, the main effect of power on risk-taking was not replicated in the context of consequential designs in several studies conducted by other researchers (see Studies 3 and 5, Jordan et al., 2011; Hiemer & Abele, 2012; Maner et al., 2007; Ronay & Von Hippel, 2010). Therefore, a possible explanation could be that there is no effect of power on risk-taking after all. Future replication studies on power and risk-taking may want to clarify this state of affairs and consider the effect size in the presence and absence of consequential designs.

Our research addresses concerns raised about the potential demand characteristics in power research and the pervasive use of the PP method (Lonati et al., 2018; Sturm & Antonakis, 2015; Schaerer et al., 2018), as well as about the problems of using manipulation checks before measuring the dependent variable (Ejelöv & Luke, 2020; Fayant et al., 2017; Hauser et al., 2018).

We also contribute to the literature discussing experimenter demand effects (de Quidt et al., 2018; de Quidt et al., 2019; Nichols & Maner, 2008; Zizzo, 2010). Our results highlight the importance of implementing rigorous research designs. We suggest researchers to be prudent about revealing possible clues either through the manipulation method or via manipulation checks because otherwise, with a little help from demand characteristics, researchers may oftentimes find what they hypothesized even if the effect is not really there.

References

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.

Anderson, C., & Galinsky, A. D. (2006). Power, optimism, and risk-taking. *European Journal of Social Psychology, 36*(4), 511–536.

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

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*(7), 462–479.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: whatever happened to actual behavior? *Perspectives on Psychological Science, 2*(4), 396–403.

Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. *The Leadership Quarterly, 26*(2), 101–122.

Briñol, P., Petty, R. E., & Stavraki, M. (2012). Power increases the reliance on first-impression thoughts. *Revista de Psicología Social, 27*(3), 293–303.

Carney, D. R., Cuddy, A. J., & Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological Science, 21*(10), 1363–1368.

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization, 81*(1), 1–8.

Chester, D., & Lasko, E. (2019). *Construct validation of experimental manipulations in social psychology: Current practices and recommendations for the future*. Retrieved from: https://doi.org/10.31234/osf.io/t7ev9

Chivers, T. (2019). What's next for psychology's embattled field of social priming. *Nature, 576*, 200–202.

Colquitt, J.A. (2008). From the editors: Publishing laboratory research in AMJ: A question of when, not if. *Academy of Management Journal*, *51*(4), 616–620.

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

de Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, *108*(11), 3266–3302.

de Quidt, J., Vesterlund, L., & Wilson, A. J. (2019). Experimenter demand effects. In A. Schram & A. Ule (Eds.), *Handbook of research methods and applications in experimental economics* (pp. 384–400). Cheltenham, UK: Edward Elgar Publishing.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, *9*(3), 522–550.

Doldor, E. (2017). From politically naïve to politically mature: Conceptualizing leaders' political maturation journey. *British Journal of Management, 28*(4), 666-686.

Doyen, S., Klein, O., Pichon, C. L., & Cleeremans, A. (2012). Behavioral priming: It's all in the mind, but whose mind?. *PLOS One, 7*(1), 1–7.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2010). The accentuation bias: Money literally looms larger (and sometimes smaller) to the powerless. *Social Psychological and Personality Science*, *1*(3), 199–205.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., & Waymouth, S. (2009). Who is being deceived? The experimental demands of wearing a backpack. *Psychonomic Bulletin & Review*, *16*(5), 964–969.

Ejelöv, E., & Luke, T. J. (2020). "Rarely safe to assume": Evaluating the use and interpretation of manipulation checks in experimental social psychology. *Journal of Experimental Social Psychology*, *87*, 1–13.

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review*, *27*(1), 31–41.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fayant, M. P., Sigall, H., Lemonnier, A., Retsin, E., & Alexopoulos, T. (2017). On the limitations of manipulation checks: An obstacle toward cumulative science. *International Review of Social Psychology*, *30*(1), 125–130.

Flynn, F. J., Gruenfeld, D., Molm, L. D., & Polzer, J. T. (2011). Social psychological perspectives on power in organizations. *Administrative Science Quarterly*, *56*(4), 495–500.

French, J. R. P., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150–167). Ann Arbor, MI: University of Michigan Press.

Furr, R. M. (2009). Personality psychology as a truly behavioral science. *European Journal of Personality*, *23*(5), 369–401.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453–466.

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, pp. 421–460). Washington, DC: American Psychological Association.

Gilder, T. S., & Heerey, E. A. (2018). The role of experimenter belief in social priming. *Psychological Science, 29*(3), 403–417.

Goldstein, N. J., & Hays, N. A. (2011). Illusory power transference: The vicarious experience of power. *Administrative Science Quarterly*, *56*(4), 593–621.

Gordon, R. (2011). Leadership and power. In A. Bryman, D. L. Collinson, K., Grint, B., Jackson, M., Uhl-Bien (Eds.), *The Sage Handbook of Leadership* (pp. 195–202). London: Sage.

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.

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

Guinote, A. (2007b). Behavior variability and the situated focus theory of power. *European Review of Social Psychology, 18*(1), 256–295.

Harris, C. R., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PLOS One, 8*(8), 1–9.

Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are manipulation checks necessary?. *Frontiers in Psychology*, *9*(998), 1–10.

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.

Jonas, K. J., Cesario, J., Alger, M., Bailey, A. H., Bombari, D., Carney, D., ... Jackson, B. (2017). Power poses–where do we stand?. *Comprehensive Results in Social Psychology, 2*(1), 139–141.

Jordan, J., Sivanathan, N., & Galinsky, A. D. (2011). Something to lose and nothing to gain: The role of stress in the interactive effect of power and stability on risk taking. *Administrative Science Quarterly, 56*(4), 530–558.

Kahneman, D. (2012, September 26). A proposal to deal with questions about priming effects. Retrieved from https://www.nature.com/news/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review, 110*(2), 265–284.

Kidd, R. F. (1976). Manipulation checks: Advantage or disadvantage?. *Representative Research in Social Psychology, 7*(2), 160–165.

Kühnen, U. (2010). Manipulation checks as manipulation: Another look at the ease-of-retrieval heuristic. *Personality and Social Psychology Bulletin, 36*(1), 47–58.

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2012). Power increases social distance. *Social Psychological and Personality Science, 3*(3), 282–290.

Lejuez, C. W., Aklin, W., Daughters, S., Zvolensky, M., Kahler, C., & Gwadz, M. (2007). Reliability and validity of the youth version of the Balloon Analogue Risk Task (BART–Y) in the assessment of risk-taking behavior among inner-city adolescents. *Journal of Clinical Child and Adolescent Psychology, 36*(1), 106–111.

Lejuez, C. W., Aklin, W. M., Zvolensky, M. J., & Pedulla, C. M. (2003). Evaluation of the Balloon Analogue Risk (BART) as a predictor of adolescent real-world risk-taking behaviors. *Journal of Adolescence, 26*(4), 475–479.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84.

Lewellyn, K. B., & Muller-Kahle, M. I. (2012). CEO power and risk taking: Evidence from the subprime lending industry. *Corporate Governance: An International Review*, *20*(3), 289–307.

Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, *64*, 19–40.

Magee, J. C., & Galinsky, A. D. (2008). 8 social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals*, *2*(1), 351–398.

Magee, J. C., & Smith, P. K. (2013). The social distance theory of power. *Personality and Social Psychology Review*, *17*(2), 158–186.

Maner, J. K., Gailliot, M. T., Butz, D. A., & Peruche, B. M. (2007). Power, risk, and the status quo: Does power promote riskier or more conservative decision making?. *Personality and Social Psychology Bulletin*, *33*(4), 451–462.

Nave, C. S., Feeney, M. G., & Furr, R. M. (2018). Behavioral observation in the study of personality and individual differences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *The SAGE handbook of personality and individual differences: The science of personality and individual differences* (pp. 317–340). London: SAGE.

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *The Journal of General Psychology*, *135*(2), 151–166.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783.

Orne, M. T. (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143–179). New York, NY: Academic Press.

Pannenberg, M. (2010). Risk attitudes and reservation wages of unemployed workers: evidence from panel data. *Economics Letters*, *106*(3), 223–226.

Pashler, H., Rohrer, D., & Harris, C. R. (2013). Can the goal of honesty be primed?. *Journal of Experimental Social Psychology*, *49*(6), 959–964.

Patterson, M. L. (2008). Back to social behavior: Mining the mundane. *Basic and Applied Social Psychology*, *30*(2), 93–101.

Patterson, M. L., Giles, H., & Teske, M. (2011). The decline of behavioral research? Examining language and communication journals. *Journal of Language and Social Psychology*, *30*(3), 326–340.

Perdue, B. C., & Summers, J. O. (1986). Checking the success of manipulations in marketing experiments. *Journal of Marketing Research*, *23*(4), 317–326.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, *12*(4), 531–544.

Podsakoff, P. M., & Podsakoff, N. P. (2019). Experimental designs in management and leadership research: Strengths, limitations, and recommendations for improving publishability. *The Leadership Quarterly*, *30*(1), 11–33.

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.

Rinderknecht, R. G. (2019). Effects of participant displeasure on the social-psychological study of power on Amazon's Mechanical Turk. *SAGE Open*, *9*(3), 1–13.

Ronay, R., & Von Hippel, W. (2010). Power, testosterone, and risk-taking. *Journal of Behavioral Decision Making, 23*(5), 473–482.

Rosnow, R. L., & Rosenthal, R. (1997). *People studying people: Artifacts and ethics in behavioral research.* New York, NY: Freeman.

Rucker, D. D., Dubois, D., & Galinsky, A. D. (2011). Generous paupers and stingy princes: Power drives consumer spending on self versus others. *Journal of Consumer Research, 37*(6), 1015–1029.

Schaerer, M., Lee, A. J., Galinsky, A. D., & Thau, S. (2018). Contextualizing social power research within organizational behavior. In D. L. Ferris, R. E. Johnson, & C. Sedikides (Eds.), *The self at work: Fundamental theory and research* (pp. 194–221). New York, NY: Routledge.

Schmid, P. C. (2018). Less power, greater conflict: Low power increases the experience of conflict in multiple goal settings. *Social Psychology, 49*(1), 47–62.

Schmid, P. C., Kleiman, T., & Amodio, D. M. (2015). Power effects on cognitive control: Turning conflict into action. *Journal of Experimental Psychology: General, 144*(3), 655–663.

Shanks, D. R., Newell, B. R., Lee, E. H., Balakrishnan, D., Ekelund, L., Cenac, Z., … Moore, C. (2013). Priming intelligent behavior: An elusive phenomenon. *PLOS One*, *8*(4), 1–10.

Sigall, H., & Mills, J. (1998). Measures of independent variables and mediators are useful in social psychology experiments: But are they necessary?. *Personality and Social Psychology Review, 2*(3), 218–226.

Smith, P. K., Jostmann, N. B., Galinsky, A. D., & van Dijk, W. W. (2008). Lacking power impairs executive functions. *Psychological Science, 19*(5), 441–447.

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.

Sturm, R. E., & Antonakis, J. (2015). Interpersonal power: A review, critique, and research agenda. *Journal of Management, 41*(1), 136–163.

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.

Tost, L. P., Gino, F., & Larrick, R. P. (2013). When power makes others speechless: The negative impact of leader power on team performance. *Academy of Management Journal, 56*(5), 1465–1486.

West, S. G., & Brown, T. J. (1975). Physical attractiveness, the severity of the emergency and helping: A field experiment and interpersonal simulation. *Journal of Experimental Social Psychology, 11*(6), 531–538.

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

Footnotes

[1] We focus on explicit and not on implicit power priming methods because the former are the ones used more widely and the ones most prone to demand effects. When we write "power priming" we mean "explicit" priming of power.

[2] There were three attention checks. Failure in at least two of them led to exclusion from the whole study. The items of the attention checks were: "I sleep less than an hour per night", "I eat a computer everyday", and "All my friends are aliens".

[3] In the RA condition, the low-power participants (receivers) felt rather more powerful (M = 2.88, SD = 1.60) than the participants in the neutral condition (observers) (M = 2.71, SD = 1.79). Although the difference is not significant, two possible reasons come to mind explaining this unexpected pattern of results for the observers: 1) being excluding from the distribution game and not receiving any tickets from the allocator and 2) being faced with the allocator's distribution of the lottery tickets with which the observers did not necessarily agree might have lowered their felt power.

[4] The fixed payments were paid instantly at the end of the study and participants were informed that the varied payments were paid with a time lag after the study as bonus payments on the Mturk platform. Because the final payoffs had to be calculated based on the decisions participants made in each step.

[5] Adding control variables (gender, age, and ethnicity) did not change the results: main effect of the method of power manipulation, SM: $F(1, 1564) = 81.49$, $p < .001$/UM: $F(1, 1564) = 42.18$, $p < .001$, main effect of manipulation check, SM: $F(1, 1564) = 81.51$, $p < .001$/UM: $F(1, 1564) = 9.26$, $p < .01$, and their interaction effect, SM: $F(1, 1564) = 7.46$, $p < .01$/UM: $F(1, 1564) = 0.98$, $p = .32$.

[6] The potential for demand effects is significantly lower under RA without a manipulation check compared to (1) RA with manipulation check (SM: $M = 0.32$, $SD = 0.47$, $F(1,1629) = 70.24$, $p < .001$/UM: $M = 0.05$, $SD = 0.22$, $F(1,1629) = 7.46$, $p = .006$), (2) PP without manipulation check (SM: $M = 0.33$, $SD = 0.47$, $F(1,1629) = 72.06$, $p < .001$/UM: $M = 0.10$, $SD = 0.30$, $F(1,1629) = 30.53$, $p < .001$), and (3) PP with manipulation check, (SM: $M = 0.46$, $SD = 0.50$, $F(1,1629) = 164.23$, $p < .001$/UM: $M = 0.12$, $SD = 0.33$, $F(1,1629) = 46.14$, $p < .001$).

[7] Adding control variables (gender, age, and ethnicity) did not change the results: main effect of type of power manipulation, $F(1, 1560) = 5.26$, $p = .02$.

[8] However, the result of a contrast test comparing the effect of power on self-reported risk-taking in the low demand-prone condition against the effect of power in all high demand-prone condition taken together was non-significant, $F_{contrast}(1, 1625) = 0.79$, $p = .37$.

[9] Results remained non-significant when adding control variables (gender, age, and ethnicity).

[10] We did the same contrast analyses as for the consequential risk-taking measure and none was significant.

[11] The BART task and lottery task correlated significantly, $r = .23$, the lottery task and the self-report risk-taking also correlated significantly, $r = .11$, the BART and the self-report task did not correlate, $r = .02$.

Table 1

ANOVA Results for the Potential for Demand Effects

Panel A: Structured Measure (SM)

Variables	Sum of Squares	*df*	Mean Square	*F*	*p* value
Power manipulation method	16.16	1	16.16	86.39	.00**
Manipulation check	14.86	1	14.86	79.43	.00**
Power manipulation method × Manipulation check	1.37	1	1.37	7.30	.01**
Error	304.75	1629	0.19		

Panel B: Unstructured Measure (UM)

Variables	Sum of Squares	*df*	Mean Square	*F*	*p* value
Power manipulation method	2.90	1	2.90	46.75	.00**
Manipulation check	0.49	1	0.49	7.91	.01**
Power manipulation method × Manipulation check	0.06	1	0.06	0.94	.33
Error	100.88	1629	0.06		

Note. ** $p < .01$ * $p < .05$.

Table 2

ANOVA Results for Self-reported Risk-taking

Variables	Sum of Squares	*df*	Mean Square	*F*	*p* value
Power manipulation method	0.39	1	0.39	0.15	.70
Power level	11.87	1	11.87	4.47	.03*
Manipulation check	1.69	1	1.69	0.64	.42
Power manipulation method × Manipulation check	1.81	1	1.81	0.68	.41
Power level × Manipulation check	0.19	1	0.19	0.07	.79
Power manipulation method × Power level	0.88	1	0.88	0.33	.56
Power manipulation method × Power level × Manipulation check	1.22	1	1.22	0.46	.50
Error	4310.97	1625	2.65		

Note. ** $p < .01$ * $p < .05$.

Table 3

ANOVA Results for the BART Risk-taking Task

Variables	Sum of Squares	*df*	Mean Square	*F*	*p* value
Power manipulation method	27.15	1	27.15	0.51	.48
Power level	8.23	1	8.23	0.15	.69
Manipulation check	50.41	1	50.41	0.95	.33
Power manipulation method × Manipulation check	11.34	1	11.34	0.21	.64
Power level × Manipulation check	10.35	1	10.35	0.19	.66
Power manipulation method × Power level	11.51	1	11.51	0.22	.64
Power manipulation method × Power level × Manipulation check	28.86	1	28.86	0.54	.46
Error	86585.225	1625	53.28		

Note. ** $p < .01$ * $p < .05$.

Table 4

ANOVA Results for the Lottery Risk-taking Task

Variables	Sum of Squares	*df*	Mean Square	*F*	*p* value
Power manipulation method	0.17	1	0.17	0.08	.77
Power level	0.11	1	0.11	0.05	.82
Manipulation check	4.89	1	4.89	2.42	.12
Power manipulation method × Manipulation check	0.65	1	0.65	0.32	.57
Power level × Manipulation check	1.04	1	1.04	0.52	.47
Power manipulation method × Power level	4.66	1	4.66	2.31	.13
Power manipulation method × Power level × Manipulation check	0.66	1	0.66	0.33	.57
Error	2804.38	1387	2.02		

Note. ** $p < .01$ * $p < .05$.

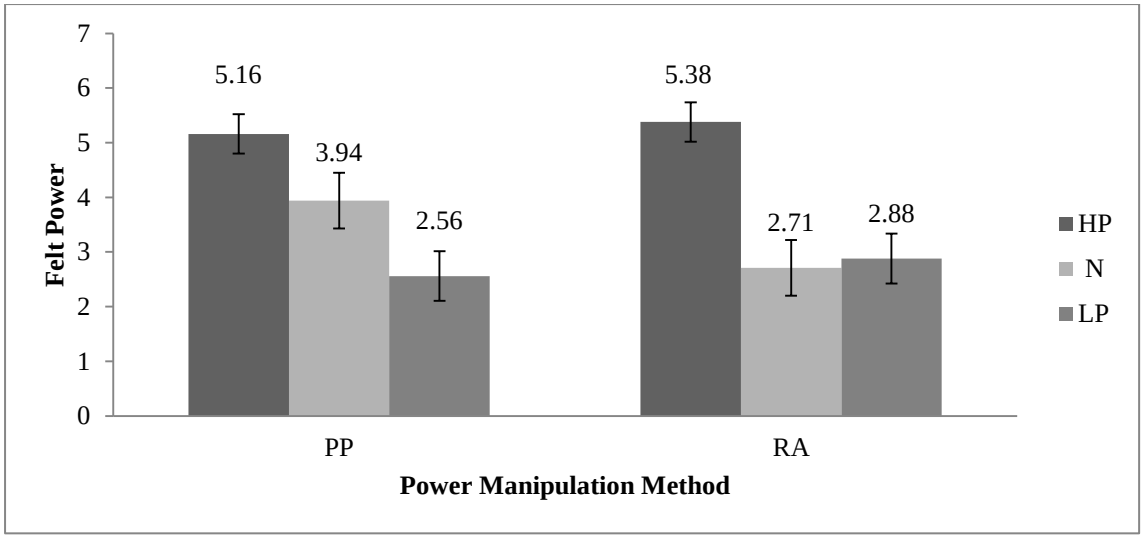


Figure 1. The effect of power manipulation method on felt power in the pre-test, phase one (*N* = 323). PP = power priming, RA = resource allocation; three power levels: HP = high-power, N = neutral, and LP = low-power.

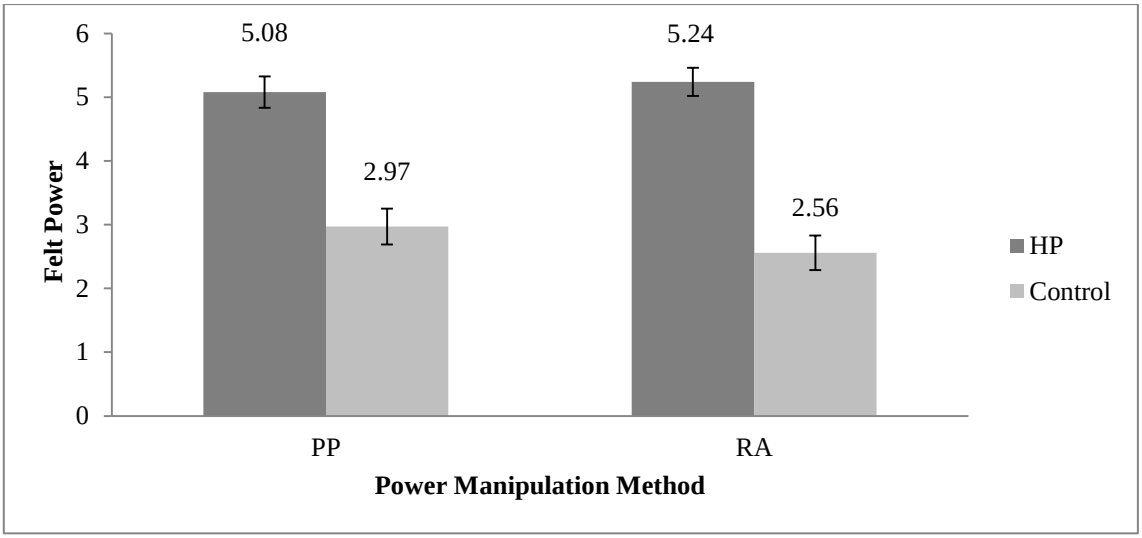
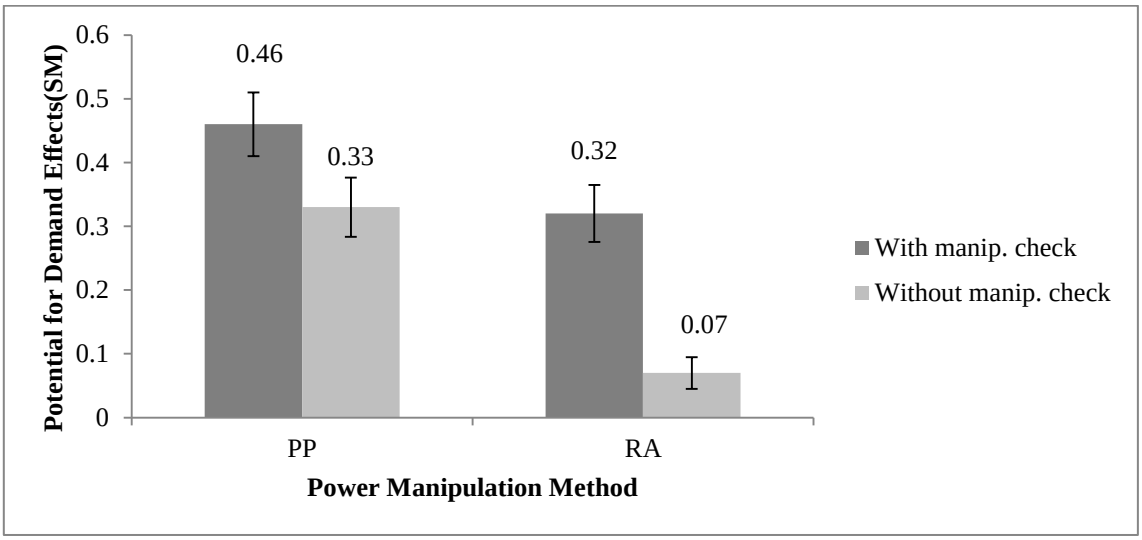


Figure 2. The effect of power manipulation method on felt power in the pre-test, phase two (*N* = 606). PP = power priming, RA = resource allocation; two power levels: HP = high-power and control.

Panel A: Structured Measure (SM)



Panel B: Unstructured Measure (UM)

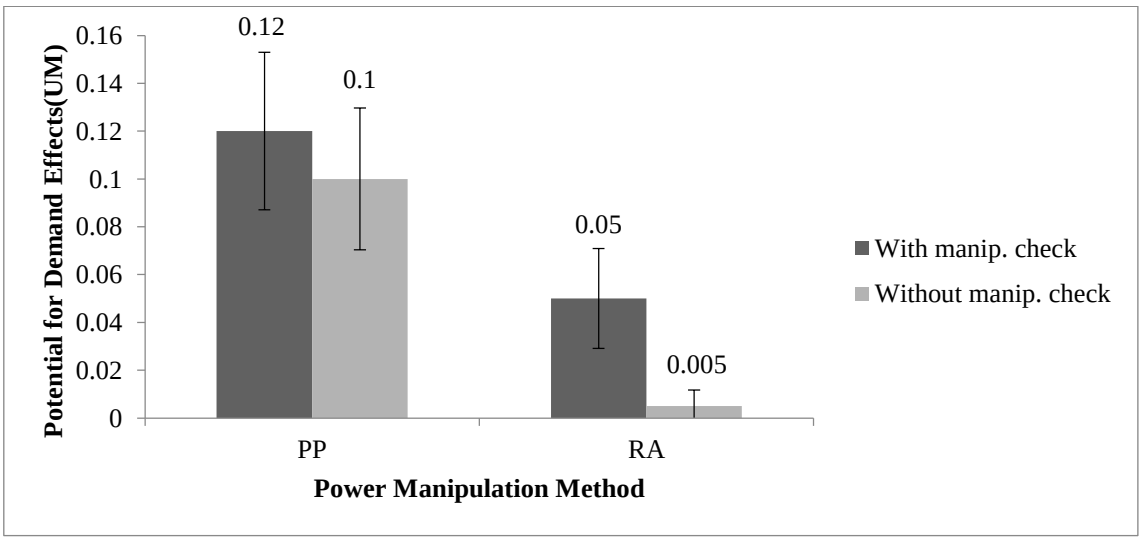


Figure 3. The effect of power manipulation method and manipulation check on the potential for demand effects using a structured measure (SM) and an unstructured measure (UM). PP = power priming, RA = resource allocation; With manip. check = with manipulation check, Without manip. check = without manipulation check.

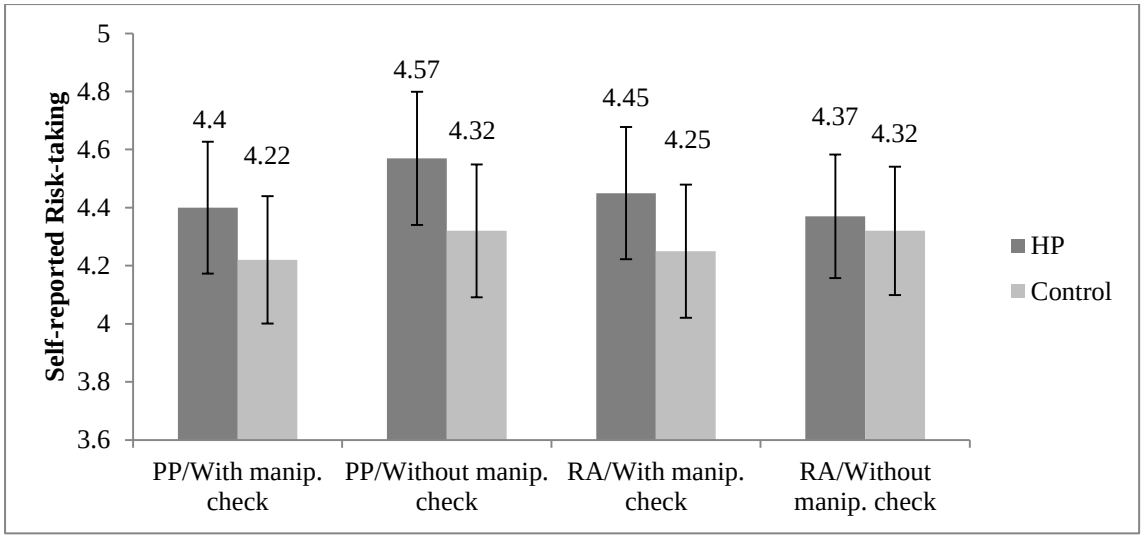


Figure 4. The effect of power manipulation method, manipulation check, and power level on self-reported risk-taking. PP = power priming, RA = resource allocation; With manip. check = with manipulation check, Without manip. check = without manipulation check; HP = high-power.

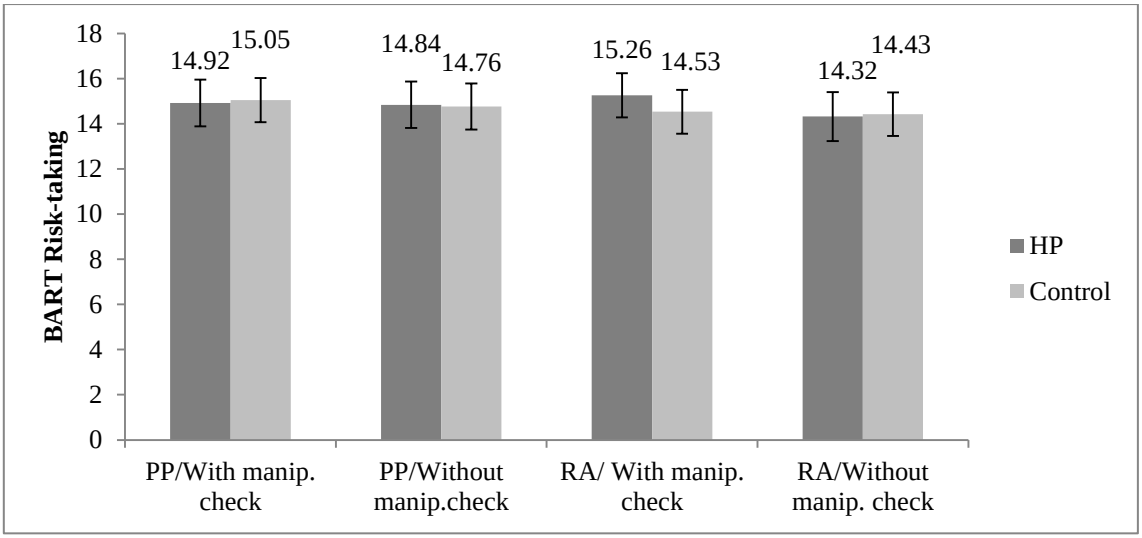


Figure 5. The effect of power manipulation method, manipulation check, and power level on BART risk-taking. PP = power priming, RA = resource allocation; With manip. check = with manipulation check, Without manip. check = without manipulation check; HP = high-power.

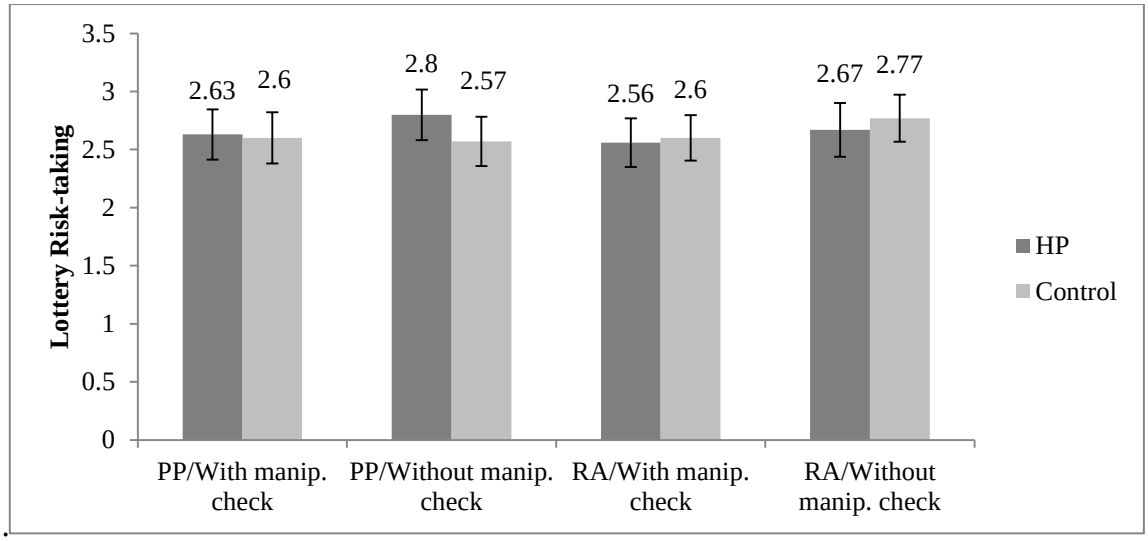


Figure 6. The effect of power manipulation method, manipulation check, and power level on lottery risk-taking. PP = power priming, RA = resource allocation; With manip. check = with manipulation check, Without manip. check = without manipulation check; HP = high-power.

Leader's Unethical Behavior Depends on the Level of Power, Machiavellianism, and Narcissism

Mahshid Khademi, Marianne Schmid Mast, Christian Zehnder

Faculty of Business and Economics, University of Lausanne, Switzerland

Author Note

Mahshid Khademi, Marianne Schmid Mast, and Christian Zehnder, Department of Organizational Behavior, Faculty of Business and Economics, University of Lausanne, Switzerland.

Correspondence concerning this article should be addressed to Mahshid Khademi, Department of Organizational Behavior, University of Lausanne UNIL- Chamberonne, Internef, CH-1015 Lausanne, Switzerland. Phone number: +41 21 692 34 50 E-mail: Mahshid.Khademi@unil.ch

This work was supported by HEC Research Fund for PhD and Postdocs. Results of this study have been previously been presented at the European Association of Work and Organizational Psychology (EAWOP) and European Business Ethics Network (EBEN) conference.

Abstract

A widespread belief suggests that power corrupts. But we posit that individuals differ in how they use their power depending on their personality characteristics. In two studies, using incentivized experimental design, we examined the effect of power and individual differences on unethical behavior. We manipulated power by the number of followers and the amount of control a leader had over the followers' outcomes. Results showed no support for a general effect of "power corrupts". However, in Study 1 we found an interaction effect between Machiavellianism and power on leader's unethical behavior: individuals with higher levels of Machiavellian orientation more engaged in unethical behavior when they are appointed to high-power position compared to low-power positions. Results of the Study 2 showed that in high-power positions with self-serving opportunities, Narcissism increased the likelihood of leader's unethical behavior. Our findings highlight the importance of selecting the right people for high stakes positions.

Keywords: power, unethical behavior, Machiavellianism, Narcissism

Leader's Unethical Behavior Depends on the Level of Power, Machiavellianism, and Narcissism

Power is known as the control and influence one has over others (Schmid Mast, 2010). We tend to believe that power has a corruptive nature and individuals with absolute power are more likely to take ethical shortcuts. But is this true for everyone in power? More precisely, does power corrupt everybody or it rather acts as a magnifier on pre-dispositional ethical qualities?

Power leads to behavioral change (Keltner, Gruenfeld, Anderson, 2003; Smith & Bargh, 2008). For instance, high-power individuals tend to pay less attention to the social environment compared to low-power individuals (Van Kleef, De Dreu, Pietroni, & Manstead, 2006), are more attuned to their situational goals, needs, and feelings (Guinote, 2007), express less compassion, when facing others' suffering (Van Kleef, Oveis, Van Der Löwe, LuoKogan, Goetz, & Keltner, 2008), and more prone to objectify social targets (Gruenfeld, Inesi, Magee, & Galinsky, 2008). But, some studies show that individual characteristics can influence the link between power and behavioral change in powerholders (e.g., Chen, Lee-Chai, & Bargh, 2001; DeCelles, DeRue, Margolis, & Ceranic, 2012). Galinsky et al., (2015), in their review, suggest that in order to harness the good part of power we need to be more vigilant in selecting the right people for power positions. Yet, little research has been conducted to understand the individual traits of the powerholders that produces positive (or negative) ethical outcomes.

Moreover, existing research on the dark side of power has mostly induced psychological power (Tost, 2015). For example, Schaerer et al. (2018) in a review article found that more than half of the studies in power domain up until 2015 are comprised of manipulations where an experience of psychological power is manipulated instead of structural and objective power. This is alarming as scholars have expressed their concern on

the issue of demand effects in such methods (Flynn, Gruenfeld, Molm, & Polzer, 2011; Sturm & Antonakis, 2015; Lonati, Quiroga, Zehnder, & Antonakis, 2018). Demand effects refer to the experimental cues and expectations that can influence participant behavior (Orne, 1962). The presence of demand characteristics can be the sources of distortion and biases in participants' outcome behavior (Zizzo, 2010). As most research conducted on the dark side of power are based on such methods (Tost, 2015) that can potentially generate the chances of demand characteristics, we felt the need to test the effect of power and unethical behavior using an alternative method in which power is induced objectively. We thus aim to further investigate the effect power on unethical conducts using a robust and rigorous design in which objective power is manipulated. Furthermore, we examine the types of individual characteristics that bring on negative behavior in individuals who are granted objective power, or, put differently, the personality characteristics that prevent powerholders from abusing their power if the opportunity presents itself.

Power and Unethical Behavior

The link between power and unethical behavior can be explained through two pathways: self-focus and disinhibition (Lammers, Galinsky, Dubois, & Rucker, 2015). Unethical behavior refers to a behavior that exerts a harmful effect on others (Brass, Butterfield, & Skaggs, 1998) and is "either illegal or morally unacceptable to the larger community" (Jones, 1991, p. 367). Power is linked to unethical behavior because feeling powerful typically generates a sense of social distance from others (Magee & Smith, 2013) that induces self-focus (Lammer et al., 2015). For instance, Dubois, Rucker, and Galinsky (2014) found that powerful individuals tended to lie more to benefit themselves whereas low-power participants were more likely to lie to benefit the others. This shows that high-power individuals might engage more in unethical behavior when it comes to their self-benefit.

Besides, according to Approach-Inhibition Theory (Keltner et al., 2003), power decreases the activation of the behavioral inhibition system. The behavioral inhibition system helps individuals to avoid unacceptable behaviors and therefore cautions a person with respect to deciding about the appropriate behavior in each situation (Gary, 1982). Individuals who experience power have low behavioral inhibition activation that leads to a release from the inhibiting effect of social norms (Hirsh, Galinsky, & Zhong, 2011). Therefore, powerholders tend to go against established social norms and be more likely to engage in unethical behaviors (Lammers, Stapel, and Galinsky, 2010; Bendahan, Zehdner, Pralong, & Antonakis, 2015). For example, Bendahan and colleagues (2015) gave high-power individuals an endowment to distribute between their followers while they could keep part of the endowment for themselves. Previously, to establish the social norm, they had asked all the participants what a responsible leader would decide in terms of the distribution of the endowment. They observed that individuals in the high-power condition diverged more from the established norm compared to those who were in low-power condition.

Power and Individual Characteristics

It seems that empirical evidence support the widespread belief that power corrupts. But it is important to understand if there are individual differences with respect to the link between power and unethical behavior. Power in general increases the correspondence between individual traits and behavior (Galinsky, Rucker, & Magee, 2015). By reducing the dependence on others, power reveals the traits and attitudes of the powerholder (Hirsh et al., 2011; Morrison, See, & Pan, 2015). For instance, Chen et al. (2001) found that individuals high on communal orientation are more likely to behave generously (i.e., invest their own time in order to decrease the working time of the subordinates) when primed with high-power, whereas exchange orientation led individuals to behave in a more self-serving manner. In another study, Decelles et al. (2012) showed that individuals with a strong moral identity who

were primed with power were less likely to engage in self-interest behavior (taking points from a shared pool of points among participants) compared to those with weak moral identity. Moral identity is "the extent to which an individual considers morality as a part of his/her self-concept" (Aquino & Reed, 2002, p. 1424). A similar personality trait that can be important in moral dilemmas is Moral Disengagement: the propensity to disengage from the distress of an unethical behavior (Bandura, 1990). According to Moore et al. (2012) Moral Disengagement ease the process of facing amoral act by reframing unethical behavior such that it is serving a greater goal. Moral Disengagement generally predicts unethical behavior (i.e., self-reported cheating, lying, and stealing in the workplace) (Moore, Detert, Klebe Treviño, Baker, & Mayer, 2012). Therefore, we can expect that Moral Disengagement facilitates an individual's path to unethical behavior especially when the individual has control over valued resources in the position of power.

Besides, Bendahan et al. (2015) suggested that future research needs to investigate the link between power and unethical behavior with regards to the dark personality traits (Machiavellianism, Narcissism, and Psychopathy). One dimension of the dark triad is Machiavellianism which refers to the use of interpersonal manipulation, deceit, and amoral viewpoints towards others (Christie & Geis, 1970). Machiavellian orientation in general, is negatively associated with ethical judgment and behaviors (Rayburn & Rayburn, 1996; Verbeke, Ouwerkerk, & Peelen, 1996). Wisse and Sleebos, (2016) in a field study found that supervisor's Machiavellian orientation increases abusive behavior only when the supervisor perceived their position as high-power rather than low-power (power was measured by a self-report scale on the amount of control the supervisor perceived to have over valued resources). Hence, we can conclude that high-power situation that provides freedom to exercise ones favored behavior, a person's Machiavellian personality could facilitate unethical behavior.

Methodological Aspects in Studying Power

Existing research that experimentally induced power has commonly used priming methods (Schaerer, Lee, Galinsky, Thau, 2018). Two of the most common ways to prime power are: Assigning a high- or low-power roles to an individual in a role play scenario (Anderson, & Berdahl, 2002) or asking participants to recall and write about a power situation in which they had power over others (or someone else had power over them for the low-power condition) (Galinsky, Gruenfeld, & Magee, 2003). These priming methods have raised many concerns among scholars, as they tend to increase the chances of generating demand effects (Sturm & Antonakis, 2015) which refer to "*subjects' awareness of the implicit aspect of the psychological experiment [that] may become the principal determinant of [their] behavior*" (Orne, 2009, p. 111). In priming methods where participants are asked to role play or to recall a time they felt powerful, the chance that the purpose of the study becomes salient increases because power notions are explicitly discussed with participants. Consequently, the effect of power priming in these cases might be an effect of the expectation the participant harbors about the goals of the study; and participant might thus behave in accordance to what they think power typically *should* entail and the way a powerful person *should* behave.

Research in behavioral economics tries to avoid the above mentioned issues by giving participants objective power (e.g., Sivanathan, Pillutla, & Murnighan, 2008; Bendahan et al., 2015). Objective power can be induced via a dictator game in which one participant takes the role of the dictator (leader) and has the task to distribute an endowment between him/herself and other group members (followers). In such settings, the power manipulation has been done by increasing the number of followers and number of distribution choices the leader has (Bendahan et al., 2015; Joshi & Fast, 2013). All the players (including the leader) know that the leader's distribution will determine the ultimate pay-offs. This way we can ensure that power affects the powerholder (leader) not by expectations or assumptions about how one

should behave but rather through the real control the leader has over the outcome of his/her followers (Strum & Antonakis, 2015).

However, one has to be careful about the notion of *power* that sometimes becomes tangled with confounding effect of *having access to more monetary gain*. For instance, Bendahan et al. (2015) used a behavioral economics approach to test the question of power and corruption. In their study, power was manipulated by the number of the discretionary choices and the number of followers each leader was assigned to. The leader's decision determined how corrupt a leader was if the distribution was against the norm accepted among the participants of the study. Results showed that high-power leaders engaged in more corrupted choices compared to the low-power individuals. The researchers concluded that power corrupts. But high-power individuals in this study, not only had more followers and more choices, but they also had access to more monetary gain. That is, the high-power individuals had bigger endowments to distribute between themselves and the followers (see the table of parameters, Bendahan et al., 2015; p.111). Being exposed to bigger endowment can lead to more enticement, and thus more corrupted behavior in such settings (Engel, 2011). Therefore, it is possible that the corrupted behavior of the powerholders was related to the size of the endowment and the ultimate gain rather than power per se.

To remedy this problem, the power manipulation in the current study is based on the real incentivized game to avoid demand effects (as was the case in Bendahan et al.) and at the same time the power holders (leaders) were given a fixed amount of money and had no gain in the endowment that they had to distribute among their followers. This procedure was designed in order to avoid confounds due to income effect.

Study 1

The goal of the current study is twofold: 1) test whether the effect of power on unethical behavior is replicated in a setting in which individuals experience "real" power over

others and 2) test whether individual differences moderate the power-unethical behavior link. As discussed earlier, research shows that power typically leads to moral depravity, but that it can also elevate moral behavior depending on the individual and contextual characteristics (Lammer et al., 2015). In this study, we set out to investigate how Moral Disengagement and Machiavellianism affect the link between power and unethical behavior. We expected that leaders high in power engage in more unethical behavior, as has already been demonstrated in the literature (H1). The main proposition of the current study is that individual characteristics of the power holder moderate the link between power and unethical behavior. Hence, we predicted that Machiavellianism moderates the effect of power on unethical behavior such that individuals who experience high-power are more likely to engage in unethical behavior when they are high in Machiavellianism compared to when they are low in Machiavellianism. We expect Machiavellianism to influence unethical behavior less for individuals with low-power (H2). In the same vein, we hypothesized that Moral Disengagement moderates the effect of power on unethical behavior such that individuals who experience high-power are more likely to engage in unethical behavior when they are high in Moral Disengagement compared to when they are low in Moral Disengagement. We expected Moral Disengagement to influence unethical behavior less for individuals with low-power (H3).

Pre-test of the Power Manipulation

To check how powerful the leaders would feel in the designed power manipulation, we conducted a manipulation check, prior to running the main study, with an independent sample of participants. By conducting a separate study, we avoided having to ask participants during the study about how powerful they felt which would have revealed the aim of the study and created the chances of demand characteristics (Antonakis, 2017). Moreover, if we asked participants at the end of the study and before measuring the outcome behavior, their final decision concerning the unethical behavior might be influenced by their answer given to the

manipulation check and their guesses about the hypothesis (Lonati et al., 2018). That is why we opt for a separate but comparable sample and asked participants to imagine themselves as the leaders and rate how powerful such leader would feel in each condition.

More precisely, two weeks before the main experiment, we recruited 60 participants for an online study in which we described the designed setting, the roles, the options, and the possible payoffs. Then, we asked participants to put themselves in the shoes of the powerholder and rate the extent to which the leader would feel powerful in one of the four conditions (i.e., low-power: 2 followers and no control over the distribution of the money for the followers, medium power (I): 2 followers and full control, medium power (II): 5 followers and no control, and high-power: 5 followers and full control).

Participants did not report a significant difference regarding the power of the leader between the two medium power manipulations. However, they rated the leader in the high-power manipulation setting significantly more powerful than the leader in the low-power manipulation setting, $t(25) = 2.40$, $p = 0.01$. Based on these results, we chose to run the full study only with the high power manipulation (five followers and absolute control) and the low power manipulation (two followers and no control).

Method

Participants. A total of 440 Master and Bachelor students were recruited from the participant pool of a university in the French speaking part of Switzerland. On average, they were on average 21.5 years old. Participants were required to be native French speakers. They were informed that their payoff depended on their decision and the decisions of their group members but on average they gained an equivalent of 17 USD for their participation.

Procedure and Measures. The study took place in a large computer room with several groups at the same time. Participants were invited to the lab in large groups (21 to 32 individuals per session). They came to the lab and then were led to their individual cubicles,

sat in front of a computer to complete a two-stage task. But, first they were randomly assigned to a 5-person groups (or 7-person groups depending on the condition of the study). In each group, one participant was randomly assigned to play the role of the leader and the rest were followers. The leaders and followers remained anonymous to each other during the study, because the interactions were managed via a computer interface and anonymously. The leaders remained leaders in the first and second stage of the task.

We manipulated leaders' power in two different ways, but always correspondingly. First, a leader's power depended on the number of followers supervised. The leader in our scenario supervised either 5 followers (high-power condition) or 2 followers (low-power condition). Second, a leader's power also depended on how much control over the followers' outcomes a leader had, which was manipulated by either leaders having absolute control over the distribution of followers' payoffs (high-power condition) or having no control over the distribution of followers' payoffs (low-power condition).

In the first stage of the task, followers had to perform in a task the idea behind this was to provide the leader with an array of followers' performances based on which the leader can decide about the distribution outcome. Such task performance also help increasing the ecological validity of the study as followers are earning their payoffs based on their performance in the task. Emphasis was put on instructing the leader that he or she had full discretion about which distribution option to choose. Whereas, in the low-power condition, the leader was instructed to push a button that would select randomly a distribution option determining the proportion of the payoffs amongst the two followers. The leader could not take any of this money for him/herself. Followers were then informed about the distribution and about the ultimate payoff they would receive. Therefore, in this stage we first randomly excluded one of the followers from the other group members and had him/her complete a task for the second stage of the experiment which will be explained later. But, all other followers

did a slider task in which they had a slider on the screen with the value ranging from 0 to 100. In each attempt of the task, participants needed to position the mouse on the number 50 on the slider. After each incorrect attempt, the slider showed the participants the outcome. They could re-try as many times as they wanted to adjust the position of the mouse again. However, participants knew that they only had two minutes to correctly position as many sliders as possible. Followers' performance was measured by the number of correctly positioned sliders in two minutes. These numbers were communicated to the leader, (via the computer interface). The leader had a fixed amount of money ($10.-) per follower at his/her disposal (that would make $20.- for the leaders in the low-power condition and $50.- for the high-power condition). Leader's task was to distribute all of the money among the followers either personally or randomly depending on the condition. In the high-power condition, the leaders were instructed to distribute the payoffs after having learned about each follower's performance. However, the leaders were not instructed to distribute the money according to performance. Three options were proposed to the leaders: (1) distribute the endowment based on the performance rates, (2) distribute the endowment evenly, and (3) the best performer wins all. After the distributions were done and the followers were informed, the leader had to move on to the second stage of the task.

As explained, in the first phase, high-power leaders directly controlled the distribution choice while in the low-power condition leaders faced a random distribution. We intended to rule out the effect of the type of distribution choices in the first stage on leader's unethical behavior in the second stage across conditions. This is why we conducted a matching process: we first ran the high-power condition to be able to record the distribution choices of high-power leaders. Then based on the distribution outcomes of the high-power leader we calibrated the random distributor for the low-power leaders to achieve to similar probability of showing the choice outcomes. For example, if the high-power leader selected the equal

distribution with 0.7 chance, the random distributor equally distributed the endowment with 0.7 chance as well. This way, the distribution choices of the leaders in both conditions are matched and cannot differentially affect the behavior of the leaders in the second stage of the task.

The second stage of the task had two phases: a production phase and a distribution phase. The idea behind the production phase was that the followers performed a task and earned their money based on their performance. The earned money then in the distribution phase will be given to the leader to be transferred to the follower. As this money was made by the follower if the leader took any part of it we could consider this an act of unethical behavior. It is important to note the difference between a selfish and an unethical behavior. A selfish act is not always necessarily an unethical behavior if it leads no harm to others (Lu, Zhang, Rucker, & Galinsky, 2017). For example lying in a random die game for gaining more money might be regarded as selfish but not unethical because it does not directly harm another stake holder. In our setting we distinguished between selfish and unethical behavior as by using a version of the dictator game in which the final endowment fully belonged to the follower. Withholding all or any part of the endowment was considered as an act of unethical behavior from the leader because it was framed as stealing from another person.

For the production phase, the last follower who was excluded from the others in stage one was asked to solve 5 anagrams in 5 minutes. An anagram is a word formed by rearranging the letters of another word. To solve an anagram, the follower needed to use all the given letters of a word to build a new word. The follower was then informed that each solved anagram was worth $2.- and his/her performance was going to be communicated to the leader, and that it was the leader who would distribute the total sum of the money earned by the follower (endowment) between him/herself and the follower and hence determine the final payoff of the follower. After 5 minutes of follower production, the leader was informed about

the follower's performance and about how much money the follower is entitled for this production.

In the distribution phase, the experimenter gave $10.- to the leader to keep plus the follower's endowment from the production. The experimenter emphasized to the leader that $10.- are for the leader to keep (reinforces the power position of the leader) and the rest of the money is what the follower earned for his/her work. Previously, an independent group of participants from the same university was provided with the information of the roles in the study and was asked what they would have considered an ethical distribution decision in this case by the leader (described in more detail below). The majority mentioned that the leader should not take any part of the follower's endowment. Prior to taking the distribution decision, we informed the leaders that the majority of their colleagues believed that the leader should not withhold any part of the follower's endowment. The leader was free to distribute the follower's endowment between the follower and him/herself as (s)he pleased.

Note that this time and unlike in the first stage of the study, the leader was allowed to take a part or the entire endowment for him/herself. In fact, at this stage, the leader was playing a dictator game with the endowment that entirely was the result of the follower's effort. Although the follower was passive and had no power to refuse or change the decision of the leader, in contrast to other dictator games, in our setting, the leader had a valuable piece of information: He/she knew about the property right of the follower in the endowment.

Moreover, we were interested not only on the selfish distribution of the endowment but also a destructive unethical behavior. That is why we added an element of inefficiency in leader's unethical behavior and by that we mean increasing the leader's payoff -by taking the follower's money- is socially inefficient. That is, as the leader increases his/her payoff, the follower's payoff decreases by a greater amount. In this case, taking 1 unit of the follower's endowment would value 0.8 units for the leader which made the overall outcome worse off if

the leader decided to take advantage of his position. In sum, the unethical behavior of the leader can be seen as a selfish and socially destructive action.

The leader distributed the payoff via the computer system by writing the amount that he/she decided to give to the follower and this amount was communicated to the follower as his/her final payoff from this stage of the game. The leader's total payoff was calculated as the fixed participation fee in the experiment (10.-) plus the amount of money taken from the follower. The follower's total payoff was calculated as the fixed participation fee in the experiment (10.-) plus the amount of money accorded by the leader. The leaders and the followers were unaware of other the final payoffs of the different roles. At the end, the experimenter called the participants one by one, paid the calculated final payoff.

Norm elicited. As previously mentioned, unethical behavior is defined as a behavior that is against the accepted social norm and hurt others. To label an act as ethical or unethical we needed to understand what would be considered as an acceptable behavior from the leader in such situations. Therefore, we had conducted the norm elicitation phase two weeks before the experiment. Sixty-one students from the same university as the participants of the experiment, were invited to the norm elicitation study. Participants were provided with the information of the roles in the designed game, the options available for different roles, and the ultimate payoffs of the each role. Then they were asked what they would have considered an ethical distribution by the leader who has an endowment of $10.- (The endowment assumed to be $10.- because the anagrams were designed simple enough so that every follower could solve them). The participants had to write down the amount that they think the leader can take from the follower endowment between 0 to 10). What the majority considered appropriate was the accepted norm for the experimental design. Among the participants, 57% mentioned that the leader should not take any part of follower's endowment. Thus, we informed the participants in the experiment that the majority of their colleagues believed that the leader

should not withhold any part of follower's endowment. We explicitly conveyed this information to all the leaders to ensure that participants in leader roles were informed what is considered as an appropriate behavior. This way any deviance from the accepted norm can be considered as a conscious act against the accepted social norm and thus unethical.

Leader's unethical behavior. According to the principle of distributive justice, people should be paid based on their performance (Adams, 1965; Cook & Hegtvedt, 1983). It therefore seems that, given that the leader was not implicated in the performance of the follower, he/she should leave the follower's endowment entirely to him/her. The elicited norm from the comparable sample also emphasized that a responsible leader should not take anything from the follower's endowment. We communicated the accepted social norm by the majority of the participants to the leaders prior to their decision in the last stage of the game. Therefore any amount withheld was a deviance from the accepted social norm and therefore considered unethical behavior (Bendahan et al., 2015). This variable is dichotomous. Meaning that, any amount bigger than zero that the leader kept for him/herself was considered as an act of unethical behavior (M = .58, SD = .50).

Leader gain. The amount that leaders withheld from followers endowment was considered as the leader gain (M = 3.06, SD = 3.36).

Machiavellianism. After the leaders were appointed and before they started the game, they were asked to complete the 20-item Mach-IV scale (Christie and Geis, 1970) measuring their Machiavellian orientation. This scale contains cynical and positive views towards human nature (e.g., "*Generally speaking, men won't work hard unless they're forced to do so*" or "*Most people are brave*") and interpersonal tactics in accordance with these point of views (e.g., "*Honesty is the best policy in all cases*" or "*It is wise to flatter important people*"). All items are rated on a seven-point scale from 0 (strongly disagree) to 7 (strongly agree). We

averaged participants' ratings on these four dimensions to create a Machiavellianism score (Cronbach's Alpha = .78., *M* = 3.40; *SD* = 0.58).

Moral Disengagement. After the leaders were appointed and before they started the game, we asked to also report their Moral Disengagement degree. We used Moral Disengagement scale with 8 items developed by Moore, Detert, Klebe Treviño, Baker, and Mayer (2012). Participants are asked to determine the extent to which they agree to 8 statements on a scale from 1 (completely disagree) to 7 (completely agree) such as "*Taking something without the owner's permission is okay as long as you're just borrowing it*" or "*People who get mistreated have usually done something to bring it on themselves*". The items were averaged to obtain a score for Moral Disengagement (Cronbach's Alpha = .73, *M* = 2.91; *SD* = 0.88).

Distribution choices. We recorded the leaders' decision in the first stage of the game as a categorical variable coded as "*distribute the endowment based on the performance rates* = 1" (40% of the times), "*distribute the endowment evenly* = 2" (55% of the times), and "the *best performer wins all* = 3" (5% of the times).

Controls. We controlled for the age and the gender of the leaders as research showed that both of these elements (being male and in younger age group) are good predictors of unethical conducts (Ruegger & King, 1992). We also controlled for the followers performance because it determined the final endowment in the hands of the leader. Note that in the second stage of the game, a follower had to build five anagrams in a limited amount of time (5 minutes). Then, the leader had to distribute this endowment between him/herself and the follower. The amount of the endowment depended on the number of anagrams made by the follower ($2.- for each anagram). We considered the follower's performance (the number of anagrams made by the followers) as the control variable (*M* = 4.63, *SD* = .64).

Results

To test Hypothesis 1 and see whether higher power predicts unethical behavior in leaders, we conducted a Probit model (see Table 1). The result disconfirms a main effect of power on unethical behavior ($\beta = 0.15, p = .61,$). There was no significant difference between the likelihood of leader's unethical behavior in the high-power condition (M = 0.6, SD = 0.50) compared to the low-power condition (M = 0.55, SD = 0.50) (see table 1).

Hypothesis 2 predicted that the effect of power on leader's unethical behavior would be moderated by Machiavellian orientation. The result of a Probit model on the interaction between Machiavellianism and power when controlled for age, gender, and follower's performance revealed a significant interaction effect between power and Machiavellianism explaining the unethical behavior of leaders (β = 1.55, p = .01) (reported in Table 2). For participants in the high-power condition, Machiavellianism was positively related to unethical behavior (β = 1.60, p = .001); but, for low-power leaders the relationship between Machiavellianism and unethical behavior was non-significant (β = .10, p = .78). Taking out the controls did not change our findings (interaction effect remained significant, β = 1.39, p = .019). For example, the slopes in the high- and low-power condition were significantly different ($\chi^2(1) = 6.16, p = .01$). These results confirmed that Machiavellianism moderates the relationship between power and leader's unethical behavior or in other words power only corrupt individuals who are high on Machiavellianism. We graphed our findings in *Figure 1*.

Hypothesis 3 posited that the effect of power on leaders' unethical behavior would be moderated by Moral Disengagement. We ran the same Probit regression analysis but this time with Moral Disengagement and power in the interaction (again controlled for age, gender of the leader, and follower's performance). The result was non-significant ($\beta = .44, p = .23$; see Table 3).

Additional analyses

Although we did not have a prediction about what influenced the amount of money the leaders would take from theirs followers, we ran further analysis to better understand which of the independent variables relate to leader gain. We first tested whether power could explain leader gain when controlled for leader's age, leader's gender and follower's performance. To ensure correct inference, we estimated all models with robust standard errors. The result of an OLS regression again showed no significant effect of power on leader gain ($\beta = .53$, $p = .50$). That is, there was no difference between the gains of the leaders in the high-power condition (M = 3.25, SD = 3.36) compared to the low-power condition (M = 2.88, SD = 3.38), $t(78) = 0.50$, $p = .31$. But, when we introduced Machiavellianism and Moral Disengagement to the regression, we found a significant main effect of Machiavellianism on leader gain (β = 1.35, p = .04). One unit increase in Machiavellian orientation (for example from 3 to 4) increased leader gain by 68.05%. Leader's Moral disengagement was not related to leader gain. Also none of the interaction terms (the interaction between power and Machiavellianism and the interaction between power and Moral Disengagement) were significant.

Discussion

In Study 1, we test how power affects leaders' unethical behavior and how Machiavellianism and Moral Disengagement influence the link between power and unethical behavior. We found no relationship between power and unethical behavior. This result seems to be against the existing findings in the literature as powerholders tend to cheat, lie, and engage in self-serving behavior (Galinsky, Magee, Inesi, & Gruenfeld, 2006; Lammers, Stapel, & Galinsky, 2010). Yet, when we look closer, we notice that most of the research conducted with respect to the dark side of power used methods that induce psychological power (Tost, 2015). It has been shown that such methods convey an opportunistic (as oppose to responsible) frame of power that facilitates behaviors in the same line (Tost & Johnson, 2019). It is possible that our manipulation method helped participants to enforce the

responsibility frame of the power and thus hinder powerholders against the dark side of power.

But, the interaction between Machiavellianism and power significantly predicted unethical behavior of the leaders; individuals with higher level of Machiavellianism when assigned to high-power positions tend to display unethical behavior compared to those who were assigned to low-power positions. This results imply that power can have a corruptive effect on those who are prone to manipulate and exploit others for personal gains (= Machiavellianism). When appointed to high-power positions, Machiavellians are more likely to go against the social norms and engage in unethical behavior. Our findings indicate that Machiavellians tend to take advantage when they are appointed to power positions.

However, our findings did not confirm the moderating effect of Moral Disengagement on the link between power and unethical behavior. We believe that some details regarding the setting of the experiment could explain why. As discussed earlier, Moral Disengagement refers to the ability to disengage from the distress of an unethical behavior (Bandura, 1990). In our design, right before the leader took the (un)ethical decision, the social norm was communicated to them: the leaders were reminded of what was expected from them and the moral costs that they should go through by withholding some of the follower's endowment. By communicating the social norm, we might have made leaders aware of what was the moral thing to do in the situation. This awareness before the final decision could have shielded the leaders from the influence of Moral Disengagement.

Study 2

Our results in study 1 showed that power does not engender unethical behavior. But, individual differences interact with power and trigger unethical behavior. Yet, one criticism on the design of the first study could be that power generally has some benefits for the powerholders. Power is a force that serves the superior as well as the subordinates. But, the

way we manipulated power in our first study did not give powerholders a chance to benefit from their power. Leaders in our high-power condition had valued resources at hand but did not have any opportunity for self-benefit. Therefore, there is a possibility that such design induced a form of power that lacked a dimension of power (control on one's own outcome). To answer the question whether power with self-serving opportunity results in the same findings, we ran our second study. Moreover, we added the full range of the Dark Triad (Machiavellianism, Narcissism, and Psychopathy) as the moderators of the link between power and unethical behavior (suggested by Bendahan et al., 2015). Machiavellianism has shown to be linked to opportunistic and unethical tendencies (Al Aïn, Carré, Fantini-Hauwel, Baudouin, & Besche-Richard, 2013). Narcissism as another dimension of Dark Triad refers to insatiable need for recognition and admiration, a strong sense of self-importance and entitlement, and the desire for authority and arrogance (Emmons, 1997). Narcissist leaders are mostly driven by the need for dominance and power (Rosenthal & Pittinsky, 2006). Another dimension of dark triad is Psychopathy. Psychopaths are known for displaying disinhibited behavior accompanied with lack of remorse and empathy. Psychopath leaders has been vastly studied in businesses literature under the title of corporate sociopaths, organizational psychopath etc. (Boddy, 2017; Westerlaken & Woods, 2013). It seems that psychopathic tendencies are more often found in higher managerial levels rather than junior workers (Boddy, Ladyshewsky, & Galvin, 2010); probably because these tendencies provide leaders with the sufficient charm and ease to exploit each opportunity to reach to the top (Babiak & Hare, 2006). Therefore, as the first hypothesis we test if power increases unethical behavior (H1). We would like to replicate the moderating effect of Machiavellianism on power and unethical behavior as those with higher level of Machiavellian orientation being more unethical in high-power positions (H2). We also expect that Narcissism and Psychopathy

simply follow the same pattern and positively moderates the link between power and unethical behavior (H3, H4).

Method

Participants. One hundred and nighty five participants were recruited from a Swiss university. On average, they were 21.5 years old and 57.6% of the participants were male. Participants were required to be native French speakers. They were informed that their payoff depended on their decision and the decisions of their group members but on average they gained an equivalent of 17 USD for their participation.

Procedure and Measures. The procedure is similar to the first study with some minor changes explained below. Participants came to the lab in large groups and were sat in front of the computer and randomly assigned the role of either the leader or the follower and start a two stage task. In the first stage of the experiment leaders were given 5 followers and 60 USD. All the leaders were given the endowment and they had to distribute this amount among themselves and their followers. Power was manipulated by the amount of control the leaders had over the distribution. In high-power condition, the leader was at the liberty to directly distribute the $60 endowment among him-/herself and the followers while in the low-power condition the leader was asked to press a button and a random generator decided about the distribution. In both conditions, the leaders received some pieces of information about the followers (e.g., age, gender, and the university they study at) and then, the endowment was distributed.

In the second stage, the leader was paired with another follower that was not involved in the first stage of the study. Actually the follower had been working to complete the slider task (used in the first stage of study 1). The follower had to adjust 5 sliders correctly and each slider is worth $2. If the follower adjusted all the sliders correctly he/she would earn 10 USD. However, the follower is informed that it is the leader who has the right to confirm this

transaction. The leader was given $10 as his/her wage then (s)he received what the follower earned by completing the slider task. Then, the norm was communicated to the leader (as the last study) and (s)he was asked to determine if (s)he wanted to take a part of the amount received from follower's earning. The difference between the power manipulation in this study and the former study is that the power manipulation is done in a way that the powerholder can benefit from their power as well.

To avoid income effects among the leaders in the high- and low-power condition and rule out any confounding effect, we made sure that leaders in both conditions received similar gains from the endowment in the first stage. To do so, we first ran the study with only high-power conditions. Then, we made a profile for all leaders in the high-power condition by considering their Dark Triad and their distribution decisions. To avoid income effect, for the low-power condition, when a leader was pressing the button to randomly distribute the endowment, an algorithm started to detect the closest profile of the leader in the high-power condition and then matched the earnings of the leaders in low-power condition to the leader in high-power condition. This way we could rule out the income effect on unethical behavior across the two conditions.

Leader's unethical behavior. We measured leader's unethical behavior in the same way as in Study 1 by the deviance from the accepted social norm; we considered any amount kept by the leader from the follower's endowment as unethical behavior. This variable is dichotomous. Any amount bigger than zero that the leader kept for him/herself was considered as an act of unethical behavior (M = .58, SD = .50).

Leader gain. The amount that leaders withheld from the followers' $10 was considered as the leader gain. There were so many leaders that did not take any part of the followers' endowment. Hence, we observed over dispersion (M = 3.063, SD = 3.36, *Variance*

= 11.27) and high level of number of zeros (42.5% zero) that we further will address in our analysis.

***Machiavellianism*.** After the leaders were appointed and before they started the game, they were asked to complete the 20-item Mach-IV scale (Christie & Geis, 1970) measuring their Machiavellian orientation. This scale contains cynical and positive views towards human nature (e.g., "*Generally speaking, men won't work hard unless they're forced to do so*" or "*Most people are brave*") and interpersonal tactics in accordance with these point of views (e.g., "*Honesty is the best policy in all cases*" or "*It is wise to flatter important people*"). All items are rated on a seven-point scale from 0 (strongly disagree) to 7 (strongly agree). We averaged leaders' ratings on the 20 dimensions to create a Machiavellianism score (Cronbach's Alpha = .76., *M* = 4.03; *SD* = 0.41).

***Narcissism*.** After the leaders were appointed and before they started the game, we used Short Dark Triad's Narcissism scale developed by Gamache, Savard, and Maheux-Caron (2018). Participants are asked to determine the extent to which they agree to 4 statements on a scale from 1 (completely disagree) to 7 (completely agree) such as "*I tend to want others to pay attention to me*". The items were averaged to obtain a score for Nercissism (Cronbach's Alpha = .77, *M* = 3.95; *SD* = 1.31).

***Psychopathy*.** We used Short Dark Triad's psychopathy scale developed by Gamache, Savard, and Maheux-Caron (2018) with 4 items. Participants are asked to determine the extent to which they agree to 4 statements on a scale from 1 (completely disagree) to 7 (completely agree) such as "*I tend to exploit others towards my own goal*". The reliability among the items were not high enough to be able to compile the items so we did not consider this variable in our analyses (Cronbach's Alpha = .55).

***Controls*.** We controlled for the age, gender as previous study but we added the faculty in which the participants are enrolled and the day of the experiment. Because participants who

are enrolled in the Faculty of Business and Economics are known to be more familiar with the structure of the economic games and tuned to think in terms of economic gain and costs; therefore they would more often display the optimum behavior in their final choices regardless of the contextual factors of the experiments. Moreover, the last day of the experiment was the very last day of the semester and the rate of unethical behavior surprisingly raised. Therefore, we controlled for the session of data collection in the last day of the semester.

Results

To test whether power predicts unethical behavior in leaders (H1), we used a Probit model estimating the effect of power on leaders' unethical behavior while controlling for leaders' age, gender, the faculty they study, and the day of the experiment. The results disconfirmed the main effect of power on unethical behavior ($\beta = -.15, p = .55$). That is, the probability of engaging in an unethical behavior for the leaders in high-power condition (*M* = .74; *SD* = .06) was not significantly different from the leaders in low-power condition (*M* = .76; *SD* = .05). But, when we entered the individual characteristics to the Probit model (adding Narcissism and Machiavellianism), we observed a main positive effect of Machiavellianism ($\beta = 0.96, p < .01$) and Narcissism ($\beta = 0.21, p = .03$) explaining the probability of unethical behavior (see table 4).

To see the effect of individual characteristics in combination with power, we interacted once power with Machiavellianism and once power with Narcissism to explain leader's unethical behavior. The results disconfirmed the interaction between power and Machiavellianism ($\beta = -0.77,\ p < .24$). It seems that the effect of Machiavellianism as a moderator in previous study is now concentrated in a main effect of power with self-serving opportunities on unethical behavior. Machiavellians in situations with self-serving opportunities indulge their position regardless of the power level they are granted.

But, Narcissism positively interacted with power explaining leader's unethical behavior ($\beta = 0.39,\ p < .04$). A simple slop analysis showed that Narcissism positively and significantly predicted unethical behavior of the leaders in high-power condition, β = 0.42, p = .01; but this link is non-significant in low-power condition, β = .11, p = .45; $\chi 2 = 4.36,\ p = .04$ (see table 5, 6, and figure 2).

Additional analyses. We also tested whether high-power leaders took more money from their followers. The amount that leaders took from their followers in the second stage was considered as the leader gain. An OLS regression on leader gain controlled for leaders' age, gender, the faculty in which they study, and the sessions of the experiment showed no effect of power on the amount the leader withheld (β = .57, p = .41). But when we added both individual characteristics, Machiavellianism positively influenced leader gain, β = 2.08, p < .01. But we did not find a main effect of Narcissism (β = .31, p = .26).

We interacted the individual characteristics separately and in combination with power while having considered the same control variables to explain leader gain. The results disconfirmed an interaction effect between power and Machiavellianism explaining but, Narcissism positively interacted with power explaining leader gain ($\beta = 1.07,\ p = .03$). Higher level of Narcissism explained leader gain in high-power positions ($\beta = 0.83,\ p = .025$) but not in the low-power positions ($\beta = -0.02,\ p = .96;\ \chi 2 = 4.36,\ p = .04$).

To rule out the effect of leader's behavior (income from the distribution) in the first stage on leaders gain in the second stage we used an instrumental variable regression. We regressed leader gain in the second stage on the amount of money the leaders took for themselves in the first stage using leader's power, Narcissism, Machiavellianism, age, gender, the faculty in which they study, and the sessions of experiment as the instruments. We found no significant effect of the amount received by the leader in the first stage on leader gain in the second stage ($\beta = .07,\ p = .12$). In our analyses we excluded Narcissism and the last

day of experiment from the second stage of the regression and the result of the *F*-test on the power of excluded instruments were in the accepted range, $F(2,117) = 10.50$, $p < 0.01$ (Stock & Yogo, 2002). The Sargan chi-square test of over-identification also suggested that the instruments are valid, $p = .75$ (Sargan, 1958).

Post-test. To ensure that our manipulation of power was robust and to better understand what triggers the different outcomes in Study 1 and Study 2 we conducted a post-test. In this post-test, we intended to see if the power manipulation in our two studies induced different feeling in individuals in their power role. We recruited 250 participants in an online platform (Mturk) and randomly assigned them to a 2 (high-power vs. low-power) by 2 (with self-serving opportunity vs. without self-serving opportunity) design. We described the leaders' roles in our two studies. Then, we asked the participants to report their feelings if they were in the shoes of the leaders. Participants had to rate to what extent they felt for example, responsible, opportunistic, authoritarian, and powerful. We found that participants in power roles with self-serving opportunity, felt much more opportunistic in comparison to the participants who had no self-serving opportunity, $F(1, 248) = 11.17$, $p < .01$. This could explain why individuals on average stole much more in self-serving conditions when compared to no self-serving conditions, $\chi 2(1) = 6.75$, $p < .01$. Leaders in power conditions with no self-serving opportunity less often stole from their followers ($M = 0.58$, $SD = 0.50$) compared to the leaders in self-serving conditions ($M = 0.75$, $SD = 0.43$). Same pattern of results replicated for the amount that leaders stole from their followers. In conditions that power was induced with self-serving opportunities the leaders took much more from their yet unknown followers in the second stage compared to the leaders in the conditions without self-serving opportunity, $F(1, 201) = 8.06$, $p < .01$.

Discussion

In this study, we added an additional element (self-serving opportunity) to our power manipulation by letting high-power leaders to benefit from their position. Although the leaders could benefit from their power this time, our results again disproved the corruptive effect of power. One reason could simply be that "power does not corrupt". It is individuals who corrupt power. We found that individual characteristics interacted with levels of power and explained leaders' unethical behavior. More specifically, Narcissism increased the chance of leader's unethical behavior in high-power positions compared to low power positions. Moreover, a significant main effect of Machiavellianism and Narcissism were demonstrated on leader's unethical behavior. These results imply that power can have a revealing effect: in high-power positions with self-serving opportunity, Narcissism increases the chance and the intensity of unethical behavior (amount leaders took from their followers). Those individuals who are prone to exploit others for personal gains (=Narcissists), when appointed to high-power position, more often engage in unethical behavior.

Narcissists are accompanied with a grandiose sense of self-importance and entitlement. It seems that being the sole decision maker of a group of individuals reinforced their belief of self-importance. Therefore, they might engage more often in degrading others and self-serving behavior. The fact that high-power leaders had the control over others' outcomes can boost narcissists' sense of grandiosity and fed their need for dominance and control. With such boosted self-entitlement, narcissist's exploitativeness could be activated and thus possibly lead to degrading others' rights and further aggressions (Reidy, Zeichner, Foster, & Martinez, 2008). In the same vein, our results showed that Narcissist in the high power roles with self-serving opportunities were more likely to exploit their yet unknown followers.

Moreover, a recent research on the link between Narcissism and reduced generosity shows the mediating effect of lower perspective taking skills (Böckler, Sharifi, Kanske,

Dziobek, & Singer, 2017). We know that power in general, diminishes perspective taking (Galinsky, Magee, Inesi, & Gruenfeld, 2006). So in high-power positions, Narcissists are doubly affected by the impaired perspective taking, and therefore tend to exploit their counterparts more often. As Glad (2002) argued, a malign Narcissist generally engage in self-serving behavior up to certain boundaries that the social context allows; but when absolute power is present all sorts of grandiose fantasies are attainable.

Summary and Concluding Discussion

In two studies we investigated the effect of power and individual characteristics on leaders' unethical behavior. In the first study we showed that Machiavellian and in the second study Narcissism interacted with power and determined the propensity for unethical behavior. Our findings are in line with the claim of the researchers pinpointing the revealing effect of power (Chen et al., 2001; Galinsky, et al., 2015; Overbeck & Droutman, 2013). Indeed, powerholders act more in accordance with their traits and attitudes and set out to indulge their current needs because power makes individuals more focused on the self, one's own goals, and needs (Guinote, 2007; Rucker, Galinsky, & Dubois, 2012).

The present work did not replicate the results of the former research on the corruptive nature of power. In both studies, power was manipulated in a structural format. It has been shown that different operationalizations of power led to different behavioral outcomes (Keltner et al., 2003; Smith & Bargh, 2008; Tost & Johnson, 2019). For instance, most of the studies that project the dark outcomes of power such as lying, corruptive, and opportunistic behavior (e.g., Galinsky, Magee, Inesi, & Gruenfeld, 2006; Gruenfeld, et al., 2008; Lammers, et al., 2012) used methods that manipulated power via power priming (Tost, 2015). Whereas, most studies that unveil the bright side of power induced structural power (e.g., Tost, Wade-Benzoni, & Johnson, 2015; Tost & Johnson, 2019; Wisse & Rus, 2012). A theoretical reason behind these controversial results is explained by Tost (2015): structural manifestation of

power is more associated with responsibility and accountability; while psychological power (induced mostly by recalling task) is more linked to opportunistic and competitive side of power (e.g., Overbeck & Park, 2001, 2006; Schmid Mast, et al., 2009; Scholl, Sassenberg, Scheepers, Ellemers, & de Wit, 2017; Tost, Wade-Benzoni, & Johnson, 2015). Moreover, another study showed that various self-construal of powerholders (a collective self-construal as opposed to personal self-construal) predicted leader's self-interested behavior (Wisse & Rus, 2012). In our study as power was induced in a structural setting, powerholders could hold a rather collective construal of power which led to a non-significant effect of "power corrupts". Future research might want to investigate the effect of different construal of power (responsibility as oppose to opportunity) on unethical behavior as it seems that different construals of the self and one's own power have a major influence on powerholders' behavioral outcomes. It is important to be particularly vigilant on the ways power is manipulated and what construal of the self and the power position we convey by our designs.

Another reason for the non-significant effect of power on unethical behavior could be that in our design the high- and low-power conditions were located in the relatively powerful range. In other words, leaders in both high- and low-power positions -although significantly different- felt rather powerful in their positions. In fact, what we had in our study could be seen as a comparison between different levels of high power. Although we designed our control condition in a way that conveyed the least amount of power to the individuals (having no control over the outcomes of the followers), it could be that in our design all the leaders either in high-power or control condition -although significantly different- felt rather powerful in their positions. It is debatable if a better control condition was used, we could have detected the corruptive effect of power. Indeed, one challenge in power studies is defining a proper control condition especially when power is manipulated objectively and in a structural format.

Future research should try to replicate our finding using a design with clearer distinction between the high and neutral power conditions.

Moreover the interaction effect between Machiavellianism and power was not replicated in the second study. Several factors could explain why Machiavellian leaders in high-power positions with self-serving opportunity did practice self-indulgence similar to leaders in low-power positions. From a theoretical point of view, Machiavellians in a morally obscure situation are thought to be less concerned about the (un)ethical aspect of their actions and tend to behave in a way that guarantees their desired outcome (Bass, Barnett, & Brown, 1999). More specifically, they more often act towards distrust, amoral manipulations, and devaluation of ethical behavior when it comes to personal gain (Dahling, Whitaker, & Levy, 2009). In Study 2, we gave the leaders self-serving opportunities to take a part or the whole amount of an endowment for themselves. That self-serving opportunity could have facilitated the tendencies of the Machiavellians to indulge in both high- and low- power, thus change the moderating effect of power in study 1 to a strong main effect in Study 2. We basically showed that once Machiavellians are in power positions with self-serving opportunities, they tend exert unethical behavior regardless of the amount of power given to them. These results are alarming because it has been shown that Machiavellians are in fact more attracted to power and hence, tend to emerge as leaders more often (Corzine, 1997; Dahling, Whitaker, & Levy, 2009; Fehr, Samson, & Paulhus, 1992; Mael, Waldman, & Mulqueen, 2001). Personnel selection for high-power positions becomes therefore crucial, as it seems that "*the corrupts go for power rather than that power corrupts*" (Van Prooijen & Van Lange, 2014, p 84).

Our study also benefits from manipulating objective power (with leaders having real followers and control over their valued resources). That, to a great extent, protects the design from possible confounds and demand effects (Strum & Antonakis, 2015). However, our power manipulation had some limitations as well. Because we set out to avoid deceptions

towards the participants, we incorporated real followers and real money in the study design which requires very large resources and sample size. Moreover, our baseline condition could be selected in a way to better distinguish high-power and neutral power condition. For example adding an observer to the leader-follower relation or a neutral player could be an alternative control conditions. Future research should try to replicate our findings using a rather efficient design with a proper control condition.

References

Adams, J. S. (1965). Inequity in social exchange. *Advances in experimental social psychology*, *2*, 267-299.

Adams, J. M., Florell, D., Burton, K. A., & Hart, W. (2014). Why do narcissists disregard social-etiquette norms? A test of two explanations for why narcissism relates to offensive-language use. *Personality and Individual Differences*, *58*, 26-30.

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

Aquino, K., & Reed II, A. (2002). The self-importance of moral identity.*Journal of Personality and Social Psychology*, *83*(6), 1423-1440.

Babiak P, Hare RD. 2006. *Snakes in Suits When Psychopaths go to Work*. Regan Books: New York.

Bandura A. (1990). Mechanisms of moral disengagement. In Reich W (Ed.), *Origins of terrorism: Psychologies, ideologies, states of mind* (pp. 161–191). New York, NY: Cambridge University Press.

Bargh, J. A., & Raymond, P. (1995). The naive misuse of power: Nonconscious sources of sexual harassment. *Journal of Social Issues*,*51*(1), 85-96.

Bargh, J. A., Raymond, P., Pryor, J. B., & Strack, F. (1995). Attractiveness of the underling: An automatic power→ sex association and its consequences for sexual harassment and aggression. *Journal of personality and social psychology*, *68*(5), 768-781.

Bedell, K., Hunter, S., Angie, A., & Vert, A. (2006). A historiometric examination of Machiavellianism and a new taxonomy of leadership. *Journal of Leadership & Organizational Studies*, *12*(4), 50-72.

Bendahan, S., Zehnder, C., Pralong, F. P., & Antonakis, J. (2015). Leader corruption depends on power and testosterone. *The Leadership Quarterly*,*26*(2), 101-122.

Bergman, R. (2002). Why be moral? A conceptual model from developmental psychology. *Human Development*, *45*(2), 104-124.

Blader, S. L., Shirako, A., & Chen, Y. R. (2016). Looking Out From the Top Differential Effects of Status and Power on Perspective Taking. *Personality and Social Psychology Bulletin*, *42*(6), 723-737.

Blevins, D. P., Tsang, E. W., & Spain, S. M. (2015). Count-based research in management: suggestions for improvement. *Organizational Research Methods*, *18*(1), 47-69.

Böckler, A., Sharifi, M., Kanske, P., Dziobek, I., & Singer, T. (2017). Social decision making in narcissism: Reduced generosity and increased retaliation are driven by alterations in perspective-taking and anger. *Personality and Individual Differences, 104*, 1–7.

Boddy, C. R., Ladyshewsky, R., & Galvin, P. (2010). Leaders without ethics in global business: Corporate psychopaths. *Journal of Public Affairs*, *10*(3), 121-138.

Burns, J. M. 1978. *Leadership*. New York: Harper & Row.

Campbell, W. K., Bush, C. P., Brunell, A. B., & Shelton, J. (2005). Understanding the social costs of narcissism: The case of the tragedy of the commons. *Personality and Social Psychology Bulletin, 31*(10), 1358–1368. doi: 10.1177/0146167205274855.

Campbell, W. K., Reeder, G. D., Sedikides, C., & Elliot, A. J. (2000). Narcissism and comparative self-enhancement strategies. *Journal of Research in Personality, 34*(3), 329-347.

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.

Christie, R. , & Geis, F. L. (1970). *Studies in Machiavellianism*. Academic Press.

Cook, K. S., & Hegtvedt, K. A. (1983). Distributive justice, equity, and equality. *Annual Review of Sociology, 9*(1), 217-241.

Corzine, J. B. (1997). Machiavellianism and management: A review of single-nation studies exclusive of the USA and cross-national studies. *Psychological Reports, 80*(1), 291-304.

Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (Vol. 32). CRC Press.

Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management*, *35*(2), 219-257.

Dalbert, C., Montada, L., & Schmitt, M. (1987). Belief in a just world: Validation correlates of two scales. *Psychologische Beitrage*, *29*(4), 596-615.

DeCelles, K. A., DeRue, D. S., Margolis, J. D., & Ceranic, T. L. (2012). Does power corrupt or enable? When and why power facilitates self-interested behavior. *Journal of Applied Psychology*, *97*(3), 681-689.

De Vries, M. F. K., & Miller, D. (1985). Narcissism and leadership: An object relations perspective. *Human Relations*, *38*(6), 583-601.

DeWall, C. N., Baumeister, R. F., Mead, N. L., & Vohs, K. D. (2011). How leaders self-regulate their task performance: evidence that power promotes diligence, depletion, and disdain. *Journal of Personality and Social Psychology*, *100*(1), 47-65.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2015). Social class, power, and selfishness: When and why upper and lower class individuals behave unethically. *Journal of Personality and Social Psychology*, *108*(3), 436-449.

Emmons, R. 1997. Narcissism: theory and measurement. *Journal of Personality and Social Psychology,* 52(1): 11–17.

Engel, C. (2011). Dictator games: a meta-study. *Experimental Economics*, *14*(4), 583-610.

Fehr, B., Samson, D., & Paulhus, D. L. (1992). The construct of Machiavellianism: Twenty years later. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment* (Vol. 9, pp. 77–116). Hillsdale: Erlbaum.

Finkelstein, S. (1992). Power in top management teams: Dimensions, measurement, and validation. *Academy of Management Journal*, *35*(3), 505-538.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*(6), 621-628.

Flynn, F. J., Gruenfeld, D., Molm, L. D., & Polzer, J. T. (2011). Social psychological perspectives on power in organizations. *Administrative Science Quarterly*, *56*(4), 495-500.

French, J. R., Raven, B., & Cartwright, D. (1959). The bases of social power. *Classics of Organization Theory*, 311-320.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453-466.

Galinsky, A. D., Magee, J. C., Gruenfeld, D. H., Whitson, J. A., & Liljenquist, K. A. (2008). Power reduces the press of the situation: implications for creativity, conformity, and dissonance. *Journal of Personality and Social Psychology*, *95*(6), 1450-1466.

Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, *17*(12), 1068-1074.

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 handbooks in psychology. APA handbook of personality and social psychology, Vol. 3. Interpersonal relations* (pp. 421–460). Washington, DC: American Psychological Association.

Gert, B., & Gert, J., (2016).The Definition of Morality. In *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/spr2016/entries/morality-definition .

Glad, B. (2002). Why tyrants go too far: Malignant narcissism and absolute power. *Political Psychology, 23*(1), 1-2.

Grapsas, S., Brummelman, E., Back, M. D., & Denissen, J. J. A. (2020). The "Why" and "How" of Narcissism: A Process Model of Narcissistic Status Pursuit. *Perspectives on Psychological Science, 15*(1), 150–172.

Greenbaum, R. L., Hill, A., Mawritz, M. B., & Quade, M. J. (2017). Employee Machiavellianism to unethical behavior: The role of abusive supervision as a trait activator. *Journal of Management, 43*(2), 585-609.

Grijalva, E., Harms, P. D., Newman, D. A., Gaddis, B. H., & Fraley, R. C. (2015). Narcissism and leadership: A meta-analytic review of linear and nonlinear relationships. *Personnel Psychology, 68*(1), 1-47.

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.

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

Guinote, A. (2007b). Behavior variability and the situated focus theory of power. *European Review of Social Psychology, 18*(1), 256-295.

Guinote, A., Judd, C. M., & Brauer, M. (2002). Effects of power on perceived and objective group variability: evidence that more powerful groups are more variable. *Journal of Personality and Social Psychology, 82*(5), 708-721.

Hall, J. A., & Bernieri, F. J. (Eds.). (2001). *Interpersonal Sensitivity: Theory and Measurement.* Psychology Press.

Hall, J. A., Mast, M. S., & Latu, I. M. (2015). The vertical dimension of social relations and accurate interpersonal perception: A meta-analysis. *Journal of Nonverbal Behavior, 39*(2), 131-163.

Hirsh, J. B., Galinsky, A. D., & Zhong, C. B. (2011). Drunk, powerful, and in the dark how general processes of disinhibition produce both prosocial and antisocial behavior. *Perspectives on Psychological Science, 6*(5), 415-427.

Hoogervorst, N., De Cremer, D., van Dijke, M., & Mayer, D. M. (2012). When do leaders sacrifice?: The effects of sense of power and belongingness on leader self-sacrifice. *The Leadership Quarterly, 23*(5), 883-896.

Joshi, P. D., & Fast, N. J. (2013). Power and reduced temporal discounting. *Psychological Science*, *24*(4), 432-438.

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*(2), 265-284.

Kets de Vries, M.F.R. and Miller, D. (1985) Narcissism and leadership: an object relations perspective, *Human Relations, 38*(4), pp. 583–601.

Kiazad, K., Restubog, S. L. D., Zagenczyk, T. J., Kiewitz, C., & Tang, R. L. (2010). In pursuit of power: The role of authoritarian leadership in the relationship between supervisors' Machiavellianism and subordinates' perceptions of abusive supervisory behavior. *Journal of Research in Personality*, *44*(4), 512-519.

Lammers, J., Galinsky, A. D., Dubois, D., & Rucker, D. D. (2015). Power and morality. *Current Opinion in Psychology, 6*, 15-19.

Lammers, J., Galinsky, A. D., Gordijn, E. H., & Otten, S. (2012). Power increases social distance. *Social Psychological and Personality Science, 3*(3), 282-290.

Lammers, J., & Maner, J. (2016). Power and attraction to the counternormative aspects of infidelity. *The Journal of Sex Research*, *53*(1), 54-63.

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

Lammers, J., Stapel, D. A., & Galinsky, A. D. (2010). Power increases hypocrisy moralizing in reasoning, immorality in behavior. *Psychological Science, 21*(5), 737-744.

Lammers, J., Stoker, J. I., Jordan, J., Pollmann, M., & Stapel, D. A. (2011). Power increases infidelity among men and women. *Psychological Science, 22*(9), 1191-1197.

Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Gallucci, A. (2013). Sex, power, and money: Prediction from the Dark Triad and Honesty–Humility. *European Journal of Personality*, *27*(2), 169-184.

Lerner, M. J. (1980). *The belief in a just world: A fundamental delusion.* Plenum Press.

Long, J. S. 1997. *Regression models for categorical and limited dependent variables*. Sage Publications.

Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata*. Stata press.

Lu, J. G., Zhang, T., Rucker, D. D., & Galinsky, A. D. (2018). On the distinction between unethical and selfish behavior. In K. Gray, J. Graham (Eds.), *Atlas of moral psychology: Mapping good and evil in the mind (*pp. 465-474*)*. New York, NY: Guilford Press.

Machiavelli, N. (1966). *The prince*. New York: Bantam Books.

Mael, F. A., Waldman, D. A., & Mulqueen, C. (2001). From scientific work to organizational leadership: Predictors of management aspiration among technical personnel. *Journal of Vocational Behavior*, *59*(1), 132-148.

Magee, J. C., & Galinsky, A. D. (2008). 8 Social Hierarchy: The Self-Reinforcing Nature of Power and Status. *The Academy of Management Annals*, *2*(1), 351-398.

Magee, J. C., & Smith, P. K. (2013). The social distance theory of power. *Personality and Social Psychology Review*, *17*(2), 158-186.

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

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

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science*, *12*(2), 183-204.

Overbeck, J. R., & Droutman, V. (2013). One for all: Social power increases self-anchoring of traits, attitudes, and emotions. *Psychological Science*, *24*(8), 1466-1476.

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.

Overbeck, J. R., Tiedens, L. Z., & Brion, S. (2006). The powerful want to, the powerless have to: Perceived constraint moderates causal attributions. *European Journal of Social Psychology*, *36*(4), 479-496.

Piff, P. K., Stancato, D. M., Côté, S., Mendoza-Denton, R., & Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proceedings of the National Academy of Sciences*, *109*(11), 4086-4091.

Rayburn, J. M., & Rayburn, L. G. (1996). Relationship between Machiavellianism and type A personality and ethical-orientation. *Journal of Business Ethics*, *15*(11), 1209-1219.

Reidy, D. E., Zeichner, A., Foster, J. D., & Martinez, M. A. (2008). Effects of narcissistic entitlement and exploitativeness on human physical aggression. *Personality and Individual Differences*, *44*(4), 865-875.

Rosenthal, S.A. and Pittinsky, T.L. (2006) Narcissistic Leadership, *Leadership Quarterly, 17*(6), pp. 617–633.

Rucker, D. D., Galinsky, A. D., & Dubois, D. (2012). Power and consumer behavior: How power shapes who and what consumers value. *Journal of Consumer Psychology*, *22*(3), 352-368.

Ruegger, D., & King, E. W. (1992). A study of the effect of age and gender upon student business ethics. *Journal of Business ethics*, *11*(3), 179-186.

Sargan, J. D. (1958). The Estimation of Economic Relationships Using Instrumental Variables. *Econometrica*, *26*, 393-415.

Sassenberg, K., Ellemers, N., & Scheepers, D. (2012). The attraction of social power: The influence of construing power as opportunity versus responsibility. *Journal of Experimental Social Psychology*, *48*(2), 550-555.

Sassenberg K., Ellemers N., Scheepers D., Scholl A. (2014). "Power corrupts" revisited: the role of construal of power as opportunity or responsibility. In J. W. Van Prooijen & P. Van Lange (Eds.). *Power, politics, and paranoia: Why people are suspicious of their leaders*, (pp. 73–87). Cambridge, UK: Cambridge Univ. Press

Schaerer, M., Lee, A. J., Galinsky, A. D., & Thau, S. (2018). Contextualizing social power research within organizational behavior. In D. L. Ferris, R. E. Johnson, & C. Sedikides (Eds.), *SIOP organizational frontiers series. The self at work:*

Fundamental theory and research (pp. 194–221). Routledge: Taylor & Francis Group.

Schmid Mast, M. S. (2010). Interpersonal behavior and social perception in a hierarchy: The interpersonal power and behavior model. *European Review of Social Psychology, 21*(1), 1-33.

Schmid Mast, M., Jonas, K., & Hall, J. A. (2009). Give a person power and he or she will show interpersonal sensitivity: the phenomenon and its why and when. *Journal of Personality and Social Psychology*, *97*(5), 835-850.

Schmid Mast, M., Khademi, M., & Palese, T. (2020). Power and social information processing. Cur*rent Opinion in Psychology, 33*, 42-46.

Scholl A., Ellemers N., Sassenberg K., Scheepers D. (2015) Understanding Power in Social Context: How Power Relates to Language and Communication in Line with Responsibilities or Opportunities. In R. Schulze & H. Pishwa (Eds.). *The exercise of power in communication* (PP. 312-334). Palgrave Macmillan, London.

Scholl, A., Sassenberg, K., Scheepers, D., Ellemers, N., & de Wit, F. (2017). A matter of focus: Power-holders feel more responsible after adopting a cognitive other-focus, rather than a self-focus. *British Journal of Social Psychology*, *56*(1), 89-102.

Slabu, L., & Guinote, A. (2010). Getting what you want: Power increases the accessibility of active goals. *Journal of Experimental Social Psychology, 46*(2), 344-349.

Smith, P. K., & Bargh, J. A. (2008). Nonconscious effects of power on basic approach and avoidance tendencies. *Social Cognition, 26*(1), 1-24.

Stinson, F. S., Dawson, D. A., Goldstein, R. B., Chou, S. P., Huang, B., Smith, S. M., ... & Grant, B. F. (2008). Prevalence, correlates, disability, and comorbidity of DSM-IV narcissistic personality disorder: results from the wave 2 national epidemiologic

survey on alcohol and related conditions. *The Journal of Clinical Psychiatry*, *69*(7), 1033.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, *20*(4), 518-529.

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

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin*, *146*(1), 30.

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.

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

Tost, L. P., & Johnson, H. H. (2019). The prosocial side of power: How structural power over subordinates can promote social responsibility. *Organizational Behavior and Human Decision Processes*, *152*, 25-46.

Tost, L. P., Wade-Benzoni, K. A., & Johnson, H. H. (2015). Noblesse oblige emerges (with time): Power enhances intergenerational beneficence. *Organizational Behavior and Human Decision Processes*, *128*, 61-73.

Van Kleef, G. A., De Dreu, C. K., Pietroni, D., & Manstead, A. S. (2006). Power and emotion in negotiation: Power moderates the interpersonal effects of anger and happiness on concession making. *European Journal of Social Psychology*, *36*(4), 557-581.

Van Kleef, G. A., Oveis, C., Van Der Löwe, I., LuoKogan, A., Goetz, J., & Keltner, D. (2008). Power, distress, and compassion: Turning a blind eye to the suffering of others. *Psychological Science, 19*(12), 1315-1322.

Van Prooijen, J. W., & Van Lange, P. A. (Eds.). (2014). *Power, politics, and paranoia: Why people are suspicious of their leaders*. Cambridge University Press.

Weber, M. (2009). *The theory of social and economic organization*. Simon and Schuster Inc.

Westerlaken, K. M., & Woods, P. R. (2013). The relationship between psychopathy and the Full Range Leadership Model. *Personality and Individual Differences, 54*(1), 41-46.

Williamson, T. (2008). The good society and the good soul: Plato's Republic on leadership. *The Leadership Quarterly, 19*(4), 397-408.

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment, 88*(2), 205-219.

Wisse, B., & Rus, D. (2012). Leader self-concept and self-interested behavior: The moderating role of power. *Journal of Personnel Psychology, 11*(1), 40-48.

Wisse, B., & Sleebos, E. (2016). When the dark ones gain power: Perceived position power strengthens the effect of supervisor Machiavellianism on abusive supervision in work teams. *Personality and Individual Differences, 99*, 122-126.

Table 1

Probit Regression Estimates Predicting Leader's Unethical Behavior from Power

	Unethical Behavior
Age	-.02(0.39)
Gender	-.27(0.94)
Follower's performance	-.02(0.11)
Power	.15(0.51)
Constant	.59 (0.43)
Psudo R-sq	0.017

Note. N= 80; Unstandardized beta coefficients are reported along with robust z-statistics in parentheses; pseudo R-square is based on the Cox and Snell (1989) method; power dummy codded as 2 = High-power and 1=Low-power; gender dummy coded as 0 = female and 1 = male;* $p < .05$**$p < .001$.

Table 2

Probit Regression Estimates Predicting Leader's Unethical Behavior from Power and Leader's Machiavellianism

Model	(1)	(2)
Age	-.005(0.11)	-.03(0.63)
Gender	-.46(1.52)	-.45(1.38)
Follower's performance	-.06(0.29)	-.13(0.54)
Power	.17(0.57)	4.99**(2.53)
Machiavellianism	.66**(2.54)	.08(0.22)
Power× Machiavellianism		1.55**(2.67)
Constant	1.47(.88)	1.30(0.67)
Pseudo R-sq	0.087	0.159

Note. N=80; Unstandardized beta coefficients are reported along with robust z-statistics in parentheses; pseudo R-square is based on the Cox and Snell (1989) method; power dummy codded as 2 = High-power and 1=Low-power; gender dummy coded as 0 = female and 1 = male; ** $p < .05$ * $p < .1$.

Table 3

Probit Model Estimates Predicting Leader's Unethical Behavior from Power and Leader's Moral Disengagement

Model	(1)	(2)
Age	-.002(0.04)	-.01(0.27)
Gender	-.38(1.26)	-.42(1.34)
Follower's performance	-.003(0.01)	-.02(0.07)
Power	.20(0.69)	-1.06(0.99)
Moral Disengagement	.26(1.44)	.08(0.38)
Power× Moral Disengagement		.44(1.21)
Constant	-.61(0.38)	.41(0.25)
Pseudo R^2	.043	.062

Note. *N*= 80; Unstandardized beta coefficients are reported along with robust z-statistics in parentheses; pseudo R-square is based on the Cox and Snell (1989) method; power dummy codded as 2 = High-power and 1=Low-power; gender dummy coded as 0 = female and 1 = male; ** $p < .05$, * $p < .1$.

Table 4

Probit Regression Estimating Leader's Unethical Behavior from Power, Machiavellianism, Narcissism, and Control Variables

	(1)	(2)
Age	-.01(0.20)	.01(0.20)
Gender	-.03(0.13)	-.17(0.62)
Faculty	.36(1.20)	.03(0.10)
Sessions	-.25(0.69)	-.06(0.15)
Power	-.15(0.60)	-.003(0.01)
Machivellianism		.96**(2.83)
Narcissism		.21**(2.20)
Constant	.87(1.08)	-4.01**(2.46)
Psudo R-sq	.016	.104

Note. N= 128; Unstandardized beta coefficients are reported along with robust z-statistics in parentheses; pseudo R-square is based on the Cox and Snell (1989) method; power dummy codded as 1 = High-power and 0=Low-power; gender dummy coded as 0 = female and 1 = male; ** $p < .05$, * $p < .1$.

Table 5

Probit Regression Estimating Leader's Unethical Behavior from Power and Leader's Machiavellianism

Model	(1)	(2)
Age	-.01 (0.24)	-.01(0.25)
Gender	-.07(0.28)	-.09(0.32)
Faculty	.21(0.67)	.19(0.59)
Sessions	-.06(0.17)	-.02(.06)
Power	-.04(0.15)	3.0(1.16)
Machiavellianism	.96**(3.03)	1.39**(2.74)
Power× Machiavellianism		-.77(1.17)
Constant	-2.95**(2.00)	-4.67**(2.28)
Pseudo R-sq	0.074	0.082

Note. *N*= 128; Unstandardized beta coefficients are reported along with robust z-statistics in parentheses; pseudo R-square is based on the Cox and Snell (1989) method; power dummy codded as 1 = High-power and 0=Low-power; gender dummy coded as 0 = female and 1 = male; ** $p < .05$, * $p < .1$.

Table 6

Probit Model Estimating Leader's Unethical Behavior from Power and Leader's Narcissism

Model	(1)	(2)
Age	-.01 (0.23)	-.01(0.15)
Gender	-.15(0.54)	-.17(0.62)
Faculty	.20(0.63)	.18(0.56)
Sessions	-.24(0.65)	-.24(.66)
Narcissism	.22**(2.26)	.06 (0.47)
Power	-.12(0.48)	-1.60**(2.18)
Power × Narcissism		.39**(2.09)
Constant	-.19 (0.22)	.49(0.53)
Pseudo R-sq	.049	.075

Note. N= 128; Unstandardized beta coefficients are reported along with robust z-statistics in parentheses; pseudo R-square is based on the Cox and Snell (1989) method; power dummy codded as 1 = High-power and 0=Low-power; gender dummy coded as 0 = female and 1 = male; ** $p < .05$, * $p < .1$.

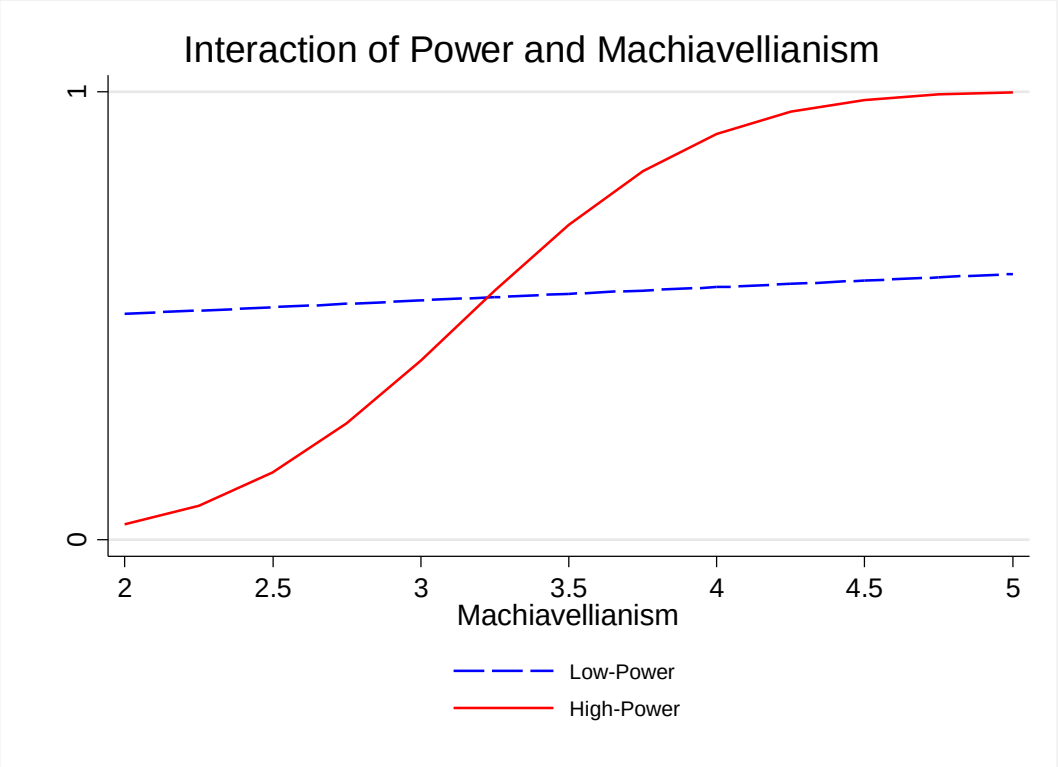


Figure 1. The interaction effect between power and Machiavellianism on unethical behavior.

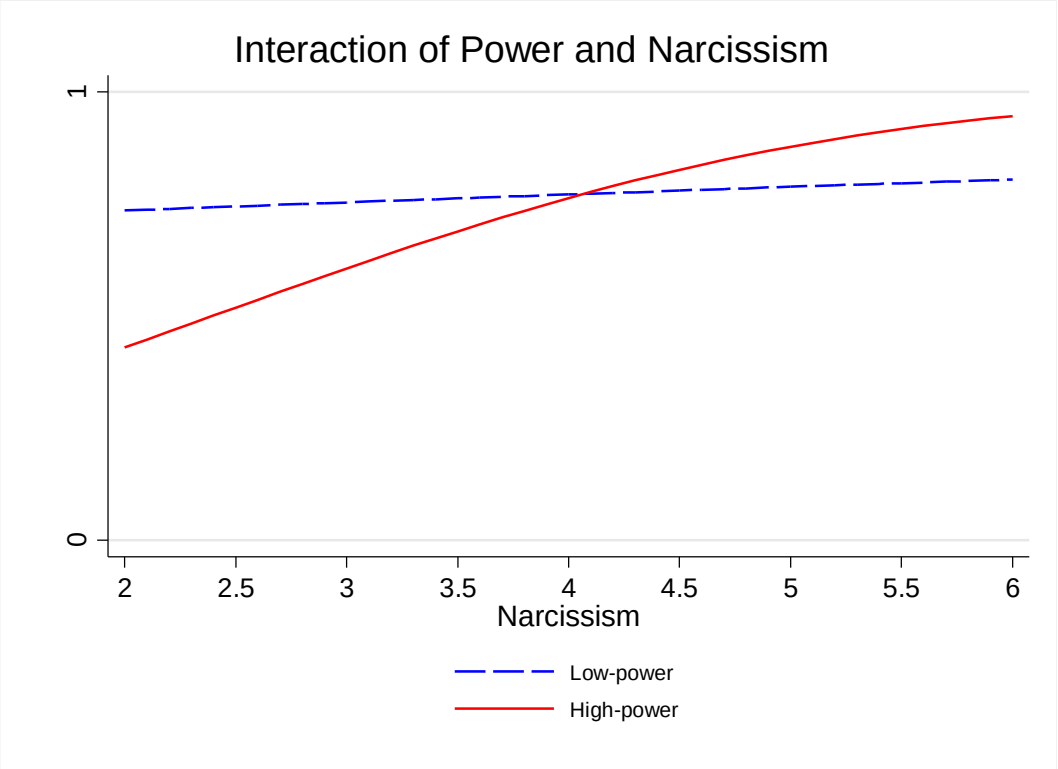


Figure 2. The interaction effect between power (with self-serving opportunity) and Machiavellianism on unethical behavior.

From Hierarchical to Egalitarian: Hierarchy Steepness Depends on Speaking Time Feedback and Task Interdependence

Mahshid Khademi, Marianne Schmid Mast, and Denise Frauendorfer

Faculty of Business and Economics, University of Lausanne, Switzerland

Author Note

Mahshid Khademi, Marianne Schmid Mast, and Denise Frauendorfer, Department of Organizational Behavior, Faculty of Business and Economics, University of Lausanne, Switzerland.

This work was supported by HEC Fund for Research under the Grant number 26051319. Results of this study previously presented at the 15th Swiss Psychological Society conference and 19th Eawop Congress, and presently accepted in the journal of "Group Dynamics: Theory, Research, and Practice".

Abstract

An individual's speaking time behavior in groups is influenced by many elements. We investigate whether giving feedback on speaking time about how much each group member talks and whether group task interdependence influence hierarchy emergence in problem solving groups. We also aim to investigate how these effects emerge by looking at whether initial over- and under-participators increase or decrease their speaking time contribution. Using an incentivized experimental design, we randomly assigned 405 participants to same-gender 3-person groups (70 all-female and 65 all-male groups). Then groups received either individual or group speaking time feedback and were either confronted with solving a relatively high or low interdependent task. Results show that groups that received individual speaking time feedback built flatter hierarchies, mostly due to over-participators decreasing their speaking time, and that in groups faced with a rather interdependent task, steeper hierarchies emerged, mostly due to under-participators decreasing their speaking time. This study highlights the importance of taking into account that environmental factors such as making salient each individuals speaking time or task type (i.e., degree of interdependence) affect how hierarchies emerge.

Keywords: Hierarchy steepness, speaking time feedback, task interdependence, team communication, small group

From Hierarchical to Egalitarian: Hierarchy Steepness Depends on Speaking Time Feedback and Task Interdependence

When individuals interact in so-called leaderless groups, meaning groups with no initial formal hierarchy, the distribution of the individual speaking times becomes less egalitarian over time and a dominance hierarchy emerges (Bales, 1970; Bales, Strodtbeck, Mills, & Roseborough, 1951; Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013; Gatica-Perez, 2009; Jayagopi, Hung, Yeo, & Gatica-Perez, 2009; Sanchez-Cortes, Aran, Schmid Mast, & Gatica-Perez, 2011; Schmid Mast, 2001). Such emerging hierarchies differ in how steep they are, meaning that the individual differences in speaking time among group members can be more or less pronounced. The steepness of a hierarchy can be operationalized as the standard deviation of the speaking time among the group members (Schmid Mast, 2002). If the standard deviation is high, group members differ in their individual speaking time contributions (steeper hierarchy) more so than if the standard deviation is low (flatter hierarchy). We set out to shed light on whether and how speaking time feedback and task type affect hierarchy steepness emergence in problem-solving groups.

Speaking Time Feedback and Hierarchy

Research has shown that when leaderless groups are provided with individual speaking time feedback, group members form less steep hierarchies (Bachour, Kaplan, & Dillenbourg, 2010; DiMicco, Hollenbach, Pandolfo, & Bender, 2007; Kim, Chang, Holland, & Pentland, 2008; Terken & Sturm, 2010). Speaking time feedback is operationalized as feedback on how much each group member talks during the group discussion. This individual speaking time feedback is visually displayed so that all group members can see it and it is provided in a dynamic way and in real-time, meaning that a device registers individual speaking time contributions while the discussion or interaction unfolds. Different devices have been used to provide this dynamic real-time individual speaking feedback to group

members. In one study, a table was developed that displayed the relative amount of time each group member spoke during the group discussion via lights fixed on a table. The more lights lit up for one group member, the more that group member spoke (Bachour et al., 2010). In another study, the distribution of speaking time among group members was displayed on the phones of each group member and the phones were visible on a table which the group members sat around for discussion (Kim et al., 2008). In yet another study, the relative speaking time of each group member was projected as a graph on a wall, visible for all group members (DiMicco et al., 2007). In all of these studies, individual speaking time was assessed via microphones and the relative distribution of speaking time among group members was displayed in a dynamic way in real time as the discussion unfolded and was visible to all group members.

Speaking time feedback has shown to result in emerging hierarchies that are flatter compared to when no such information is provided (DiMicco et al., 2007; Terken & Sturm, 2010). However, a major criticism that can be formulated against these findings is that many of the studies did not properly design their control condition. In most studies, the control condition consists of the absence of speaking time feedback (DiMicco et al., 2007; Kim et al., 2008; Terken & Sturm, 2010). Comparing individual speaking time feedback with no feedback at all might bring about the difference in the hierarchical structure not because the feedback was individual, but because there was feedback (compared to no feedback). Obtaining feedback induces cognitive load for the group members (compared to not obtaining any feedback at all) and maybe this is why a hierarchy was less likely to form. Or, the fact of being observed (Hawthorne effect; Roethlisberger & Dickson, 1939; Mayo, 1993) – necessary for being able to provide feedback - made norms of equality more salient to group members resulting in less steep hierarchies. So, a good control condition is one that also puts cognitive load on group members and also gives the participants the feeling of being observed.

Moreover, one can argue that obtaining feedback about speaking time can create demand effects, meaning making speaking time behavior salient among group members can interfere with hierarchy formation. Therefore, it is important that the feedback in the control condition is also about speaking time and not about say content as in some studies (e.g., Bachour et al., 2010). This is why we created a control condition that induces a certain amount of cognitive load, in which participants obtained feedback about their behavior so that they felt equally observed as in the individual feedback condition, and in which speaking time was equally salient: the condition of group speaking time feedback (explained in more detail in the Method section). We set out to replicate the findings on the link between speaking time feedback and flatter hierarchies with a proper control condition.

Hypothesis 1: *Individual speaking time feedback compared to group speaking time feedback results in flatter hierarchies.*

This replication is important for two reasons: 1) We aim to verify whether we find the same result as reported in the literature by using a proper control condition and 2) if we want to understand how the flatter hierarchies come about in the individual speaking time feedback condition, we need to first establish that they actually do become flatter.

How Does Speaking Time Feedback Affect Hierarchy

In leaderless groups, the time one spends talking indicates how much that person is regarded as the leader of the group discussion (Bass, 1949). Those who speak more during group interactions are perceived as being more dominant (Schmid Mast, 2002) and people who speak a lot during group interactions readily emerge as leaders (Ridgeway & Berger, 1986). The distribution of speaking time among group members is therefore a crucial objective indicator of one's hierarchical position within a group (Schmid Mast, 2001).

The emergence of a hierarchy happens in the early stages of a group interaction and remains relatively stable over time (Kalma, 1991). We ask whether the type of speaking time

feedback influences how people change their speaking contributions and thus how a hierarchy emerges? Hypothesis 1 states that hierarchy steepness will be lower with individual speaking time feedback than with group speaking time feedback. If that is true, it can happen in two ways: 1) the initial over-participators speak less and/or 2) the initial under-participators speak more. Therefore, Hypotheses 1.1 and 1.2 are simply a logical consequence of Hypothesis 1. We want to find out whether it is rather the over- or the under-participators (or both) who change their behavior based on the type of feedback they obtain.

Studies have shown that when a group is exposed to individual speaking time feedback, over-participators (those who normally dominate the conversation) tend to significantly reduce their comments (Bachour et al., 2010; DiMicco et al., 2007). In the present research, we therefore test whether receiving feedback on individual speaking time leads people who initially speak much and dominate the discussion, to diminish their speaking contribution during the group discussion and whether people who initially speak little, increase their speaking contribution.

> Hypothesis 1.1: *Speaking time feedback moderates the link between the initial (warm-up) level of participation (initial over-, average-, and under-participators) and speaking time behavior during the group discussion such that that initial over-participators decrease their participation in the group discussion when receiving individual speaking time feedback (as compared to control).*
>
> Hypothesis 1.2: *Speaking time feedback moderates the link between the initial (warm-up) level of participation (initial over-, average-, and under-participators) and speaking time behavior during the group discussion such that that initial under-participators increase their participation in the group discussion when receiving individual speaking time feedback (as compared to control).*

Effects of Task Interdependence on Hierarchy

Besides speaking time feedback, we also suspect that the nature of the task the group members are confronted with might influence the steepness of the emerging hierarchy (Anderson & Brown, 2010). Understanding how the emergence of a hierarchy is altered depending on the task is important because some tasks might afford a steeper hierarchy than others (Halvey, Chou, & Galinsky, 2011). No research has looked at how task type affects the emergence of social hierarchies.

From a functional approach, hierarchy enhances group effectiveness by reducing the inner group conflicts (Cantimur, Rink, & van der Vegt, 2016). The more the task is complex, the higher the chance of conflict. Therefore, groups faced with complex tasks might be in need for a control mechanism through which these conflicts can be resolved or prevented (Keltner, Van Kleef, Chen, & Kraus, 2008). Some research highlights the prominent role of a hierarchical structure for good group performance in highly interdependent groups faced with problem solving tasks (Ronay, Greenaway, Anicich, & Galinsky, 2012) or in the context of sport teams (Halevy, Chou, Galinsky, & Murnighan, 2012). This suggests that the formation of a hierarchy in an interdependent group can be a response to the need for coordination and management of potential conflicts. Interdependence reflects the extent to which members need to interact with one another in order to accomplish the group task successfully (Cropanzano & Ambrose, 2015). An example of a low interdependent task would be when the team members separately work on the sections of a standard report to be aggregated in a final document later. In contrast, a highly interdependent task requires the team members to incorporate each and every member's expertise, input, and feedback in order to pool it and to come to a consensus decision. According to Expectation States Theory (EST), hierarchies are formed based on performance expectations group members have for each other (Berger, Fisek, Norman, & Zelditch, 1977). Group members quickly estimate the relative quality of

each member's potential contribution to the task solution, called performance expectation. Members for whom the other group members have higher performance expectations are given more opportunities to contribute (e.g., are given more floor time) and therefore emerge as more influential in the group. In relatively homogenous groups, performance expectations can form around behavior that signals competence. For instance, dominant group members behave in such a way that they are perceived as more competent by fellow group members (Anderson & Kilduff, 2009). They, for instance, express their opinions more freely (Kalma, Visser, & Peeters, 1993) and assertively (Aries, Gold, & Weigel, 1983) which signals competence.

With respect to the interdependent task that requires discovering who possesses which information and to uncover the unshared information (which research shows is difficult for groups, Stasser & Titus, 2003) this means that group members need to ponder information contributed by different group members; they scrutinize the other group members for signs of task competency to form their performance expectations. Hierarchies will then build quickly around those expectations as predicted by EST. Groups might therefore form steeper hierarchies when confronted with a high interdependent task than when faced with a relatively low interdependent task. We posit that when groups work on complex tasks that require information sharing (e.g., interdependent tasks), steeper hierarchies will emerge than when groups work on less interdependent tasks.

> Hypothesis 2: *Steeper speaking time hierarchies emerge when group members are faced with a task that is relatively high in interdependence (compared to low in interdependence).*

Similar as with speaking time feedback, one can ask *how* steeper hierarchies emerge with regards to task interdependence. Again, Hypotheses 2.1 and 2.2. are logical deductions of Hypothesis 2. If task interdependence is responsible for steeper hierarchies to emerge compared to relatively lower task interdependence, then this needs to come from initial over-

participators to increase their participation and/or initial under-participators to decrease their participation. This way, a more pronounced hierarchy would form in the interdependent task condition. Therefore, we expect initial over-participators to increase their speaking time compared to when confronted with a low interdependent task.

Hypothesis 2.1: *Task interdependence moderates the link between the initial (warm-up) level of participation (initial over-, average-, and under-participators) and speaking time behavior during the group discussion such that that initial over-participators increase their participation in the group discussion when the task was relatively high in interdependence (as compared to relatively low in interdependence).*

In a similar but opposite vein, members who initially contribute less will be given less opportunities when the group works on an interdependent task. Hence:

Hypothesis 2.2: *Task interdependence moderates the link between the initial (warm-up) level of participation (initial over-, average-, and under-participators) and speaking time behavior during the group discussion such that that initial under-participators decrease their participation in the group discussion when the task was relatively high in interdependence (as compared to relatively low in interdependence).*

The Present Study

In the present article, we set out to test how providing group members with information about how much each group member talks during the group discussion (individual speaking time feedback) and how the level of interdependence among group members necessary for solving the group task, affect the steepness of the emergent speaking time hierarchy. Moreover, we set out to investigate how those changes in hierarchy steepness come about when looking at how initial over- and under-participators change their speaking

time contribution depending on the type of speaking time feedback obtained and depending on the type of task they are confronted with.

Method

Participants

A total of 405 participants (210 women, 195 men) interacted in same-gender groups of 3 people. Men and women form hierarchies differently in groups (Schmid Mast, 2001); males form more hierarchically structured groups compared to females in their first interactions. Moreover, in a meta-analysis, same gender groups showed stronger associations between speaking time behavior and dominance (Schmid Mast, 2002). Therefore, in our study we used same gender groups and controlled for the gender in our analysis.

The sample size was calculated based on a power analysis (Faul, Erdfelder, Lang, & Buchner, 2007) assuming an alpha level of .05, a power of .80 (conventional), and an assumed medium effect size (rho=0.2) which gave us a minimum of 30 observations per group. Participants were mostly Master and Bachelor students from the participant pool of a large university in the French speaking part of Switzerland. The proposal of the study was approved by the ethical committee of the university.

We did not assess participants' ethnicity, but because they come from a student participant pool we know the distribution of the ethnicity in the subject pool which typically consists of 75% White. On average, they were 21 years old ranging from 17 to 47. Participants were required to be native French speakers. Participants were remunerated for their participation with the equivalent of 20 US$ and a possible bonus equivalent to 5 US$.

Procedure

Participants were recruited for a study in which they needed to complete a task in a 3-person group. We selected 3-person groups, firstly for logistic reasons, to have a reasonably clear visualization of the speaking time feedback. Secondly, the task used in the study was

originally designed for groups of three people. We therefore decided not to change the task and proceeded with three-member groups. Groups were subject to a 2 (type of speaking time feedback: individual vs. group) by 2 (task interdependence: high vs. low) between-subjects design. Groups were randomly assigned to one of the 4 conditions.

Warm-up. Upon arrival, participants introduced themselves in a 5-min warm-up discussion. In this warm-up phase, we asked participants to briefly introduce themselves while becoming acquainted with receiving feedback on their speaking time. Also, the group was instructed on how to read the feedback and how to understand the meaning of the numbers on the wall. While the group members were interacting, we registered each member's speaking time.

Group discussion. Groups then received the instructions concerning the problem-solving task. Before working on the task as a group, group members were given 5 min to solve the task individually. In all conditions, groups were told that each person might have unique pieces of information and that we thus encourage them to exchange the information they have. This is the typical instruction for hidden profile tasks: groups are informed that the information given to the members might not be entirely identical (Lu, Yuan, & McLeod, 2012; Stasser & Titus, 1985; Stasser, Taylor, & Hanna, 1989). Participants then had 15 min to discuss their individual solutions in the group and reach a consensus decision. They were informed that if the group discovered the correct answer, they would be paid a bonus (equivalent of 5 US$ per group member). We left a timer on the table so that the groups were able to observe the remaining discussion time. One minute before the end of the discussion, the experimenter raised a sign showing that the group had to come to a final decision. When time was up, the experimenter stopped the recording and collected the group solution. Then, a questionnaire was distributed asking the participants to report how difficult they found the task (perceived task difficulty). Participants then learned about the correct solution for the task and were paid and thanked.

Microcone device. Speaking time was assessed with the Microcone (Innovative microphones for groups, 2017). The Microcone consists of an array of 6 microphones in one device placed in the middle of a table around which speakers sit. It can identify, register, and record the voices of up to 6 individuals. Applications that connect to Microcone provide real time information about the speaking behavior of the group and also each individual. The Microcone is able to distinguish the voices of the speakers in a group and calculate the percentages of the speaking time for the group as a whole and also for each group member individually. For example, it can provide feedback that the group communicated 90% of the time from the start of the discussion. The proportion of the dead-air time was taken into account to calculate this feedback. Microcone can also provide individualized feedback. For example, in a group of three when everybody speaks perfectly equally the Microcone will provide the following feedback: Speaker 1: 30%, Speaker 2: 30%, and Speaker 3: 30 %. Overlaps in the discussion were taken into account for all speakers involved. Therefore, the sum of the speaking percentages is generally over 100%. The dead-air time was neglected when calculating the relative contribution of group members.

Manipulation of speaking time feedback. During the group discussion, participants obtained real-time speaking time feedback. We used Microcone to provide such feedback while the discussions unfolded. Groups were randomly assigned to either obtain individual speaking time feedback or group speaking time feedback (= control). In both conditions, this information was projected on the walls of the room around the discussion group, so that all group members could easily see the feedback in real time (*Figure 3*). In the individual speaking time feedback condition, each member's contribution was recorded and the percentages of the members' speaking time relative to the entire time of the group discussion was calculated. Recordings of the speaking time started from the moment the first person started talking. Therefore, the group saw each member's percentage of speaking time in relation to the entire group's speaking time on the wall (*Figure 1).* In the control condition, feedback was given about the cumulative speaking time of the entire group relative to the duration of the discussion. That is, during the discussion, the group members saw how much they talked as a group (*Figure 2*). The researchers did not provide any information about the

interpretation of the feedback and about whether the members should change their behavior based on the feedback during the group discussion or not.

Manipulation of task interdependence. We used an existing personnel selection task in which an airline company is hiring a new pilot for long-distance flights (Schulz-Hardt, Brodbeck, Mojzisch, Kerschreiter, & Frey, 2006). The resumes of four applicants were presented to the participants, but only one applicant (applicant C) was the best candidate for the position based on the favorable ratio of positive compared to negative points in his resume (7 positive and 3 negative points, all other applicants had 4 positive and 6 negative points).

Task interdependence refers to the degree to which the completion of a task depends on the combined inputs of all group members (Wildschut, Lodewijkx, & Insko, 2001). The high interdependent task was presented as a hidden profile task: Each member of the group had unique pieces of information about the applicants that were hidden from the other members (i.e., different unique pieces of information are known to different group members) (Schulz-Hardt et al., 2006). More specifically, the pieces of information within the group were distributed equally in such a way that each member had 6 pieces of information about each candidate. Among these 6 pieces, two were uniquely held by each member and the other 4 pieces were shared, meaning that all group members had them. This made the value of each member's contribution equal because each member had an equal number of hidden pieces of information. Finding the best applicant required sharing information among all group members. Group members needed to discuss each piece of information so that by accumulating the pieces of information, the correct solution emerged. The best solution cannot be detected by relying only on the shared information (Schulz-Hardt et al., 2006). This type of task is relatively interdependent because the group members need to share all the information and especially the unique (unshared) information they hold.

We also created a low interdependent version of this task by simply providing each group member with all the information needed to solve the task. The members received 10

pieces information about each candidate and therefore had the full information for selecting the best candidate. This necessitated much less exchange of information during the group discussion. Because all group members were able to individually identify the best applicant given they all possessed the same pieces of information and they all possessed all pieces of information about all applicants, the sharing of information was not needed once the group members discovered that each group member had all the information needed at hand. In both conditions, members were required to memorize the information about the applicants prior to the group discussion and no notes or information sheets were allowed during the discussion. In the low-interdependent groups, the members had to recall more pieces of information (10) compared to the interdependent teams (6). In the relatively low interdependent task, the memorization was harder (more pieces of information), but in the relatively high interdependent task, remembering all the unique pieces of information of the other group members was taxing. We therefore assumed that the higher number of pieces of information to remember in one condition would be balanced out by the number of unique pieces of information that needed to be discovered and retained in the other condition. Nevertheless, we measured perceived task difficulty (see below).

Measures

Initial individual speaking time. In the *warm-up phase*, group members got acquainted with each other. They had to introduce themselves and provide information about themselves and their education. We recorded each participant's speaking time during the 5-min warm-up phase and classified them into over- and under-participators. Because initial individual speaking time was not normally distributed, we used the cut-off points proposed by Hayes (2017): over-participators (N = 69), average-participators (N = 271), and under-participators (N = 65) were based on the cut-off points of 16^{th} and 84^{th} percentile of the distribution suggested by Hayes. Over-participators were those whose speaking time was in

the first 16th percent of the speaking time distribution in the warm-up phase. Correspondingly, under-participators were those whose speaking time was in the last 16th percent of the speaking time distribution in the warm-up phase. The rest were categorized as average participators.

Individual speaking time. During group discussion, individual speaking time was assessed by the Microcone. Each member's speaking time contribution was calculated as the percentages of the member's speaking time relative to the entire time of the group discussion (M = 39.58; SD = 11.60).

Hierarchy. We used the standard deviation of the average speaking time across group members as a measure of emergent hierarchy (Schmid Mast, 2001, 2002). To illustrate, in a 3-person-group, if every group member talks about the same amount of time, the hierarchy would be flat (equal to zero). Larger numbers indicate a more hierarchical organization within a group.

Perceived task difficulty. After groups finalized their solution and finished the group discussion, participants reported how difficult they perceived the group task to be, using one item on a Likert scale ranging from 1 (*very easy*) to 7 (*very hard*). Higher scores indicated higher task difficulty (M = 3.13; SD = 1.83). We averaged the task difficulty scores among group members (Cronbach's Alpha = .63) to obtain a measure of perceived task difficulty on the group level.

Control variables. Past research has shown that in general it is necessary to statistically control for individual level attributes that enhance the effect of that attribute on the group level. As we measured social hierarchy using the speaking time variable, we needed to control for the variables that are related to individual speaking time. Therefore, we asked participants to report their fluency in the French language, level of education, age, and gender at the end of the study to be considered as control variables.

Results

Effects of Speaking Time Feedback and Task Interdependence on Hierarchy

To test Hypotheses 1 and 2, we calculated a 2 (speaking time feedback: individual vs. control) by 2 (task interdependence: high vs. low) ANOVA with speaking time hierarchy as the dependent variable. Confirming Hypothesis 1, results showed a significant speaking time feedback main effect, $F(1,126) = 4.32$, $p = .040$, showing that individual speaking time feedback resulted in a more egalitarian distribution of speaking time among group members ($M = 9.19$; $SD = 4.84$) compared to the control condition ($M = 11.12$; $SD = 5.20$). There was also a significant task interdependence main effect, $F(1,126) = 5.19$, $p = .025$, confirming Hypothesis 2 and showing that being faced with a high interdependent task resulted in a more hierarchical distribution of speaking time among group members ($M = 11.19$; $SD = 5.12$) than being faced with a relatively low interdependent task ($M = 9.09$; $SD = 4.88$). The interaction effect was not significant, $F(1,126) = 1.02$, $p = .314$.

Understanding How Individual Speaking Time Feedback Makes Hierarchies Flatter

We tested whether initial over-, average-, and under-participators differed with respect to how much they spoke during the group discussion with a linear contrast analysis showing that indeed, over-participators spoke more ($M = 48.65$; $SD = 10.00$) than average ($M = 39.17$; $SD = 10.38$), and under-participators ($M = 31.66$; $SD = 11.69$), $F_{\text{contrast}}(2, 134) = 35.13$, $p < 0.001$.

We predicted that speaking time feedback moderates the relation between initial participation and speaking time during the group discussion such that initial over-participators reduce their speaking time in the group discussion in the presence of individual speaking time feedback as compared to control (Hypothesis 1.1) and that under-participators increased it compared to control (Hypothesis 1.2). To test these hypotheses, we created three dummy variables: under-participator (is 1 when the speaking time is in the first 16^{th} percent and 0

otherwise), average-participator (is 1 when the speaking time is in the larger than 16th percent and smaller than 84th percentile and 0 otherwise), and over-participator (is 1 when the speaking time is in the last 16th percent and 0 otherwise). We cluster regressed group discussion individual participation on speaking time feedback (individual vs. control), and the dummy variables of the over- and under-participators (we excluded the average participators dummy variable to have it as the reference base), the interaction of the latter variables, as well as task interdependence, age, study level, participant gender, and French level as controls (we used these controls because they were either related to individual speaking time in the warm-up or in the group discussion;[12] see Table 1). Taking out the controls did not change the results.

There was a significant interaction effect between speaking time feedback and the initial individual participation of the over-participators (*Figure 4*) (β= -7.35, p < .01) confirming Hypothesis 2.1: initial over-participators significantly decreased their speaking time during group discussion in the individual speaking time feedback condition compared to the group speaking time feedback condition. Contrary to Hypothesis 2.2, there was no significant interaction effect between speaking time feedback and the initial individual participation of the under-participators (β= 2.27, p = .48).

Understanding How Task Interdependence Makes Hierarchies Steeper

To understand how task interdependence made hierarchies steeper (Hypothesis 2 was confirmed), we investigated whether for individuals with different initial individual participation (over- and under-participators), task interdependence influenced speaking time contribution during the group discussion differently. In other words, we tested whether initial over-participators increased their speaking time behavior in the presence of a relatively high interdependent task (Hypothesis 2.1) and whether under-participators decreased it (Hypothesis 2.2). We used the same dummy variables (under-participator, average-

participator, and over-participator) explained in the last section to test these hypotheses. We cluster regressed group discussion individual participation on task interdependence (high vs. low), the dummy variables of initial participation (over-and under-participator), the interaction of the latter variables, as well as speaking time feedback, age, study level, participant gender, and French level as controls (because they were either related to individual speaking time in the warm-up or in the group discussion – see footnote 1 and Table 2). Again, taking out the controls did not change the results.

We found a significant interaction effect between task interdependence and the initial individual participation level of the under-participators (β= -10.14, $p < .01$), but not for the over-participators (β= 3.29, $p = .23$) (*Figure 5*). Contrary to Hypothesis 4.1, over-participators faced with a high interdependent task (as compared to being faced with a low interdependent task) did not increase their speaking time during the group discussion. However, confirming Hypothesis 4.2, under-participators faced with high task interdependence (as opposed to low task interdependence) significantly decreased their speaking time during group discussion.

Discussion

The present study was aimed at understanding how differences in speaking time among group members in task groups emerge and how speaking time feedback influences how pronounced the emerging hierarchy is as well as how task type influences the emergence of such hierarchies. We showed that individual speaking time feedback leads to flatter hierarchies and that high task interdependence drives groups to form steeper hierarchies. Our results also show that individual speaking time feedback resulted in flatter hierarchies because over-participators decreased their speaking time contribution over time (and not under-participators increasing their speaking time). Also, task interdependence resulted in steeper

hierarchies because under-participators decreased their speaking time contribution over time (and not over-participators increasing their speaking time).

Effects of Speaking Time Feedback

We replicated existing empirical findings (Bachour et al., 2010; Kim, Chang, Holland, & Pentland, 2008) showing that individual speaking time feedback during a group discussion resulted in flatter emerging hierarchies. The reason why we wanted to replicate this result was to test whether it is robust with respect to a more appropriate control condition. Previous work has either used no feedback (Terken & Sturm, 2010; DiMicco et al., 2007; Kim et al., 2008) or feedback about other things (e.g., content) (Bachour et al., 2010) as the control condition. We now show that individual speaking time feedback resulting in flatter hierarchies is not due speaking time being salient or the fact of being observed (Hawthorne effect (Mayo, 1993; Roethlisberger & Dickson). It is really receiving *individual* speaking time feedback that most likely activates an equality norm that then becomes a self-fulfilling prophecy within the group (Ridgeway & Berger, 1986), making for flatter hierarchies. Seeing a visual display of the relative individual speaking time dynamically evolving in real time during a group discussion equalizes the individual contributions because those who have a tendency to speak a lot restrain from speaking (and not that those who have a tendency to speak less initially, speak more). We showed that over-participators significantly decreased in speaking time, whereas the increase in speaking time for the under-participators was not significant. In sum, egalitarian distribution of speaking time -when there is individual speaking time feedback present- is more driven by over-participators decreasing their participation than by under-participators increasing their participation.

Effects of Task Type

Our results also shows that when groups work on a relatively high interdependent task that necessitates information sharing, steeper hierarchies emerge compared to groups working

on relatively low interdependent tasks. We manipulated the type of task to test its effect on hierarchy formation within a group. The more a task necessitates interaction, exchange, and discussion (relatively high interdependent task), the more differences in individual speaking time contributions emerge. This finding confirms our hypothesis and is important because typically, scholars posit (Halevy, Chou, & Galinsky, 2011; Halevy et al., 2012) that interdependent tasks necessitate flatter hierarchies for good outcomes. We show, however, that in leaderless groups, group members organize themselves in a more hierarchical way when faced with an interdependent task. Whether this steeper hierarchy is conducive to better group performance is not clear and future research might want to tackle this question.

Such steeper hierarchies are formed because under-participators further decrease their participation but not because over-participators increasing theirs. When faced with high compared to low interdependent tasks, under-participators drive the effect more so than over-participators.

Interestingly, how steeper or flatter hierarchies emerge are not the same for speaking time feedback and for task type. Speaking time feedback has an effect on the over-participators (who will speak less so that the hierarchy will become flatter) whereas task interdependence influences under-participators more (in that they will speak even less so that the hierarchy will become steeper). Future research might want to systematically test which aspect the environment influences initial under- or over-performers more. When the equal opportunity norm is activated, this seems to stop over-participators, but when a hierarchy is called for, the under-participators hold back even more.

The high and low interdependent tasks were different in terms of difficulty.[13] The relatively low interdependent task was perceived to be easier than the high interdependent task. This is in line with findings of scholars suggesting that the preference for a hierarchical formation is most prevalent in cases of high task difficulty (Samuelson, 1991). When redoing

the ANOVA reported in the Result section, the task type effect becomes marginally significant ($F(1, 126) = 3.20$, $p = .076$). There was, however, no significant relation between perceived task difficulty and hierarchy ($r = .14$, $p = .11$). Future research on the effect of task structure on the emergence of a hierarchy might want to disentangle the effect of task structure and task difficulty. We deliberately opted for using exactly the same stimulus material, meaning the same task content and simply change the structure of the task. This was not possible without affecting task difficulty at the same time.

Moreover, our results explain the emergent hierarchy formation in same-gender groups and we do not know whether the results would be different for opposite-gender groups. We opted for all-male and all-female groups because the association between speaking time and dominance is stronger for same-gender groups compared to mixed-gender groups (Schmid Mast, 2002). Future research could try to replicate our finding in mixed-gender groups.

Conclusion

This study contributes to the understanding of the formation of hierarchies in small groups. The effect of the group task on hierarchy formation has not been the subject of much research attention so far. Results from this study show that for tasks that necessitate relatively high levels of information sharing (interdependent tasks), emerging hierarchies are steeper. Moreover, we confirm findings from the literature showing that individual speaking time feedback influences group dynamics towards a more egalitarian distribution of speaking time. We also show that when the group has a tendency to increase the hierarchy steepness because of task interdependence, then it is done via under-participators holding back even more (and not over-participators taking more of the floor) whereas when the group has a tendency to decrease the hierarchy steepness because of individual speaking time feedback, this is accomplished mostly by over-participators holding back (and not by under-participators

speaking up).

It is important to understand how hierarchies emerge and under which conditions they might become steeper or flatter because many of us regularly work in small, leaderless groups. Given that speaking time represents the opportunity to influence others, it is important to understand how it is distributed within a group. Providing speaking time feedback to work groups can be an efficient way of equalizing people's opportunities to contribute and is relatively easy to implement using different devices or simply applications on the phones of group members. For example, the concept of the flipped classroom is very popular nowadays. In flipped classroom models, students read the material and do the assignments at home and come to class to engage in a teacher-guided discussion around a topic (DeLozier & Rhodes, 2017). The flipped classroom model enhances learners' motivation and engagement (Chen, Lui, A. & Martinelli, 2017). However, if the class discussion is dominated by over-participators then the active learning processes will be disrupted. One way to avoid this could be to give students speaking time feedback. However, given that we showed that speaking time feedback influences over-participators, there might be a particular need for other methods to encourage under-participators to contribute.

Our results suggest that the way a task is structured can in itself create unequal amounts of speaking time. Therefore, when presenting a task to a team, one might want to initially provide everybody with equal amounts of task relevant information so that contributions might be more evenly distributed. For instance, in a meeting on the decision to launch a new product in an organization, it is best to provide the decision maker with a fully documented report on different aspects of the product development and share the pieces of information before the final decision making meeting. We would like to emphasize that, unequal distribution of speaking time is neither good nor bad, and that it is something that happens and we often are not aware of it.

References

Anderson, C., & Kilduff, G. J. (2009). Why do dominant personalities attain influence in face-to-face groups? The competence-signaling effects of trait dominance. *Journal of Personality and Social Psychology*, *96*(2), 491-503.

Anderson, C., & Brown, C. E. (2010). The functions and dysfunctions of hierarchy. *Research in Organizational Behavior, 30*, 55-89.

Anderson, C., John, O. P., & Keltner, D. (2012). The personal sense of power. *Journal of Personality, 80*(2), 313-344.

Aries, E. J., Gold, C., & Weigel, R. H. (1983). Dispositional and situational influences on dominance behavior in small groups. *Journal of Personality and Social Psychology, 44*, 779–786.

Bachour, K., Kaplan, F., & Dillenbourg, P. (2010). An interactive table for supporting participation balance in face-to-face collaborative learning. *IEEE Transactions on Learning Technologies, 3*(3), 203-213.

Bales, R. F. (1970). *Personality and interpersonal behavior*. New York: Holt, Rinehart and Winston.

Bales, R. F., Strodtbeck, F. L., Mills, T. M., & Roseborough, M. E. (1951). Channels of communication in small groups. *American Sociological Review, 16*(4), 461-468.

Bass, B. M. (1949). An analysis of the leaderless group discussion. *Journal of Applied Psychology, 33*(6), 527-533.

Berger, J., Fisek, M. H., Norman, R. Z., & Zelditch, M. (1977). *Status characteristics and interaction: An expectation states approach*. New York: Elsevier.

Cantimur, Y., Rink, F., & van der Vegt, G. S. (2016). When and why hierarchy steepness is related to team performance. *European Journal of Work and Organizational Psychology*, *25*(5), 658-673.

Chen, F., Lui, A. M., & Martinelli, S. M. (2017). A systematic review of the effectiveness of flipped classrooms in medical education. *Medical Education*, *51*(6), 585-597.

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103.

Cropanzano, R., & Ambrose, M. L. (Eds.). (2015). *The Oxford handbook of justice in the workplace*. Oxford University Press.

DeLozier, S. J., & Rhodes, M. G. (2017). Flipped classrooms: a review of key ideas and recommendations for practice. *Educational Psychology Review*, *29*(1), 141-151.

DiMicco, J. M., Hollenbach, K. J., Pandolfo, A., & Bender, W. (2007). The impact of increased awareness while face-to-face. *Human-Computer Interaction*, *22*(1), 47-96.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.

Fiske, S. T. (2010). Interpersonal stratification: status, power, and subordination. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 941–982). Hoboken, NJ: Wiley.

Frauendorfer, D., Schmid Mast, M., Sanchez-Cortes, D., & Gatica-Perez, D. (2015). Emergent power hierarchies and group performance. *International Journal of Psychology*, *50*(5), 392-396.

Gatica-Perez, D. (2009). Automatic nonverbal analysis of social interaction in small groups: A review. *Image and Vision Computing*, *27*(12), 1775-1787.

Halevy, N., Chou, E., & Galinsky, A. (2011). A functional model of hierarchy: Why, how, and when vertical differentiation enhances group performance. *Organizational Psychology Review, 1*(1), 32-52. doi: 10.1177/2041386610380991

Halevy, N., Chou, E. Y., Galinsky, A. D., & Murnighan, J. K. (2012). When hierarchy wins: Evidence from the national basketball association. *Social Psychological and Personality Science, 3*(4), 398-406. doi: 10.1177/2041386610380991

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford publications.

Innovative microphones for groups. (2017). Retrieved from http://www.dev-audio.com/support/downloads/.

Jayagopi, D. B., Hung, H., Yeo, C., & Gatica-Perez, D. (2009). Modeling dominance in group conversations using nonverbal activity cues. *IEEE Transactions on Audio, Speech, and Language Processing*, *17*(3), 501-513.

Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: theory and research.* Edina, MN: Interaction Book Company.

Kalma, A. (1991). Hierarchisation and dominance assessment at first glance. *European Journal of Social Psychology*, *21*(2), 165-181.

Kalma, A. P., Visser, L., & Peeters, A. (1993). Sociable and aggressive dominance: Personality differences in leadership style?. *The Leadership Quarterly*, *4*(1), 45-64.

Kim, T., Chang, A., Holland, L., & Pentland, A. S. (2008). Meeting mediator: enhancing group collaborationusing sociometric feedback. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, USA*, 457-466.

Kulyk O., Wang J., & Terken J. (2006). Real-time feedback on nonverbal behavior to enhance social dynamics in small group meetings. In S. Renals, S. Bengio (Eds.), *Machine*

Learning for Multimodal Interaction (MLMI 2005): Vol 3869. *Lecture Notes in Computer Science* (pp. 150-161). Berlin, Heidelberg: Springer.

Lu, L., Yuan, Y. C., & McLeod, P. L. (2012). Twenty-five years of hidden profiles in group decision making: A meta-analysis. *Personality and Social Psychology Review, 16*(1), 54-75.

Mayo, E. 1933. *Human problems of an industrial civilization*. New York: Macmillan.

Ridgeway, C. L., & Berger, J. (1986). Expectations, legitimation, and dominance behavior in task groups. *American Sociological Review, 51*, 603–617. doi: 10.2307/2095487

Roethlisberger, F. J., & Dickson., W. J. (1939). *Management and the worker.* Cambridge, Mass: Harvard University Press.

Ronay, R., Greenaway, K., Anicich, E. M., & Galinsky, A. D. (2012). The path to glory is paved with hierarchy: When hierarchical differentiation increases group effectiveness. *Psychological Science, 23*(6), 669-677.

Rulke, D. L., & Galaskiewicz, J. (2000). Distribution of knowledge, group network structure, and group performance. *Management Science, 46*(5), 612-625.

Sanchez-Cortes, D., Aran, O., Mast, M. S., & Gatica-Perez, D. (2011). A nonverbal behavior approach to identify emergent leaders in small groups. *IEEE Transactions on Multimedia, 14*(3), 816-832.

Schmid Mast, M. (2001). Gender differences and similarities in dominance hierarchies in same-gender groups based on speaking time. *Sex Roles, 44*(9-10), 537-556.

Schmid Mast, M. (2002). Dominance as expressed and inferred through speaking time. *Human Communication Research, 28*(3), 420-450.

Schulz-Hardt, S., Brodbeck, F. C., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology, 91*(6), 1080-1093.

Stasser, G., & Titus, W. (1985). Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*(6), 1467-1478.

Stasser, G., Taylor, L. A., & Hanna, C. (1989). Information sampling in structured and unstructured discussions of three- and six-person groups. *Journal of Personality and Social Psychology*, *57*(1), 67-78.

Terken, J., & Sturm, J. (2010). Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing, 14*(8), 703-714.

Wildschut, T., Lodewijkx, H. F., & Insko, C. A. (2001). Toward a reconciliation of diverging perspectives on interindividual-intergroup discontinuity: The role of procedural interdependence. *Journal of Experimental Social Psychology*, *37*(4), 273-285.

Footnotes

[12] We cluster regressed the speaking time behavior of members (in the warm-up phase and in the group discussion) on individual characteristics and found that individuals who had a higher education level (β = 3.06, p = .047), spoke more in the warm-up phase and that participants who were female (β = 1.97, p = .003), who were older (β = 0.45, p = .05), and who had better French skills (β = 4.01, p = .001) spoke more during the group discussion. Therefore, in our regression analyses we controlled for participant age, gender, level of French, and educational level.

[13] We tested whether perceived task difficulty differed between the low and high interdependent task. Results showed a significant effect of task type on perceived difficulty, $t(128) = 5.99$, $p < 0.001$, such that groups with high interdependent task perceived their task as more difficult (M = 3.78; SD = 1.35) than those who faced with the low interdependent task (M = 2.49; SD = 1.10).

Table 1

Regression Estimates Predicting Individual Speaking Time from the Interaction between Speaking Time Feedback and Initial Participation Level, Task Interdependence, and the Control Variables

Model	(1)	(2)	(3)
Age	.46**(2.10)	.46**(2.14)	.45**(2.11)
Female	1.05*(1.84)	1.09**(1.89)	1.00*(1.71)
French level	3.36**(3.38)	3.41**(3.45)	3.41**(3.44)
Study level	-1.41(0.87)	-1.77(1.10)	-1.36(0.84)
Task interdependence	-.47(0.86)	-.60(1.08)	-.38(0.66)
Speaking time feedback	.51(0.96)	1.78**(2.48)	.15(0.20)
Over-participants	9.12**(7.40)	13.08**(7.31)	9.12**(7.41)
Under-participants	-7.67**(5.09)	-7.64**(5.96)	-8.79**(4.16)
Over-participants×Sp. Feedback		-7.35**(2.86)	
Under-participants×Sp. Feedback			2.27(0.71)
Constant	20.56**(4.10)	20.03**(4.01)	20.62**(4.09)
R^2	0.212	0.226	0.213

Note. $N = 405$; clustered robust *t*-statistics in parentheses; gender dummy coded as 2 = female and 1 = male; ** $p < .05$ * $p < .1$; French level coded as 1 = fluent, 2 = bilingual, 3 = mother tongue; Study level coded as 1 = Bachelor's and 2 = Master's level

Table 2

Regression Estimates Predicting Individual Speaking Time from the Interaction between Task Interdependence and Initial Participation Level, Speaking Time Feedback, and the Control Variables

Model	(1)	(2)	(3)
Age	.46**(2.10)	.44**(2.02)	.41*(1.90)
Gender	1.05*(1.84)	.97*(1.68)	.68*(1.11)
French level	3.36**(3.38)	3.36**(3.39)	3.20**(3.23)
Study level	-1.41(0.87)	-1.41(0.87)	-1.43(0.88)
Task interdependence	-.47(0.86)	-1.03(1.40)	1.17(1.57)
Speaking time feedback	.51(0.96)	.56(1.05)	.08(0.13)
Over-participants	9.12**(7.40)	7.34**(3.53)	9.09**(7.41)
Under-participants	-7.67**(5.09)	-7.66**(5.09)	-2.34(1.01)
Over-participants×task interdependence		3.29(1.22)	
Under-participants×task interdependence			-10.14**(3.29)
Constant	20.56**(4.10)	21.25**(4.11)	21.90**(4.33)
R^2	0.212	0.215	0.237

Note. N = 405; clustered robust t-statistics in parentheses; gender dummy coded as 2 = female and 1 = male; ** $p < .05$ * $p < .1$; French level coded as 1 = fluent, 2 = bilingual, 3 = mother tongue; Study level coded as 1 = Bachelor's and 2 = Master's level

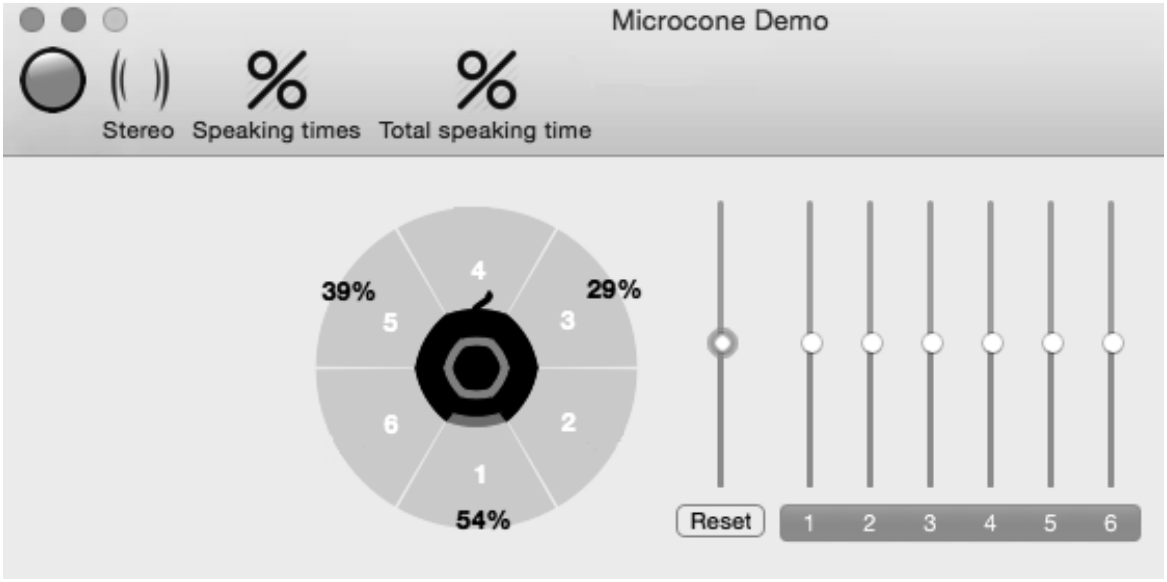


Figure 1. Display of the individual speaking time feedback. Participants knew which number in the circle represented which participant and could observe the percentages of speaking time change in real time for all 3 group members. They were instructed to ignore the other information on the display (which remained unchanged during the group discussion). In this example, one group member spoke 54% of the time while the other group members spoke 39% and 29% respectively. The cumulative percentage across all participants often rose over 100% because of overlapping speech of the group members.

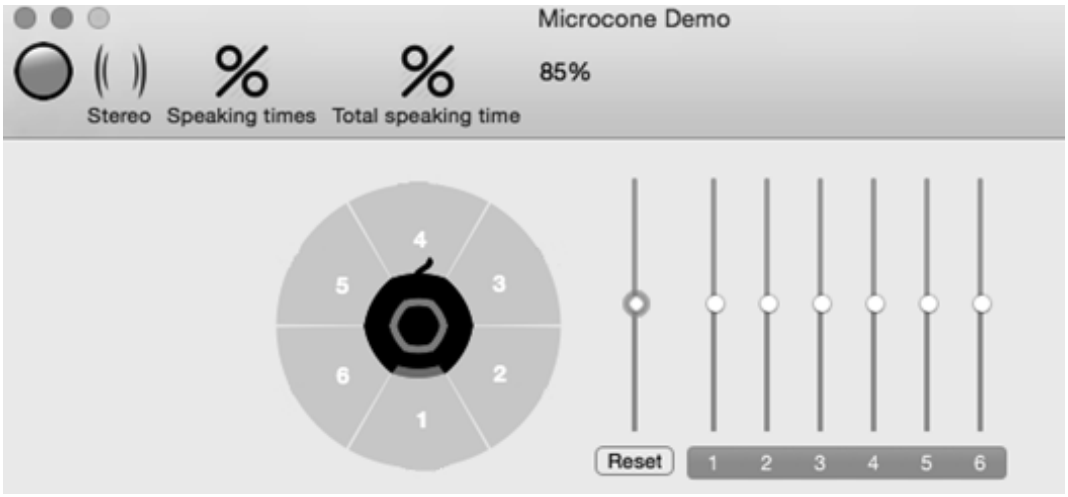


Figure 2. Display of the group speaking time feedback. Participants were instructed to focus on the percentage written next to "total speaking time" – here 85% - and they were informed about what that number meant. They were also instructed to ignore the other information on the display (which remained unchanged during the group discussion).

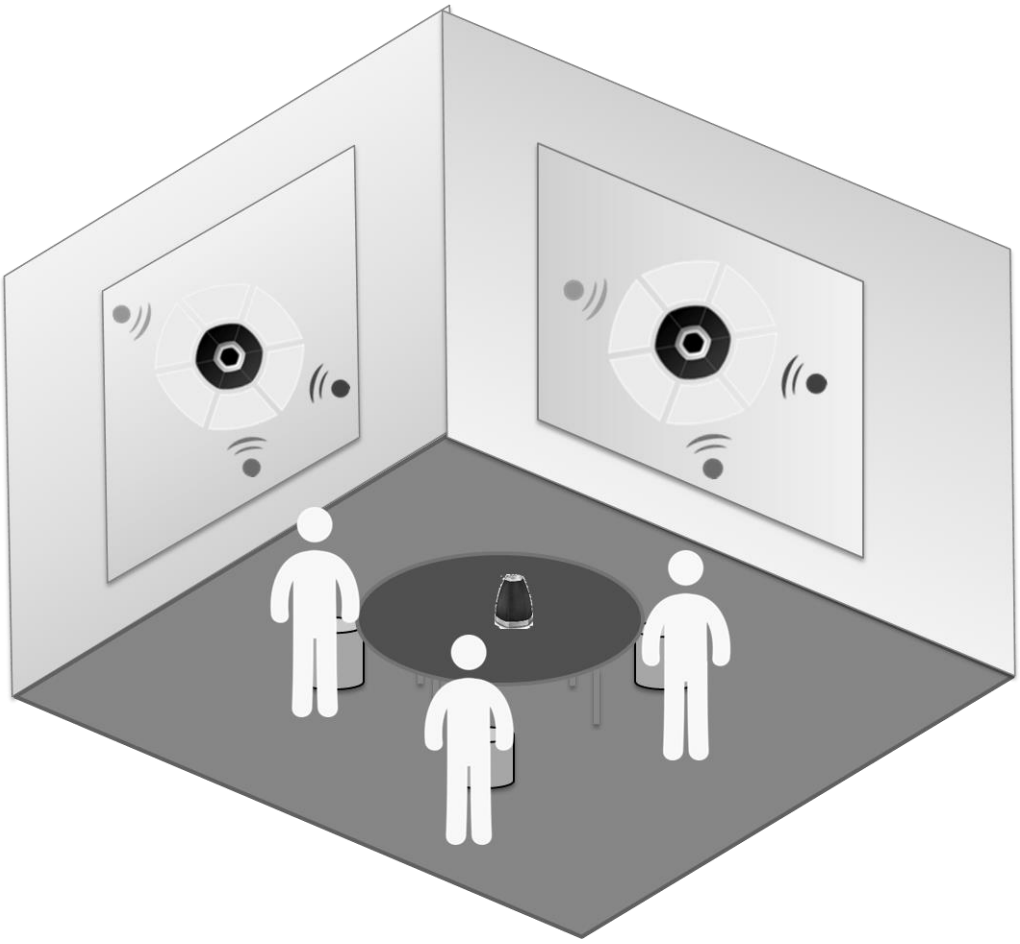


Figure 3. Display of the speaking time feedback in the room where group discussion occurred. Participants were instructed to find the feedback either according to their speaking time or according to group speaking time and they were informed about what that number meant.

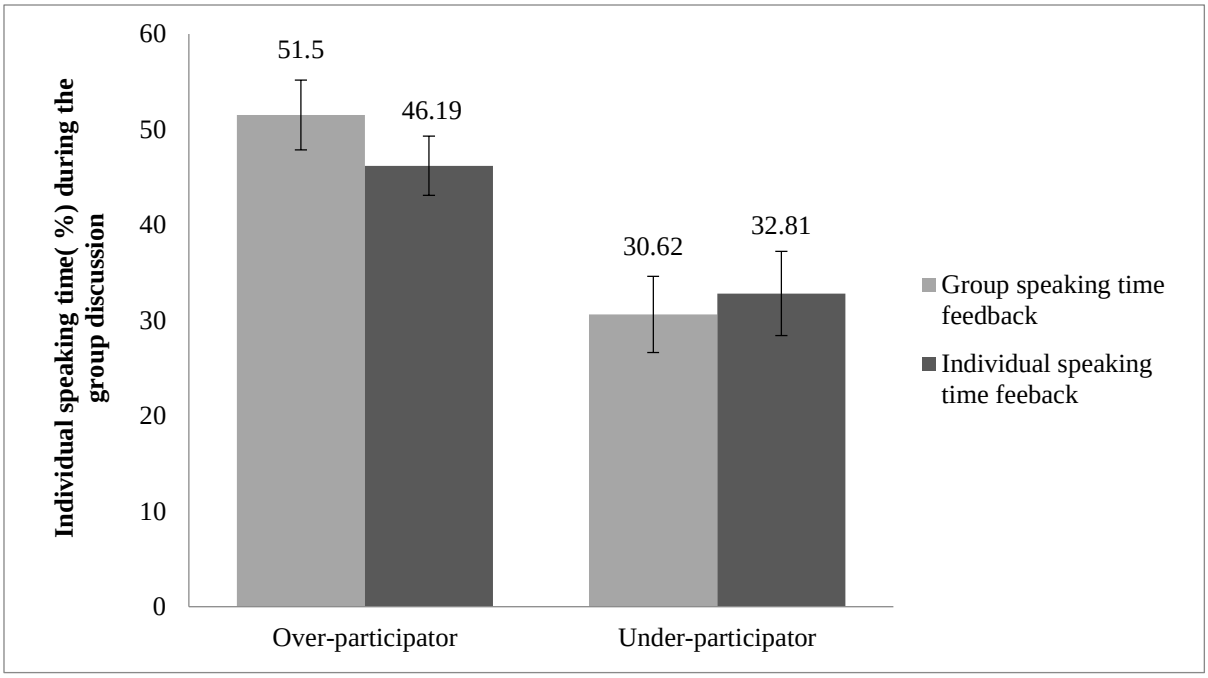


Figure 4. Interaction effect of speaking time feedback and initial individual participation

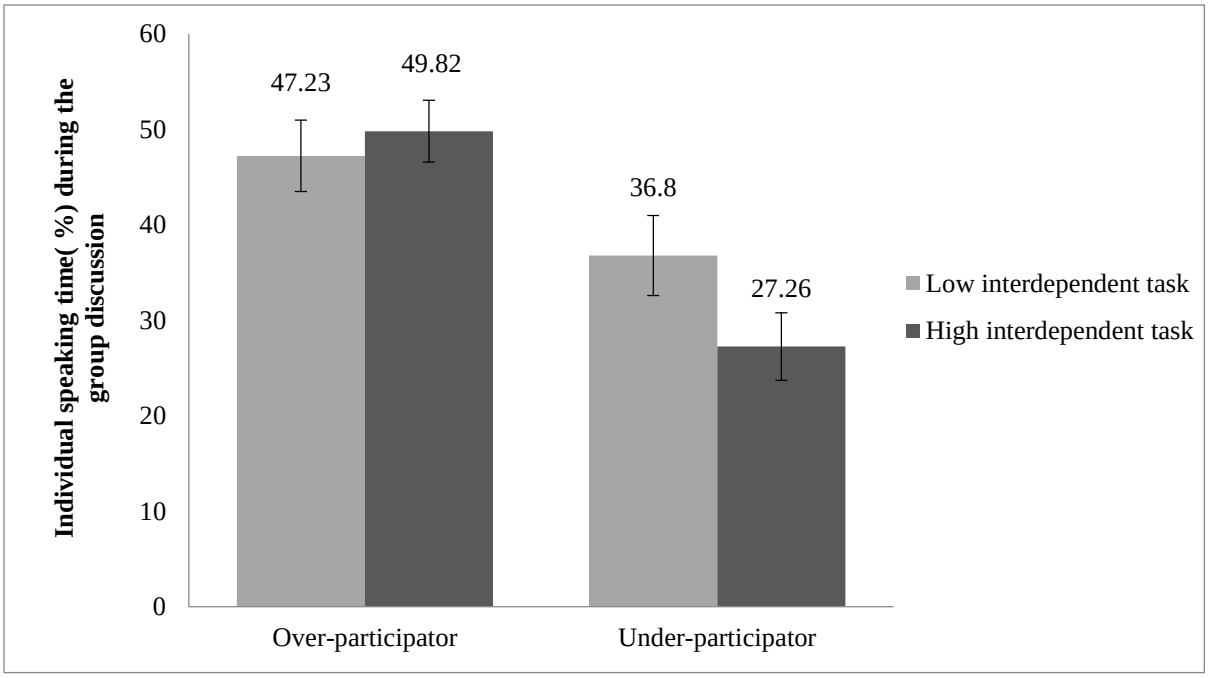


Figure 5. Interaction of task interdependence and initial individual participation